



HAL
open science

Systemic Financial Risk Analysis: from the Sectoral to the Aggregate Perspective

Arthur Stalla-Bourdillon

► **To cite this version:**

Arthur Stalla-Bourdillon. Systemic Financial Risk Analysis: from the Sectoral to the Aggregate Perspective. Business administration. Université Paris sciences et lettres, 2022. English. NNT: 2022UPSLD048 . tel-04098775

HAL Id: tel-04098775

<https://theses.hal.science/tel-04098775>

Submitted on 16 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris-Dauphine

**Analyse des Risques Financiers Systémiques: de la
Perspective Sectorielle à la Perspective Agrégée**

Soutenue par

**Arthur STALLA-
BOURDILLON**

Le 24 Novembre 2022

Ecole doctorale n° ED 543

Ecole doctorale SDOSE

Spécialité

Sciences de Gestion

Composition du jury :

Christian, BROWNLEES Professeur Associé, Universitat Pompeu Fabra	<i>Rapporteur</i>
Éric, JONDEAU Professeur, HEC Lausanne	<i>Rapporteur</i>
Frédérique, BEC Professeur, CY Cergy Paris Université	<i>Examineur</i>
Marie, BRIERE Professeur Associé, Université Paris- Dauphine, Amundi	<i>Examineur</i>
Matthieu, BUSSIÈRE Directeur, Banque de France	<i>Examineur</i>
Serge, DAROLLES Professeur, Université Paris- Dauphine	<i>Président du Jury</i>
Gaëlle, LE FOL Professeur, Université Paris- Dauphine	<i>Directrice de thèse</i>

UNIVERSITÉ PARIS-DAUPHINE

GFR Finance

Systemic Financial Risk Analysis: from the Sectoral to the Aggregate Perspective

Arthur STALLA-BOURDILLON

PhD Thesis

November, 24th 2022

Thesis supervisor :

Gaëlle LE FOL Université Paris-Dauphine

Jury president :

Serge DAROLLES Université Paris-Dauphine

Jury members :

Christian BROWNLEES Universitat Pompeu Fabra (*rapporteur*)

Eric JONDEAU HEC Lausanne (*rapporteur*)

Frédérique BEC CY Cergy Paris Université

Marie BRIERE Université Paris-Dauphine, Amundi

Matthieu BUSSIÈRE Banque de France

ACKNOWLEDGEMENT

The academic time and the institutional time have different rhythms. Sometimes they are synchronized and harmonious, sometimes they simply are not. A bit like Giovanni Drogo in *The Tartar Steppes*, and with all the caution that should be taken in such comparisons, I tried to juggle during this PhD with two temporalities, the short term one and the long term one, navigating between quantitative policy tasks and long-horizon research objectives. Being at the same time a PhD student at Dauphine and an economist at Banque de France was an unbelievable opportunity for me, both intensive and extremely rewarding. Playing on both sides, as one friend framed it, can be difficult, but I like to think that I tried to give my best at it.

At the end of this journey, I must say that I owe a particular intellectual debt to my supervisor, Gaëlle le Fol, who was constantly present throughout the thesis. Finally leaving my student life behind me, I can now assert from experience that a young researcher needs advice not only on the substance of his projects, but also on the implicit rules that govern the academic world. Gaëlle provided both. I cannot stress enough how her role was essential in the completion of this thesis.

I am also indebted to my hierarchy at Banque de France, as they enabled me to pursue this project from the beginning. I would like to thank, among others, the people who opened for me the PhD door while I was still uncertain about its implications, Laurent Clerc and Julien Idier, as well as my line managers who supported me in the last few years, Pierre-Yves Gauthier, Natacha Isslame-Rocher, Caroline Jardet, Sophie Haincourt, Nicolas Même and Nicolas Chatelais. Nicolas also played on two sides here in some sense, as he acted both as my *chef de pôle* and as my co-author for the second chapter of my thesis. I definitely think he managed to juggle with the two roles though.

My gratitude goes also to my two other co-authors Menzie Chinn and Lukas Boeckelmann. I have to say that having Lukas on my sides for my first research paper definitely supplied me with the exact amount of research insights, new ideas, work motivation and humourous ironical detachment that I needed at that time. I am also grateful to the other economists who helped me at some point along the way, among which Cyril Couaillier, Hugues Dastarac, Albert Dorval, Tristan Jourde, Marko Novakovic, Dilyara Salakhova, Maéva Silvestrini and Valerio Scalone.

I also thank the jury members of my PhD thesis, Frédérique Bec and Matthieu Bussière for hav-

ing agreed to join my dissertation defence, Marie Brière and Serge Darolles for their insightful comments as well as for their presence at our yearly meetings within the thesis committee, and finally my *rapporteurs*, Christian Brownlees and Éric Jondeau, who gave me extremely valuable feedbacks and advice. I am deeply honored that they all agreed to attend my PhD defence.

Eventually, I would like to thank my friends and my family for their support during the last four years. Special thanks are of course dedicated to my parents, Aline and Bertrand, for having welcomed the idea of a second doctor in the family (unrelated, however, to medical studies), to my older brothers, Julien and Martin, and to my twin, Hugo, who provided me with the unmatched complicity that twinning can beget. A final attention is specifically devoted to Marie and to her ceaseless affectionate backing, she knows the profound emotional liability that I owe her.

GENERAL INTRODUCTION /
INTRODUCTION GÉNÉRALE

General introduction

The role of financial markets from a central bank point of view

Financial markets play different roles in our contemporary economy. First, they are mainly designed to be an efficient tool for capital allocation, providing at the same time financial resources to companies and remunerations for capital-holders ([Wurgler, 2000](#)).

Second, this allocation operates mainly through the price discovery process. As market agents confront their opinions on financial markets, an equilibrium price will emerge from the aggregation of various bid and ask-offers, thus providing the agents with a signal of where capital is most needed. As a result, and in line with Hayek market theory of price aggregation ([Hirshleifer, 1973](#)), financial markets act as a powerful informational tool to reduce the multidimensional investor opinions into a single metric: the level of asset prices.

Third, financial markets pool together economic agents that may have diverging consumption needs depending on the potential future states of the world. As financial assets may yield returns in some of these states only, agents can trade them with one another so as to smooth their consumption paths. Risk sharing constitutes thus an essential aspect of financial markets.

It has yet been identified that, despite their usefulness, financial markets can also be seen as source of risk for the economy for two main reasons. On one hand, financial markets are characterized by booms and busts processes whereby, fueled by so-called market exuberance ([Shiller, 2015](#)), asset prices slowly increase and depart from their fundamental values before suddenly decreasing. These swift price variations are not painless for the real economy. During the build-up of the financial bubbles, as price increases are not related to fundamental factors, they may induce capital misallocation ([Miao et al., 2015](#)). When asset prices burst, the sharp repricing can lead to cascading defaults of financial institutions, and then have broader macroeconomic impacts. These bankruptcies may be caused by various non-linearities and threshold effects that are present in the financial sectors, such as unanticipated fund redemptions or sudden margin calls ([Malkiel, 2010](#)).

On the other hand, as stated before, financial networks may be appropriate for risk-sharing purposes between agents, but they are also central in the transmission of shocks. In other words,

financial markets bear a risk for the economy not only because of the boom/bust processes underlined above, but also because they can play a role in propagating shocks originating in one market to another market. These contagion patterns take place for various reasons, for example because different institutions exhibit cross-linkages between themselves (Brunetti et al., 2019), because they hold similar asset portfolios (Greenwood et al., 2015), or because two assets have related characteristics in the eyes of investors.

As a result, financial markets and central banks share a deeply intertwined relationship, both when we consider the functions assigned to the financial markets or the risks they are associated with. Central bank mandates may differ over the world or across time, but most of the time they involve price stability (Smets, 2003), sometimes combined with an unemployment target, as it is notoriously the case for the FED (Thornton et al., 2012), and, since the Great Financial Crisis, they often include financial stability objectives (Kim et al., 2018). Consequently, the relationship between central banks' actions and financial markets is multiform and depends on the central bank target that we are considering. We shed light here on the three main dimensions, in our view, along which financial markets and central banks interact with each other.

Firstly, the principal tool used by central banks to maintain price stability is the policy rate, sometimes combined with unconventional monetary measures such as asset purchase programs or forward guidance policies. In other words, in their fundamental functioning, central banks have to operate with financial markets, in this precise case with the bond markets, to meet their targets.

Secondly, within their financial stability mandate, central banks have to monitor the buildup of risks in financial markets. This involves, *inter alia*, tracking asset price valuations to gauge if a bubble is forming (Geis et al., 2018), assessing financial interconnections to evaluate the potential spillovers from one market to another (Alter and Beyer, 2014) or overseeing the resilience of market participants to determine if they can cope with significantly negative shocks.

Thirdly, financial markets have an informational utility for central banks given their ability, as stated above, to aggregate investors' opinions. This use can take several different forms. Central banks for example rely on inflation-linked swaps (ILS) or on the pricing of inflation-protected securities to assess market participants' expectations of short-term or long-term future inflation (Bauer et al., 2015). In this case, central banks exploit these market-based metrics as an indicator

of the good transmission of their monetary policy, reflected here by the potential good anchoring of investors' inflation expectations. But central banks may also rely on financial markets variables as a quantitative input when they form their forecasts regarding the future states of the business cycle. Indeed, many market-based indicators have proven to yield predictive content for economic activity forecasting, such as the term spread (Chinn and Kucko, 2015) or the dividend yield (Lan et al., 2020) and are therefore part of the central bank toolkit for business cycle monitoring.

This PhD thesis aims at shedding a new light on the use of financial market data from the central banks' perspective, essentially with respect to their financial stability mandates and to their need to gauge future economic activity. The main message throughout the thesis chapters is that a considerable amount of information is obscured if we only rely on aggregate variables (as is the case for a significant part of the macro-financial literature). As we will try to make it convincing in the thesis, the use of micro/sectoral financial data can prove useful for various purposes, including for answering research questions from a macro-perspective.

Historically, a lot of attention in the macro-financial literature has been devoted to aggregate financial data

Regarding the previous subjects discussed above (financial bubbles, market informativeness, financial contagion) the literature has historically relied on aggregate financial market variables, most of the time with a focus on the United States. Macro-variables are seen as very convenient for various reasons. They are indeed easily available and cover, compared to micro-datasets, relatively long time periods. Additionally, especially when it comes to financial market variables, micro-datasets may turn out to be relatively noisy. Fama and French (1988) and Polk et al. (2006) for instance stressed how popular metrics like the dividend yield or the price earnings ratio can be unreliable at the stock-level, especially due to the difficulty to obtain trustful estimates of book data (earnings, asset size etc.). On top of that, on the methodological side, some of the econometric models used to deal with these questions, such as Vector Auto-Regressive models (VARs), may put an upper limit on the amount of variables that could be considered in the estimation. In that regard, aggregate variables appeared as natural candidates to study these questions within sparse models.

With respect to the study of asset valuations, and of potential departures from fundamental values in financial bubble processes, most of the foundational papers relied on aggregate data. [Shiller et al. \(1983\)](#) for example underlined that, after building ex-post estimates of the fundamental value of the Standard and Poor's Composite Stock Price Index or of the Dow Jones Industrial Average, macro-equity prices were too volatile to be only influenced by fundamental factors. In a similar fashion, subsequent studies also aimed at assessing the relationship between asset prices and their fundamental values, but again most of the time with index-level data. This is the case, among others, of [Lee et al. \(1999\)](#) who estimate proxies of the Dow Jones fundamental value with cointegration techniques. Another option to gauge the level of market price-efficiency is to evaluate to what extent stock prices are driven by future cash flow news. This approach has been notably used by [Campbell \(1991\)](#) and [Campbell and Ammer \(1993\)](#), also on index-level data.

The informational role of the financial market prices with regards to future economic activity has been extensively studied in the literature too. Here again, macro-level data have been the main tool used to investigate on this issue. On the bond market side, a long stream of papers has reported the good performance of the term spread, the difference between short- and long-dated government security rates, to predict future recessions ([Stock and Watson, 1989](#), [Estrella and Hardouvelis, 1991](#), for the US, a result extended by [Chinn and Kucko, 2015](#), on other advanced economies). Some index-level equity variables turned out to be helpful for business cycle forecasting as well, such as stock returns ([Binswanger, 2000](#), [Ólan Henry et al., 2004](#), [Croux and Reusens, 2013](#), [McMillan, 2021](#)), stock price growth ([Chen and Ranciere, 2019](#)) or the dividend yield ([Lan et al., 2020](#)).

Eventually, as underlined above, in the contagion literature seminal papers often relied on aggregate country-level data given the dimensionality problem that can occur if one wants to include sectoral variables in the framework. This is notably the case for papers relying on Structural VARs (SVARs) whose identification strategies often only allow for a limited amount of variables. As a results, original studies in this field privileged market-level data over sectoral time series, as in [Diebold and Yilmaz \(2009\)](#) for stock returns and stock volatility or in [Diebold and Yilmaz \(2012\)](#) for an extension of the framework towards the bond markets and the foreign exchange markets.

Relying on financial micro/sectoral data can turn profitable to answer macroeconomic research questions

Although potentially cumbersome in terms of data availability or of dimensionality, focusing on a micro/sectoral-perspective on these issues may prove useful for three main reasons.

First financial variables are potentially affected by different factors whether we consider them at the micro-level or at the macro-level. If we focus for example on the equity markets, one could show, following the present value formula of [Campbell and Shiller \(1988\)](#), that index-level dividend yield (x_t) can be decomposed likewise:

$$x_t = \frac{\kappa}{(1 - \rho)} + \sum_{j=1} \rho^{j-1} E_t[r_{t+j} - \Delta cf_{t+j}] \quad (1)$$

Where $E_t[r_{t+j}]$ represents expected returns and $E_t[\Delta cf_{t+j}]$ expected cash flows (κ and ρ are constant parameters). In other words, this identity formula states that current equity valuations ratios essentially depend on two factors: future cash flow growth or future expected returns/discount rates. The same equation can be written for the company or the sector i :

$$x_{i,t} = \frac{\kappa_i}{(1 - \rho_i)} + \sum_{j=1} \rho_i^{j-1} E_t[r_{i,t+j} - \Delta cf_{i,t+j}] \quad (2)$$

These two identities can help us formalizing why equity price behavior may differ between the micro- and the macro-stage. Samuelson initially had the intuition that stock-level returns were mainly driven by news about firms' profitability, and that these latter were averaged out in the aggregate so that index-level returns were on the reverse mostly affected by variations in discount rates ([Jung and Shiller, 2005](#)).

Such phenomenon of contrasting behaviors between micro-returns and macro-returns has been empirically documented. [Sadka and Sadka \(2009\)](#) for example reported that the positive relationship between earning growth and returns at the micro-level turns negative at the macro-level. [Kothari et al. \(2006\)](#) underlined analogous findings between earning surprises and contemporaneous returns, according to which stock-level returns are positively correlated to idiosyncratic

earning news, whereas this positive correlation vanishes at the index-level. In a similar fashion, [Hirshleifer et al. \(2009\)](#) stress that elevated accruals predict negative future returns at the stock-level, but null or positive future returns at the index-level. All these studies share a similar reasoning: cash flows are mainly idiosyncratic, whereas discount rates are common across firms. As such, drivers of stock returns may differ greatly depending on the scale we are considering. Allegedly, these drivers are more associated with expectations about future profitability at the micro-level, but, due to diversification effects, depend more on discount rate factors (such as investor risk aversion) at the macro-level.

This contrast between micro- and macro-equity behaviors can be interesting for researchers. It is thus possible to take profit from the cross-sectional heterogeneity in the cash flow factors to extract a metric of investors' discount rates. While trying to predict future aggregate returns, [Kelly and Pruitt \(2013\)](#) for example estimate a factor model based on a cross-section of sectoral book-to-market ratios. Relying on sectoral data enables them to filter out the cash flow components in equity prices and therefore to get a more precise estimates of aggregate discount rates (with the latter turning out to perform very well in index-return forecasting).

Second, micro/sectoral data may be useful not only because micro-financial variables behave differently than macro-variables, but also because researchers can benefit from the heterogeneity of micro-asset price behaviors among themselves. In our view, this heterogeneity can serve various purposes.

As sectoral future cash flows respond differently to macroeconomic shocks, evaluating how micro-stock prices react at specific dates may be profitable to identify these shocks. [Venditti and Veronese \(2020\)](#) for example rely on airline equity prices to identify oil supply shocks, given that oil price increases may affect relatively more airline stock returns than index-returns.

In a similar fashion, sectoral equity future cash flows, and thus sectoral equity variables, respond differently to changes in future economic activity. As such, overweighting some sectors compared to other may prove fruitful in forecasting analysis. [Andersson and Agostino \(2008\)](#) showed for instance how Oil & Gas and Basic Materials industries exhibited stock returns that were more correlated with future Euro Area GDP than other sectors.

Additionally, the heterogeneity in micro-equity prices matters beyond the two cash-flow channels outlined above. A more subtle argument, outlined again by [Kelly and Pruitt \(2013\)](#), is that

growth sectors, typically composed of software companies, have future cash flows that are further away in time compared to value sectors, which exhibit more regular cash flow streams. As a result, growth sectors display a higher equity duration in comparison with value sectors, meaning that their equity valuation ratios are more sensitive to discount rate changes. This sensitivity difference can in turn be advantageous in order to purge sectoral equity valuation ratios from their discount rate components so as to better forecast future dividend growth (as in [Kelly and Pruitt, 2013](#)).

Third, sectoral time series prove helpful in contagion-analysis as they can help disentangle the causality chain in spillover processes. As an example, various papers in the literature focused on the dynamics of European sovereign interest rates or Credit Default Swaps (CDS) during the European debt crisis by looking only at country-level sovereign variables ([Ehrmann and Fratzscher, 2017](#), [De Santis and Zimic, 2018](#)). However, a substantial part of sovereign CDS dynamics in that period originated from the so-called sovereign bank-nexuses, where sovereign and bank credit risks fed each other. These nexuses can occur for various reasons. Among other, because bank risk can affect domestic sovereign risk through the “bailout channel”, that is explicit or implicit public guarantees, in case of distress of the banking sectors ([Alter and Schüler, 2012](#)), or because banks hold a significant amount of domestic sovereign bonds in their balance sheets (“balance sheet channel”, [Angeloni and Wolff, 2012](#), [Buch et al., 2016](#)). Therefore, omitting sectoral variables in the form of bank yields or bank CDS can result in a biased picture of the contagion processes, and, in turn, in a badly designed policy response.

We tried, through the lens of the three arguments outlined above, to shed light on the importance of micro/sectoral financial market variables, even when dealing with macro-financial issues. The three chapters of this PhD thesis aim at leveraging on this aspect and underlining the advantage of micro-data in three areas: stock return predictability, the use of equity variables for macro-forecasting and the contagion processes between bank and sovereign CDS markets.

Outline of the PhD thesis chapters

The first chapter of this PhD thesis compares stock return predictability, that is our ability to forecast returns, in a macro-setting (at the index-level), compared to a micro-setting (at the sectoral or at the stock-level). Note however that economic theory identifies two potential sources of

return predictability: time variation in expected returns, beta-predictability (Cochrane, 2008), or market inefficiencies, alpha-predictability. For the latter, Samuelson argued that macro-returns exhibit more inefficiencies than micro-returns, as individual stories are averaged out, leaving only harder-to-eliminate macro-mispricing at the index-level (Jung and Shiller, 2005). To evaluate this claim, we compare macro- and micro-predictability on US data to gauge if the former turns out higher than the latter. Additionally, we extend over time the methodology of Rapach et al. (2011) to disentangle the two sources of predictability.

On the result side, we show that our interpretation of Samuelson's view appears incorrect, as micro-predictability is not structurally lower than macro-predictability. There is however a diversification mechanism that plays a role between the micro- and macro-predictability time series, given that pooling the different micro-predictability series together over time yields an index that is very close to our macro-predictability estimates. Second, we find that our estimated alpha- and beta-predictability indices are coherent with their corresponding theoretical implications, thus suggesting that the two mechanisms are at play in our dataset. Notably, the alpha-predictability index appears as a theoretically based and easily updatable metric in the financial stability toolkit to spot periods of irrational exuberance.

The second chapter is dedicated to forecasting future economic activity, within a factor and on the basis of sectoral equity variables. The original idea of this work originated after the Covid-shock in March 2020, when stock prices declined abruptly, reflecting both the deterioration of investors' expectations of economic activity as well as the surge in aggregate risk aversion. In the following months however, whereas economic activity remained sluggish, equity markets sharply bounced back. This disconnect between equity values and macro-variables can be partially explained by other factors, namely the decline in risk-free interest rates, and, for the US, the strong profitability of the IT sector.

As a result, an econometrician trying to forecast economic activity with aggregate stock market variables during the Covid-crisis is likely to get poor results. The main idea of the chapter is thus to rely on sectorally disaggregated equity variables within a factor model to predict future US economic activity. We find, first, that the factor model better predicts future economic activity compared to aggregate equity variables or to usual benchmarks used in macroeconomic forecasting (both in-sample and out-of-sample). Second, we show that the strong performance of

the factor model comes from the fact that the model filters out the expected returns component of the sectoral equity variables as well as the foreign component of aggregate future cash flows, and that it also overweights upstream and value sectors that are found to be closely linked to the future state of the US business cycle.

Eventually, in the third chapter of the thesis, we propose a novel approach to quantify spillovers on financial markets based on a structural version of the Diebold-Yilmaz framework (Diebold and Yilmaz, 2009). Key to our approach is a SVAR-GARCH model that is statistically identified by heteroskedasticity, economically identified by maximum shock contribution and that allows for time-varying forecast error variance decompositions. We analyze credit risk spillovers between Eurozone sovereign and bank CDS. This means that our SVAR model includes 16 endogenous variables, enabling to encompass both sovereign and bank CDS, whereas papers close to our study focus only on sovereign CDS (De Santis and Zimic, 2018).

Methodologically, we find the model to better match economic narratives compared with common spillover approaches and to be more reactive than models relying on rolling window estimations. On the economic side, we find credit risk in the Euro Area to be less integrated than suggested by estimates based on traditional Diebold-Yilmaz approaches. We estimate that, on average, credit risk spillovers explain about 37% of the total variation in our sample, amid strong variations of the spillovers over time and in the cross section.

Introduction générale

Perspectives croisées: marchés financiers et banques centrales

Les marchés financiers jouent plusieurs rôles dans nos économies contemporaines. En premier lieu, ils ont essentiellement pour but d'allouer le capital de manière efficiente, à la fois en rémunérant les détenteurs de capital et en apportant des ressources financières aux entreprises (Wurgler, 2000).

D'autre part, cette allocation s'effectue principalement au travers du processus de formation des prix. En effet, à mesure que les intervenants de marché confrontent leurs opinions sur les marchés financiers, un prix d'équilibre va émerger de l'agrégation des différentes offres, fournissant ainsi aux agents un signal-prix indiquant là où le capital est le plus requis. De la sorte, et en lien avec les théories de Hayek sur les mécanismes de formation des prix (Hirshleifer, 1973), les marchés agissent comme un puissant outil informationnel, permettant de fondre la multiplicité des opinions des investisseurs en une seule métrique: le prix des actifs.

Enfin, les marchés financiers sont également le lieu de rencontres entre des agents économiques pouvant avoir des besoins de consommation différents en fonction des futurs états du monde. Si les actifs financiers permettent d'obtenir du rendement uniquement dans certains de ces états, les agents économiques peuvent commercer entre eux afin de lisser leur consommation. Le partage des risques constitue ainsi une dimension essentielle des marchés financiers.

Cependant, il a été également montré que, en dépit de leur utilité, les marchés financiers peuvent aussi être considérés comme source de risques pour deux principales raisons. Tout d'abord, ces derniers sont caractérisés par des processus de bulles spéculatives dans lesquels, alimentés par une forme "d'exubérance irrationnelle" (Shiller, 2015), les prix des actifs augmentent graduellement et divergent de leur valeur fondamentale avant de brutalement décroître. Ces brusques variations de prix ne sont pas sans conséquences pour l'économie réelle. Lors de l'étape de formation de la bulle, dans la mesure où les prix sont décorrélés des valeurs fondamentales, ils peuvent engendrer une allocation inefficente du capital (Miao et al., 2015). Lors de l'éclatement de la bulle, la revalorisation brutale des actifs peut mener à des défauts en cascade pour les institutions financières, et en, en retour, avoir des effets macroéconomiques significatifs. Ces défauts

peuvent avoir lieu en raison des effets de seuil et des mécanismes non-linéaires présents dans la sphère financière, comme dans le cas de rachats massifs pour les fonds d'investissement ou dans le cadre d'appels de marge non-anticipés (Malkiel, 2010).

D'autre part, comme évoqué précédemment, les interconnexions financières peuvent certes être appropriées pour le partage du risque entre agents économiques, mais elles sont également centrales dans la transmission des chocs. En d'autres termes, les marchés financiers peuvent porter un risque pour l'économie non seulement en raison des mécanismes de bulles susmentionnés, mais également parce qu'ils jouent un rôle dans la propagation de chocs négatifs de marché à marché. Ces phénomènes de contagion peuvent avoir lieu pour plusieurs raisons, par exemple lorsqu'une institution financière détient des parts d'une autre institution dans son bilan (Brunetti et al., 2019), lorsque deux institutions possèdent des portefeuilles de marché relativement semblables (Greenwood et al., 2015), ou lorsque deux actifs ont des caractéristiques similaires aux yeux des investisseurs.

Par conséquent, les marchés financiers et les banques centrales entretiennent une relation étroite, à la fois en raison des fonctions assignées aux marchés financiers mais également à cause des risques qui leur sont associés. En effet, bien que les mandats des banques centrales peuvent varier dans le temps et dans l'espace, ils incluent souvent la stabilité des prix (Smets, 2003), parfois combinée avec des objectifs sur le marché de l'emploi comme pour la FED (Thornton et al., 2012) ainsi que, depuis la crise financière de 2008, une mission de stabilité financière (Kim et al., 2018). Par conséquent, la relation entre marchés financiers et banques centrales est protéiforme et dépend de l'objectif de la banque centrale que l'on considère. Nous soulignons ici les trois principales dimensions, à nos yeux, au travers desquelles marchés financiers et banques centrales interagissent entre eux.

Tout d'abord, le principal outil utilisé par les banques centrales pour maintenir la stabilité des prix reste le taux directeur, parfois combiné avec des mesures "non-conventionnelles" comme les programmes d'achats d'actifs ou les politiques de *forward guidance*. En d'autres termes, dans leur fonctionnement même, les banques centrales doivent interagir avec les marchés financiers, dans ce cas précis essentiellement avec les marchés obligataires, pour atteindre leurs objectifs.

Deuxièmement, dans le cadre de leur mandat de stabilité financière, les banques centrales doivent surveiller les risques émanant des marchés. Cela implique, entre autre, de jauger du niveau

de valorisation des actifs afin de déterminer si une bulle est en formation (Geis et al., 2018), d'estimer les interconnexions financières afin d'apprécier les potentiels phénomènes de contagion entre marchés (Alter and Beyer, 2014) ou d'évaluer la résilience des participants de marchés afin de déterminer si ces derniers sont en mesure de résister à de larges chocs négatifs.

Troisièmement, les marchés financiers ont une utilité informationnelle compte tenu de leur capacité, comme mentionné ci-dessus, à agréger les différentes opinions des investisseurs. Cette fonction peut prendre différentes formes. Les banques centrales par exemple se reposent sur les *inflation-linked swaps* (ILS) ou sur la valorisation des actifs protégés de l'inflation pour estimer les anticipations d'inflation, sur le court terme ou sur le long terme, des investisseurs (Bauer et al., 2015). Dans ce cas précis, les banques centrales exploitent ces variables de marché pour jauger de la bonne transmission de leur politique monétaire, reflétée ici par l'ancrage des anticipations d'inflation. Mais les banques centrales peuvent également utiliser les données de marché comme variables d'entrée dans le cadre de leurs prévisions macroéconomiques. En effet, la littérature empirique a démontré que différents indicateurs de marché avaient des capacités prédictives quant au niveau de la future activité économique, et qu'ils avaient ainsi leur place dans le diagnostic de conjoncture des banques centrales. C'est notamment le cas pour le *term spread*, ou prime de terme (Chinn and Kucko, 2015), ou pour le *dividend yield* (Lan et al., 2020).

Cette thèse de doctorat vise ainsi à reconsidérer l'utilité des variables de marché pour les banques centrales, notamment dans le cadre de leur mandat de stabilité financière ainsi que dans leur besoin de prévoir le futur niveau d'activité économique. Le message principal de cette thèse est qu'une proportion considérable des informations tirées de données de marché est occultée si l'on ne considère que des variables agrégées (comme c'est le cas pour une part significative de la littérature macro-financière). Comme nous essayerons de le démontrer dans cette thèse, l'utilisation de données de marché micros/sectorielles peut s'avérer utile dans de nombreux cas, y compris dans l'étude de phénomènes à l'échelle macroéconomique.

Historiquement, l'attention de la littérature macro-financière s'est essentiellement portée sur les variables de marché agrégées

En ce qui concerne les précédents sujets évoqués (les bulles financières, la dimension informa-

tionnelle des marchés, les phénomènes de contagion financière), la littérature a historiquement utilisé des données de marché agrégées. Plusieurs raisons permettent d'expliquer cet attrait pour les variables "macros". En premier lieu elles sont relativement disponibles et couvrent, comparativement aux données "micros", des périodes de temps plus longues. Deuxièmement, en particulier en ce qui concerne les variables de marchés financiers, les données micros peuvent être relativement bruitées. [Fama and French \(1988\)](#) et [Polk et al. \(2006\)](#) soulignent par exemple comment des ratios de valorisation très usités comme le *dividend yield* ou le *price earnings ratio* peuvent ne pas être fiables à l'échelle d'une entreprise en raison de la difficulté d'obtenir des estimations précises de données bilantielles (profits, niveaux des actifs etc.). Troisièmement, d'un point de vue plus méthodologique, certains modèles économétriques utilisés pour répondre à ces questions, comme les Vecteurs Autoregressif (VARs), ont souvent une borne supérieure quant au nombre de variables qui peuvent être considérées dans leur estimation. À ce titre, les données agrégées apparaissent comme des candidats naturels pour l'étude de ces sujets dans le cadre de modèles ne tolérant qu'un nombre restreint de variables.

Pour ce qui est des problématiques de valorisation des actifs, et des potentielles divergences par rapport à leur valeur fondamentale dans le cadre de processus de bulles, la plupart des papiers pionniers sur la question se sont reposés sur des données agrégées. [Shiller et al. \(1983\)](#) par exemple soulignent que, après avoir au préalable estimé la valeur fondamentale historique du Standard and Poor's Composite Stock Price Index et du the Dow Jones Industrial Average, les prix des actions au niveau indiciel étaient trop volatiles pour être seulement influencés par des facteurs fondamentaux. Par la suite, et de façon relativement similaire, d'autres papiers ont essayé d'évaluer la relation entre le prix des actifs et leur valeur fondamentale, mais à nouveau à l'aide de données macros. C'est le cas, entre autres, de [Lee et al. \(1999\)](#) qui cherchent à estimer la valeur fondamentale du Dow Jones à l'aide de techniques de cointégration. Une autre option pour jauger du niveau d'efficience des prix de marchés est d'évaluer dans quelle mesure les prix d'action sont influencés par les nouvelles concernant leurs futurs dividendes. Cette approche a notoirement été utilisée par [Campbell \(1991\)](#) et par [Campbell and Ammer \(1993\)](#), mais également à l'échelle indicielle.

Le rôle informationnel des marchés financiers vis-à-vis du niveau de la future activité économique a été également largement étudié dans la littérature. Ici aussi, les données macros ont été le

principal outil pour l'examen de cette question. Du côté du marché obligataire, un nombre conséquent de papiers ont souligné la bonne performance du *term spread*, la différence entre les taux d'obligations souveraines à court et long terme, pour prédire les futures récessions (Stock and Watson, 1989, Estrella and Hardouvelis, 1991, pour les États-Unis, un résultat étendu aux autres économies avancées par Chinn and Kucko, 2015). Pour le marché action, d'autres variables au niveau indiciel ont également prouvé leur utilité dans le cadre des prévisions macroéconomiques, comme les rendements action (Binswanger, 2000, Ólan Henry et al., 2004, Croux and Reusens, 2013, McMillan, 2021), la croissance du prix des actions (Chen and Ranciere, 2019) ou le *dividend yield* (Lan et al., 2020).

Enfin, comme souligné ci-dessus, la littérature portant sur les phénomènes de contagion financière s'est souvent reposée sur des données à l'échelle de pays compte tenu du problème de dimensionnalité posé par l'inclusion de variables sectorielles. C'est notamment le cas pour les papiers utilisant les modèles VARs structurels (SVARs) dont les stratégies d'identification requièrent un nombre limité de variables. Par conséquent, les papiers précurseurs dans ce champ académique ont privilégié les variables agrégées au niveau indiciel par rapport aux données sectorielles (à la manière de Diebold and Yilmaz, 2009, pour les mouvements de contagion entre rendements action ou Diebold and Yilmaz, 2012, pour une extension de ce modèle au marché obligataire et au marché des changes).

L'utilisation de données de marché micros/sectorielles peut s'avérer pertinente pour l'étude de questions macroéconomiques

Bien que potentiellement problématiques en termes de disponibilité ou de dimensionnalité, les données micros/sectorielles tirées des marchés financiers peuvent être particulièrement utiles sur ces sujets pour trois raisons principales.

Tout d'abord les variables financières sont potentiellement affectées par des facteurs différents à l'échelle micro par rapport à l'échelle macro. Si l'on se concentre par exemple sur les marchés action, on peut montrer, en suivant la formule de valeur actuelle de Campbell and Shiller (1988), que le *dividend yield* (x_t) au niveau indiciel peut se décomposer de la sorte:

$$x_t = \frac{\kappa}{(1 - \rho)} + \sum_{j=1} \rho^{j-1} E_t[r_{t+j} - \Delta cf_{t+j}] \quad (3)$$

Avec $E_t[r_{t+j}]$ les rendements anticipés et $E_t[\Delta cf_{t+j}]$ les dividendes anticipés (κ et ρ étant des paramètres constants). En d'autres termes, cette identité souligne que les ratios de valorisation action dépendent de deux facteurs: de la future croissance des dividendes et des futurs rendements/du taux d'actualisation. La même équation peut être réécrite pour le secteur i :

$$x_{i,t} = \frac{\kappa_i}{(1 - \rho_i)} + \sum_{j=1} \rho_i^{j-1} E_t[r_{i,t+j} - \Delta cf_{i,t+j}] \quad (4)$$

Ces deux identités nous permettent de formaliser l'idée selon laquelle les valorisations action se comportent différemment entre la perspective macro et la perspective micro. Samuelson avait originellement l'intuition que les rendements action des firmes étaient principalement influencés par des nouvelles quant à la profitabilité de ces dernières, et que ces nouvelles n'avaient pas d'influence au niveau macro en raison de processus de diversification. Par conséquent, dans le cadre de cette hypothèse, les rendements indiciaires sont essentiellement affectés par les variations du taux d'actualisation ([Jung and Shiller, 2005](#)).

Un tel différentiel de comportement entre les rendements micros et macros a en effet été documenté empiriquement. [Sadka and Sadka \(2009\)](#) par exemple soulignent que la relation positive entre la croissance des profits et les rendements au niveau micro devient négative au niveau macro. [Kothari et al. \(2006\)](#) fournissent des résultats similaires pour les surprises quant à la profitabilité d'une firme et aux rendements contemporains de cette dernière. Au niveau micro les rendements sont positivement liés à ces surprises, mais cette corrélation disparaît au niveau indiciaire. De la même manière, [Hirshleifer et al. \(2009\)](#) relèvent le fait que des provisions élevées tendent à prédire des futurs rendements négatifs pour une entreprise, mais des rendements nuls ou positifs pour un indice. Toutes ces études partagent le même raisonnement: les dividendes reflètent des comportements idiosyncratiques tandis que les taux d'actualisation sont communs aux différentes firmes. De la sorte, les facteurs affectant les rendements action peuvent fortement varier en fonction de l'échelle que l'on considère. Ces facteurs sont plus associés aux anticipations de profitabilité au niveau micro mais, en raison de processus de diversification, ils

dépendent plus du taux d'actualisation (et donc par exemple du niveau d'aversion au risque des investisseurs) au niveau macro.

Ce contraste de comportement entre rendements micros et macros peut être bénéfique pour le chercheur. Il est en effet possible de tirer profit de l'hétérogénéité en coupe transversale de ces facteurs liés aux futurs dividendes pour estimer le niveau du taux d'actualisation requis par les investisseurs. [Kelly and Pruitt \(2013\)](#) par exemple estiment un modèle à facteurs à partir d'un échantillon de *book-to-market ratios* sectoriels. L'utilisation de données sectorielles permet de filtrer la composante-dividende du prix des actions et d'obtenir une évaluation du taux d'actualisation, ce dernier permettant en retour d'obtenir une prédiction relativement précise des futurs rendements au niveau indiciel.

Deuxièmement, les données sectorielles peuvent être bénéfiques non seulement parce que les variables financières micros se comportent différemment des variables macros, mais aussi parce que les chercheurs peuvent exploiter l'hétérogénéité des comportements entre variables micros. À nos yeux, cette hétérogénéité peut être utile sur plusieurs aspects.

Dans la mesure où les dividendes sectoriels répondent de manière différente à des chocs macroéconomiques, évaluer comment les valorisations action micros réagissent lors de dates spécifiques permet précisément d'identifier ces chocs. Ainsi, [Venditti and Veronese \(2020\)](#) utilisent le prix des actions des compagnies aériennes pour identifier les chocs d'offre de pétrole, notamment car l'augmentation des prix du pétrole correspondante est susceptible d'affecter plus les rendements des compagnies aériennes que celles des indices action agrégés.

De la même manière, les dividendes sectoriels, et donc les ratios de valorisation sectoriels, peuvent répondre différemment à des variations du futur niveau d'activité économique. Par conséquent, surpondérer certains secteurs par rapport à d'autres peut s'avérer bénéfique dans le cadre de prévisions macroéconomiques. [Andersson and Agostino \(2008\)](#) ont ainsi montré par exemple comment les secteurs des hydrocarbures et des matériaux de base présentaient des rendements action qui étaient plus corrélés au futur PIB de la Zone Euro que les autres secteurs.

Enfin, l'hétérogénéité entre les prix des actions au niveau micro importe au-delà des deux canaux liés aux dividendes sectoriels mentionnés ci-dessus. Un mécanisme plus subtil, décrit à nouveau par [Kelly and Pruitt \(2013\)](#), consiste dans le fait que les secteurs *growth*, typiquement constitués d'entreprises du secteur numérique, ont des dividendes plus éloignés dans le temps que les

secteurs *value*, qui ont tendance, eux, à présenter un profil de futurs dividendes stable dans le temps. Par conséquent, les secteurs *growth* vont présenter une *equity duration* plus forte que les secteurs *value*. Dit autrement, leurs ratios de valorisation seront plus sensibles aux variations du taux d'actualisation. Cette différence de sensibilité peut en retour être utile pour filtrer les ratios de valorisation sectoriels de leur composante liée au taux d'actualisation afin de mieux prédire la croissance des dividendes agrégés (à la manière de [Kelly and Pruitt, 2013](#)).

Enfin, les séries temporelles sectorielles peuvent être bénéfiques dans l'analyse des phénomènes de contagion dans la mesure où elles peuvent permettre d'identifier la chaîne de causalité dans les processus de propagation de chocs. À titre d'exemple, différents papiers dans la littérature se sont penchés sur les dynamiques des taux d'intérêt ou des *Credit Default Swaps* (CDS) souverains durant la crise de la dette européenne en étudiant uniquement des variables souveraines à l'échelle des pays ([Ehrmann and Fratzscher, 2017](#), [De Santis and Zimic, 2018](#)). Pourtant, une part substantielle des dynamiques de CDS sur la période provient de phénomènes de cercles vicieux de propagation du risque de crédit entre les variables souveraines et les secteurs bancaires domestiques. De tels mécanismes de rétroaction peuvent avoir lieu pour plusieurs raisons. Le risque bancaire peut, entre autres, affecter le risque souverain *via* le "canal du renflouement", à savoir des garanties publiques, explicites ou implicites, dans le cas de tension sur le secteur bancaire ([Alter and Schüler, 2012](#)), ou parce que les banques détiennent un montant significatif d'obligations souveraines dans leurs comptes (le "canal bilantiel", [Angeloni and Wolff, 2012](#), [Buch et al., 2016](#)). Par conséquent, omettre des variables sectorielles telles que les taux d'intérêt bancaires ou les CDS bancaires peut engendrer une vision biaisée du processus de contagion.

Nous avons essayé, au travers des trois arguments mentionnés ci-dessus, de souligner l'importance des données de marché micros/sectorielles, y compris pour l'étude de phénomènes macroéconomiques. Les trois chapitres de cette thèse de doctorat visent à mettre en exergue le bénéfice tiré de ces dernières dans trois domaines: la prédiction des rendements action, l'utilisation des données action pour la prévision macroéconomique et, enfin, l'étude des processus de contagion entre les marchés de CDS souverains et bancaires.

Résumé des trois chapitres de thèse

Le premier chapitre de cette thèse de doctorat compare la prédictibilité des rendements action, c'est-à-dire notre capacité à prévoir les rendements boursiers dans le futur, au niveau macro par rapport au niveau micro. Il est à noter toutefois que la théorie économique identifie deux sources de prédictibilité: la variation dans le temps des rendements anticipés, la prédictibilité-bêta (Cochrane, 2008), ou les inefficiences de marché, la prédictibilité-alpha. Pour cette dernière, Samuelson avançait l'idée que les rendements macros étaient plus inefficients que les rendements micros, dans la mesure où les facteurs (efficients) idiosyncratiques ne se transposaient pas au niveau indiciel en raison de processus de diversification. En conséquence de quoi, les rendements des indices présenteraient plus d'inefficiences que les rendements micros (Jung and Shiller, 2005). Pour évaluer cette hypothèse, nous comparons les prédictibilités micros et macros sur données américaines afin d'identifier si les premières sont effectivement moins présentes que les secondes. De plus, nous reprenons la méthodologie de Rapach et al. (2011) en l'étendant à un cadre non-constant dans le temps afin de dénouer, au cours du temps, les deux sources de prédictibilité.

Pour ce qui est des résultats, nous montrons que notre interprétation de l'intuition de Samuelson n'est pas valide dans la mesure où la prédictibilité micro n'est pas plus faible que la prédictibilité macro. Toutefois, nous montrons également que des phénomènes de diversification sont bien à l'oeuvre dans la mesure où l'agrégation des séries de prédictibilité micro au cours du temps donne un indice qui est très proche de notre série de prédictibilité macro. Deuxièmement, nous montrons que nos estimations des prédictibilités-alpha et -bêta sont cohérentes avec leurs implications théoriques. Cela suggère notamment que les deux phénomènes jouent un rôle dans notre base de données.

Le deuxième chapitre de cette thèse porte sur la prévision d'activité économique sur la base de données action sectorielles et dans le cadre d'un modèle à facteurs. L'idée originale de ce travail émergea lors du choc du Covid-19 en mars 2020 lorsque les prix d'actions chutèrent brutalement en lien avec la détérioration des perspectives d'activité économique et avec la hausse de l'aversion au risque des investisseurs. Dans les mois qui suivirent toutefois, alors que le niveau d'activité restait relativement morose, les marchés action rebondirent rapidement. Cette déconnexion entre variables boursières et variables macros peut en partie être expliquée par d'autres facteurs, notamment par le déclin des taux d'intérêt sans risque et, pour les États-Unis,

par la forte rentabilité du secteur du numérique.

Par conséquent, un économètre essayant de prédire l'activité économique pendant la crise du Covid à l'aide de données agrégées du marché action obtiendrait des résultats médiocres. La principale idée de ce chapitre est ainsi d'utiliser des données action sectorielles, dans le cadre d'un modèle à facteurs, pour prédire la future production industrielle américaine. Pour ce qui est des résultats, nous trouvons, en premier lieu, que notre modèle à facteurs prédit mieux l'activité économique par rapport aux variables agrégées du marché action mais également par rapport aux variables de référence utilisées en prédiction macroéconomique (à la fois à l'intérieur et à l'extérieur de l'échantillon). Deuxièmement, nous montrons que la bonne performance de notre modèle vient du fait qu'il filtre, parmi les variables action sectorielles, la composante liée au taux d'actualisation ainsi que la composante liée aux dividendes tirés de l'activité étrangère des firmes. Nous soulignons également que notre modèle surpondère les secteurs en amont des chaînes de valeur ainsi que les secteurs *value* qui apparaissent intimement liés au niveau du futur cycle des affaires américain.

Enfin, dans le troisième chapitre, nous proposons une nouvelle approche pour quantifier les phénomènes de contagion entre marchés financiers sur la base d'une version structurelle du modèle de (Diebold and Yilmaz, 2009). Nous nous repons essentiellement sur un modèle SVAR-GARCH identifié statistiquement par hétéroscédasticité, identifié économiquement par la contribution maximale des chocs et qui permet d'obtenir des décompositions des erreurs de prévision non-constantes dans le temps. Nous analysons la propagation des chocs de risque de crédit dans la Zone Euro entre CDS souverains et bancaires. Cela implique que notre modèle SVAR inclue 16 variables endogènes, permettant ainsi de considérer à la fois les CDS souverains et bancaires alors que les papiers proches de notre étude se focalisent la plupart du temps uniquement sur les CDS souverains (De Santis and Zimic, 2018).

En termes méthodologiques, nous trouvons que notre modèle permet de mieux identifier les chocs de crédit par rapport aux autres modèles de contagion de la littérature, et qu'il est de plus plus réactif aux événements que les modèles basés sur des estimations par fenêtres roulantes. Du point de vue économique, nous trouvons que les phénomènes de contagion expliquent seulement 37% de la variance de nos variables, avec toutefois de fortes variations dans le temps.

Bibliography

- Alter, A., and A. Beyer. 2014. The dynamics of spillover effects during the European sovereign debt turmoil. *Journal of Banking and Finance* 42:134–153.
- Alter, A., and Y. S. Schüler. 2012. Credit spread interdependencies of European states and banks during the financial crisis. *Journal of Banking and Finance* 36:3444–3468.
- Andersson, B., and A. D. . Agostino. 2008. Are sectoral stock prices useful for predicting Euro Area GDP? *ECB Working Paper* .
- Angeloni, C., and G. B. Wolff. 2012. Are banks affected by their holdings of government debt? *Bruegel Working Paper* 07.
- Bauer, M. D., E. McCarthy, et al. 2015. Can we rely on market-based inflation forecasts? *FRBSF Economic Letter* 30:1–5.
- Binswanger, M. 2000. Stock returns and real activity: Is there still a connection? *Applied Financial Economics* 10:379–387.
- Brunetti, C., J. H. Harris, S. Mankad, and G. Michailidis. 2019. Interconnectedness in the interbank market. *Journal of Financial Economics* 133:520–538.
- Buch, C. M., M. Koetter, and J. Ohls. 2016. Banks and sovereign risk: A granular view. *Journal of Financial Stability* 25:1–15.
- Campbell, J., and J. Ammer. 1993. What Moves the Stock and Bond Markets? A Variance Decomposition for LongTerm Asset Returns. *The Journal of Finance* 48:3–37.
- Campbell, J. Y. 1991. A Variance Decomposition for Stock Returns. *Economic Journal* 101:157–79.
- Campbell, J. Y., and R. J. Shiller. 1988. Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance* 43:661–676.
- Chen, S., and R. Ranciere. 2019. Financial information and macroeconomic forecasts. *International Journal of Forecasting* 35:1160–1174.

- Chinn, M., and K. Kucko. 2015. The predictive power of the yield curve across countries and time. *International Finance* 18:129–156.
- Cochrane, J. H. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21:1533–1575.
- Croux, C., and P. Reusens. 2013. Do stock prices contain predictive power for the future economic activity? A Granger causality analysis in the frequency domain. *Journal of Macroeconomics* 35:93–103.
- De Santis, R. A., and S. Zimic. 2018. Spillovers among sovereign debt markets: Identification through absolute magnitude restrictions. *Journal of Applied Econometrics* 33:727–747.
- Diebold, F. X., and K. Yilmaz. 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal* 119:158–171.
- Diebold, F. X., and K. Yilmaz. 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28:57–66.
- Ehrmann, M., and M. Fratzscher. 2017. Euro Area Government Bonds: Integration and Fragmentation During the Sovereign Debt Crisis. *Journal of International Money and Finance* 70:26–44.
- Estrella, A., and G. A. Hardouvelis. 1991. The term structure as a predictor of real economic activity. *The Journal of Finance* 46:555–576.
- Fama, E. F., and K. R. French. 1988. Permanent and Temporary Components of Stock Prices. *Journal of Political Economy* 96:246–273.
- Geis, A., D. Kapp, and K. Kristiansen. 2018. Measuring and interpreting the cost of equity in the euro area. *ECB Economic Bulletin Articles* 4.
- Greenwood, R., A. Landier, and D. Thesmar. 2015. Vulnerable banks. *Journal of Financial Economics* 115:471–485.
- Hirshleifer, D., K. Hou, and S. H. Teoh. 2009. Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics* 91:389–406.

- Hirshleifer, J. 1973. Where are we in the theory of information? *The American Economic Review* 63:31–39.
- Jung, J., and R. J. Shiller. 2005. Samuelson’s dictum and the stock market. *Economic Inquiry* 43:221–228.
- Kelly, B., and S. Pruitt. 2013. Market expectations in the cross-section of present values. *Journal of Finance* 68:1721–1756.
- Kim, H. J., H. S. Shin, and J. Yun. 2018. Monetary aggregates and the central bank’s financial stability mandate. *International Journal of Central Banking* .
- Kothari, S. P., J. Lewellen, and J. B. Warner. 2006. Stock returns, aggregate earnings surprises, and behavioral finance. *Journal of Financial Economics* 79:537–568.
- Lan, C., N. Roussanov, C. Lan, and N. Roussanov. 2020. Stock Price Movements: Business-Cycle and Low-Frequency Perspectives. *The Review of Asset Pricing Studies* 10:335–395.
- Lee, C. M., J. Myers, and B. Swaminathan. 1999. What is the intrinsic value of the dow? *Journal of Finance* 54:1693–1741.
- Malkiel, B. G. 2010. Bubbles in asset prices. In *The Oxford Handbook of Capitalism*.
- McMillan, D. G. 2021. Predicting GDP growth with stock and bond markets: Do they contain different information? *International Journal of Finance Economics* 26:3651–3675.
- Miao, J., P. Wang, and J. Zhou. 2015. Asset bubbles, collateral, and policy analysis. *Journal of Monetary Economics* 76:S57–S70.
- Polk, C., S. Thompson, and T. Vuolteenaho. 2006. Cross-sectional forecasts of the equity premium. *Journal of Financial Economics* 81:101–141.
- Rapach, D. E., J. K. Strauss, J. Tu, and G. Zhou. 2011. Out-of-Sample Industry Return Predictability: Evidence from A Large Number of Predictors. *SMU Working Paper 2-2011*.
- Sadka, G., and R. Sadka. 2009. Predictability and the earnings-returns relation. *Journal of Financial Economics* 94:87–106.

- Shiller, R. J. 2015. *Irrational exuberance: Revised and expanded third edition*. Princeton university press.
- Shiller, R. J., et al. 1983. Do stock prices move too much to be justified by subsequent changes in dividends?: Reply. *American Economic Review* 73:236–237.
- Smets, F. 2003. Maintaining price stability: how long is the medium term? *Journal of Monetary Economics* 50:1293–1309.
- Stock, J. H., and M. W. Watson. 1989. New indexes of coincident and leading economic indicators. *NBER macroeconomics annual* 4:351–394.
- Thornton, D. L., et al. 2012. The Dual Mandate: Has the Fed Changed Its Objective? *Federal Reserve Bank of St. Louis Review* 94:117–133.
- Venditti, F., and G. Veronese. 2020. Global financial markets and oil price shocks in real time. *ECB Working Paper* .
- Wurgler, J. 2000. Financial markets and the allocation of capital. *Journal of financial economics* 58:187–214.
- Ólan Henry, N. Olekalns, J. Thong, Ólan Henry, N. Olekalns, and J. Thong. 2004. Do stock market returns predict changes to output? Evidence from a nonlinear panel data model. *Empirical Economics* 29:527–540.

Contents

Acknowledgement	i
General introduction / Introduction générale	iv
1 Stock Return Predictability: comparing Macro- and Micro-Approaches	1
1.1 Introduction	3
1.2 Return Predictability in the Literature	6
1.3 Working Hypotheses	8
1.3.1 H_1 , Samuelson's view	9
1.3.2 H_2 , Cochrane's view	10
1.3.3 H_3 , Third view	12
1.4 Data and Methodology	13
1.4.1 Stock Return Data	13
1.4.2 Constructing Raw Predictability Metrics	14
1.4.3 Disentangling the Sources of Predictability	15
1.5 Empirical Results	17
1.5.1 Micro- and Macro- Raw Predictability	17
1.5.2 Alpha- and Beta-Predictability	21
1.6 Robustness checks	28
1.7 Conclusion	29
Appendices	35
1.A Appendix	35

1.A.1	List of estimated Models	35
1.A.2	Datasets	37
1.A.3	Raw Predictability series: individual graphs	39
1.A.4	Moments of the raw return predictability series	40
1.A.5	Standard errors, mean and standard deviations of raw predictability series	41
1.A.6	Individual stock return predictability	43
1.A.7	Robustness checks: alternative risk factors	46
1.A.8	Robustness checks: regression results	48
2	Forecasting Real Activity using Cross-Sectoral Stock Market Information	49
2.1	Introduction	51
2.2	Background	54
2.2.1	Theoretical Framework	54
2.2.2	Selected Literature Review	55
2.3	Model Specification and Data	58
2.3.1	A Factor Model	58
2.3.2	Data	59
2.4	In-Sample Results	60
2.5	Out-of-Sample Results	63
2.5.1	Out-of-Sample Performance	63
2.5.2	Comparison with traditional factor models	66
2.5.3	Performance by Sample Period	68
2.6	Economic Interpretation	70
2.6.1	Filtering the "return" and the "foreign cash flow" components	70
2.6.2	sec2tor overweighting	74
2.7	Conclusion	78
	Appendices	83
2.A	Appendix	83
2.A.1	The Factor model for sectoral and aggregate DYs	83
2.A.2	Additional forecasting results	85
2.A.3	Dataset - traditional factor model	87

2.A.4	Estimated factor	90
3	Structural estimation of time-varying spillovers: An application to credit risk transmission	91
3.1	Introduction	93
3.2	Estimating Spillovers in the Literature	96
3.3	Methodology	98
3.3.1	Measuring spillovers	98
3.3.2	Description of the Model	100
3.4	Data and filtering for common shocks	105
3.4.1	Data	105
3.4.2	Filtering for common shocks	106
3.5	Results	107
3.5.1	Econometric results	107
3.5.2	Economic results	115
3.6	Conclusion	128
	Appendices	135
3.A	Annex	135
3.A.1	Derivation of Forecast Error Variance Decomposition	135
3.A.2	CDS Data	137
3.A.3	Granger causality test	138
3.A.4	Data sources OLS regressions	139
3.A.5	Robustness checks	140
3.A.6	Test for identification and estimated coefficients	142
3.A.7	List of Events	143
3.A.8	IRF assumptions	146
4	Conclusion	149
	List of Figures	152
	List of Tables	154

1. Stock Return Predictability: comparing Macro- and Micro-Approaches

Abstract

Economic theory identifies two potential sources of return predictability: time variation in expected returns (beta-predictability) or market inefficiencies (alpha-predictability). For the latter, Samuelson argued that macro-returns exhibit more inefficiencies than micro-returns, as individual stories are averaged out, leaving only harder-to-eliminate macro-mispricing at the index-level. To evaluate this claim, we compare macro- and micro-predictability on US data to gauge if the former turns out higher than the latter. Additionally, we extend over time the methodology of [Rapach et al. \(2011\)](#) to disentangle the two sources of predictability. We first find that Samuelson's view appears incorrect, as micro-predictability is not structurally lower than macro-predictability. Second, we find that our estimated alpha- and beta-predictability indices are coherent with their corresponding theoretical implications, thus suggesting that the two mechanisms are at play in our dataset.

La théorie économique identifie deux sources de prédictibilité: la variation dans le temps des rendements anticipés, la prédictibilité-bêta, ou les inefficiences de marché, la prédictibilité-alpha. Pour cette dernière, Paul Samuelson avançait l'idée que les rendements macros étaient plus inefficients que les rendements micros. En effet, si les facteurs efficients affectant les rendements micros de manière idiosyncratique ne se transposent pas au niveau indiciel en raison de processus de diversification, alors les rendements macros sont essentiellement influencés par des inefficiences de marché. Pour évaluer cette hypothèse, nous comparons les prédictibilités micros et macros sur données américaines afin d'identifier si, effectivement, la prédictibilité micro s'avère moins présente que la prédictibilité macro. De plus, nous reprenons la méthodologie de [Rapach et al. \(2011\)](#) en l'étendant à un cadre non-constant dans le temps afin de dénouer, au cours du temps, les deux sources de prédictibilité. Pour ce qui est des résultats, nous montrons que notre interprétation de l'intuition de Samuelson n'est pas valide puisque la prédictibilité micro n'est pas plus faible que la prédictibilité macro. Toutefois, nous montrons également que des phénomènes de diversification sont bien à l'oeuvre dans la mesure où l'agrégation des séries de prédictibilité micro au cours du temps donne un indice qui est très proche de notre série de prédictibilité macro. Deuxièmement, nous montrons que nos estimations des prédictibilités-alpha et -bêta sont cohérentes avec leurs implications théoriques. Cela suggère notamment que les deux phénomènes jouent un rôle dans notre base de données.

Keywords: *out-of-sample return predictability; efficient market hypothesis; conditional beta pricing model; alpha predictability.*

JEL: C22, C53, G12, G14, G17

1.1. Introduction

Some forms of the Efficient Market Hypothesis (EMH) imply that stock returns are not predictable (Fama, 1970, Pesaran, 2010). Since all available information is already embedded in asset prices, changes in the latter can only be caused by the arrival of new information which is by definition unpredictable. In other words, prices should follow a random walk, and running a regression of future returns, r_{t+1} , on past information, \mathbf{X}_t , should not yield predictive content. At the same time, stock market efficiency may differ between a macro-perspective and a micro-perspective. Paul Samuelson argued in this sense (Jung and Shiller, 2005):

Modern markets show considerable micro efficiency (for the reason that the minority who spot aberrations from micro efficiency can make money from those occurrences). [...] In no contradiction to the previous sentence, I had hypothesized considerable macro inefficiency, in the sense of long waves in the time series of aggregate indexes of security prices below and above various definitions of fundamental values.

Samuelson's intuition amounts to a model where micro-returns are driven both by idiosyncratic *efficient* components and by a common *inefficient* component (as micro-inefficiencies are arbitrated away by investors). If these idiosyncratic factors are independently distributed, they will average out in the aggregate, leaving at the index-level only the inefficient component of returns. Consequently, if stock return predictability is a gauge of inefficiency, and if Samuelson's view is correct, then we should observe higher levels of predictability at the macro- than at the micro-level.

The contribution of this paper is twofold. First, we compare, over time, macro- and micro-series of return predictability. Although the literature on this subject is enormous, to our knowledge we are the first ones to conduct this exercise in a time-varying manner. Allowing time variation in our results matters, as return predictability appears largely to be a regime-dependent

phenomenon (Henkel et al., 2011, Farmer et al., 2022). Second, based on Rapach et al. (2011), we contribute to the literature aiming at identifying the drivers of return predictability by building a new indicator that is, theoretically, directly linked with market inefficiencies: the alpha-predictability index. On the result side, we first show that, contrary to Samuelson's view, aggregate returns do not exhibit higher levels of predictability compared to micro-returns. Second, we document that, as expected, our alpha-predictability index is positively linked with metrics of market effervescence.

However, more precisely, modern views of the EMH underline that a certain extent of return predictability can persist even in an efficient market setting. The aforementioned no-predictability paradigm implied that stock prices followed a random walk, and that expected returns were constant. On the contrary, Cochrane (2008) argues that, as investors' risk aversion varies over time, expected returns vary as well. Taking into account time variation in expected returns along the business cycle can therefore generate return predictability even in the absence of market inefficiencies.

To put it bluntly, in the midst of an economic crisis, investors become highly risk averse. This leads to a decline in stock prices and to an increase in expected returns. People could therefore predict that returns will be high in the future, but they are too concerned about their current situation to benefit from it. In the same strand of the literature, empirical papers also argued that this mechanism should be especially at play during economic downturns (Henkel et al., 2011, Dangl and Halling, 2012, Rapach et al., 2010).

Therefore, the interpretation of predictability is sensitive. High level of predictability can reflect inefficiencies such as investors' irrationality or market frictions. But it can also mirror variations in aggregate risk aversion. Consequently, in order to clarify our framework, we present three hypotheses that summarise the different views on return predictability

In the first hypothesis, linked with the Samuelson's view, macro-predictability should be higher than micro-predictability, especially in times of irrational exuberance (e.g. during the dot-com

bubble). The second hypothesis, in line with Cochrane's view, states that micro- and macro-predictability should not behave differently as they are influenced by the same factor: changes in aggregate risk aversion. They should therefore evolve in tandem and be higher during recessions. A third "in-between" hypothesis assumes that returns at the micro-level are driven by idiosyncratic factors that can be either efficient (e.g. news about cash flows) or inefficient (e.g. illiquidity issues). The former decrease micro-predictability, whereas the latter increase it. At the aggregate level, both types of individual factors are averaged out, so that micro-predictability can either be higher or lower than macro-predictability. Eventually, this third view is agnostic regarding the sources of macro-predictability, which can therefore be high both during recessions and during market effervescence periods.

We test the three different hypotheses on US post-war data, with an out-of-sample methodology that combines 23 models estimated on rolling windows. These models are commonly used in the return predictability literature and encompass both traditional econometric methods, factor modelling approaches and Machine Learning techniques. The large number of approaches considered here reflects the substantial model instability in forecasting returns exercises ([Timmermann, 2018](#)).

We find overall that our results corroborate the third hypothesis for at least two reasons. First micro-predictability is neither structurally higher nor lower than macro-predictability. On the contrary, micro-predictability "bounces around" macro-predictability. This result is in line with a model where micro-predictability level depends on the relative importance of efficient or inefficient idiosyncratic component of returns. Second, we extend the methodology of [Rapach et al. \(2011\)](#) in a time-varying manner so as to disentangle the sources of macro-predictability. The two resulting series, the alpha-predictability and the beta-predictability indices, should track changes in macro-predictability due to market inefficiencies and due to time-varying risk aversion, respectively. In accord with the third hypothesis, we find that the alpha-predictability index is positively associated with metrics of market exuberance, whereas the beta-predictability index

correlates with business cycle variables. This finding underlines that the two sources of return predictability are at play in our dataset, and therefore enables to reconcile the diverging views in the literature about the drivers of return predictability.

The rest of the paper is structured as follows: Section 1.2 details how the current paper is located in the return predictability literature, Section 1.3 describes the three theoretical hypotheses outlined above, Section 1.4 presents the methodology and the datasets used, Section 1.5 reports the empirical results, Section 1.6 provides different robustness checks and Section 1.7 concludes.

1.2. Return Predictability in the Literature

The literature on stock return predictability is extensive and has considerably evolved over time. Seminal papers focused on aggregate stock returns, most of the time reporting in-sample results within linear regression approaches. Various macro-financial variables appeared to have some predictive power, such as the dividend yield (Fama and French, 1988, Campbell and Shiller, 1988), the term structure of interest rates (Campbell, 1987) or the consumption-wealth ratio (Lettau and Ludvigson, 2001). Nevertheless, in sharp contrast with the previous studies, Welch and Goyal (2008) underline that the former results are hardly replicable. In a linear setting, return predictability appears as a spurious result, both in-sample and out-of-sample.

However, relying on more sophisticated techniques, subsequent papers claim to forecast future returns, although most of the time with relatively low R^2 . These innovative approaches fall mainly in three non-exclusive categories.

First, return prediction is a specific forecasting exercise in itself, as the use of a performing model by investors is likely to erase the predictability pattern the model is based upon (Timmermann, 2018). The resultant instability in the predicting relationship paved the way for forecast averaging techniques, since they enable the econometrician not to rely on the assumptions of a

single model. This includes notably simple and advanced forecast combination methods (Aiolfi and Timmermann, 2006, Rapach et al., 2010, Elliott and Timmermann, 2013, Baetje, 2018) or Bayesian Model Averaging (Dangl and Halling, 2012). Second, in line with other financial market variables, stocks returns are mostly influenced by investors' expectations. These expectations constitute an unobserved variable, but can be included in the predictive model as a latent factor. Consequently, theory-driven approaches in the form of factor models have proven to perform relatively well at different frequencies (Binsbergen and Koijen, 2010, Kelly and Pruitt, 2013). Third, given the complex structure of financial markets, it is unlikely that stock returns follow a linear process. As a result, different studies have explicitly investigated non-linear forecasting techniques. This comprises restricted linear models (Campbell and Thompson, 2008), non-linear VARs (Henkel et al., 2011), non-parametric approaches (Farmer et al., 2022), or Machine Learning methodologies (Rapach et al., 2019, Chincó et al., 2019).

Although all these analyses have exposed in-sample or out-of-sample forecastability, debate remains about the drivers of return predictability over time. Some papers underline that, in line with Cochrane's view, predictability is a countercyclical phenomenon and is therefore elevated during economic downturns (Rapach et al., 2010, Henkel et al., 2011, Dangl and Halling, 2012). On the contrary, other studies argued that returns are especially predictable in bullish financial markets (Farmer et al., 2022), while other identified specific periods of return predictability (e.g. surrounding the oil price shock of 1973, Welch and Goyal, 2008, Timmermann, 2008).

Yet, return predictability is not the only available metric to gauge market inefficiencies. One intuitive way to do so is to estimate the informative content of stock prices (Bai et al., 2016). In other words, are current prices useful to predict future cash flows? This recent work echoes older literature that evaluated to what extent stock returns were driven by future cash flows or by future returns (Campbell, 1991, Campbell and Ammer, 1993). Another method amounts to estimate a fundamental value for stock prices, and to define market inefficiency as the departure of observed prices from this estimate (Lee et al., 1999).

Most of these metrics of inefficiency are based on aggregate data. However some papers extended the above methodologies for individual stocks or for subgroups of stocks (Vuolteenaho, 2002, Cohen et al., 2003, Davila and Parlato, 2018), sometimes with indicators that evolve over time (Farboodi et al., 2020). Similarly, some studies evaluate return forecastability at the stock-level, but without reporting specifically micro-predictability (Avramov and Chordia, 2006), without time variation in the results (Rapach et al., 2011, Kong et al., 2011) or without drawing a proper micro-macro analysis (Guidolin et al., 2013, Chincó et al., 2019).

Compared to the aforementioned studies, the goal of the present paper is to compare, *over time*, *macro- and micro-predictability* so as to extract from this analysis a metric of market inefficiencies¹. This question has, to our knowledge, never been addressed in the literature.

1.3. Working Hypotheses

We formalize in this section the three hypotheses outlined above. Following Avramov (2004) and Rapach et al. (2010) we express (excess) aggregate returns as:

$$r_{t+1} = \alpha(\mathbf{X}_t) + \beta_t' \mathbf{f}_{t+1} + \epsilon_{t+1} \quad (1.1)$$

Where $\alpha(\mathbf{X}_t)$ represent the inefficient part of returns, \mathbf{f}_{t+1} a vector of portfolio-based factors capturing systematic risk, β_t the corresponding vector of factor loadings and ϵ_{t+1} a disturbance term of mean zero.

Two sources of return predictability are potentially at play here. With time t variables, the econometrician is able to predict market inefficiencies $\alpha(\mathbf{X}_t)$. Additionally, return predictabil-

¹There are also many papers, outside the return predictability literature, that underline that stock markets behave differently at the stock-level compared to the index-level. Sadka and Sadka (2009) document that the positive relationship between earning growth and returns at the micro-level turns negative at the macro-level. Kothari et al. (2006) report similar findings between earning surprises and contemporaneous returns. Eventually, Hirshleifer et al. (2009) stress that elevated accruals predict negative future returns at the stock-level, but null or positive future returns at the index level. As such, drivers of stock returns may differ greatly depending on the scale we are considering.

ity can emerge from the forecastability of risk factors if we further assume that they evolve likewise:

$$\mathbf{f}_{t+1} = g(\mathbf{X}_t) + \mathbf{u}_{t+1} \quad (1.2)$$

Where $g(\mathbf{X}_t)$ is a vector of (forecastable) conditional expected returns for the risk factors and \mathbf{u}_{t+1} a vector of mean-zero disturbance terms independent of ϵ_{t+1} .

Besides, we consider that micro-returns $r_{i,t+1}$ are affected by aggregate factors $\alpha(\mathbf{X}_t)$ and ϵ_{t+1} , but also by their individual counterparts: $\alpha_i(\mathbf{X}_t)$ and $\epsilon_{i,t+1}$ (that is, idiosyncratic inefficiencies and idiosyncratic unpredictable shocks). We assume that $\alpha_i(\mathbf{X}_t)$ and $\epsilon_{i,t+1}$ are centered around 0, and are diversified away at the macro-level. More precisely we write our system of macro- and micro-returns such as:

$$\begin{cases} r_{i,t+1} = \alpha_i(\mathbf{X}_t) + \omega_i \alpha(\mathbf{X}_t) + \beta'_{i,t} \mathbf{f}_{t+1} + \epsilon_{i,t+1} + \delta_i \epsilon_{t+1} \\ r_{t+1} = \alpha(\mathbf{X}_t) + \beta'_t \mathbf{f}_{t+1} + \epsilon_{t+1} \\ \mathbf{f}_{t+1} = g(\mathbf{X}_t) + \mathbf{u}_{t+1} \end{cases} \quad (1.3)$$

With ω_i and δ_i the exposures of $r_{i,t+1}$ to the common factors $\alpha(\mathbf{X}_t)$ and ϵ_{t+1} , respectively, and with $\epsilon_{i,t+1}$ being independent from ϵ_{t+1} and from \mathbf{u}_{t+1} . The system of equations 1.3 constitutes the basis for the three following hypotheses.

1.3.1 H_1 , Samuelson's view

Our first hypothesis is built upon Samuelson's intuition and entails several implications.

First, we consider here that $\alpha_i(\mathbf{X}_t) = 0$ given that, in line with Samuelson, arbitrageurs should eradicate micro-mispricings. Second, at the time where Samuelson expressed this idea, the theory of return predictability driven by time-varying expected returns was not formulated yet. Some studies even modelled expected returns as a constant (Samuelson, 1975). We therefore

suppose here that $g(\mathbf{X}_t) = \mathbf{c}$, with \mathbf{c} a constant vector, so that $\mathbf{f}_{t+1} = \mathbf{c} + \mathbf{u}_{t+1}$. Third, as micro-inefficiencies are arbitrated away, and since the efficient idiosyncratic news are averaged out in the aggregate, it is assumed here that micro-returns are more driven by unpredictable components than macro-returns. Consequently, return predictability should be higher in the aggregate than at the micro-level². Fourth, the common predictable factor $\alpha(\mathbf{X}_t)$ should especially be forecastable in times of elevated market inefficiency.

The System 1.3 can therefore be rewritten for H_1 as:

$$\begin{cases} r_{i,t+1} = \omega_i \alpha(\mathbf{X}_t) + \beta'_{i,t} \mathbf{f}_{t+1} + \epsilon_{i,t+1} + \delta_i \epsilon_{t+1} \\ r_{t+1} = \alpha(\mathbf{X}_t) + \beta'_t \mathbf{f}_{t+1} + \epsilon_{t+1} \\ \mathbf{f}_{t+1} = \mathbf{c} + \mathbf{u}_{t+1} \end{cases} \quad (1.4)$$

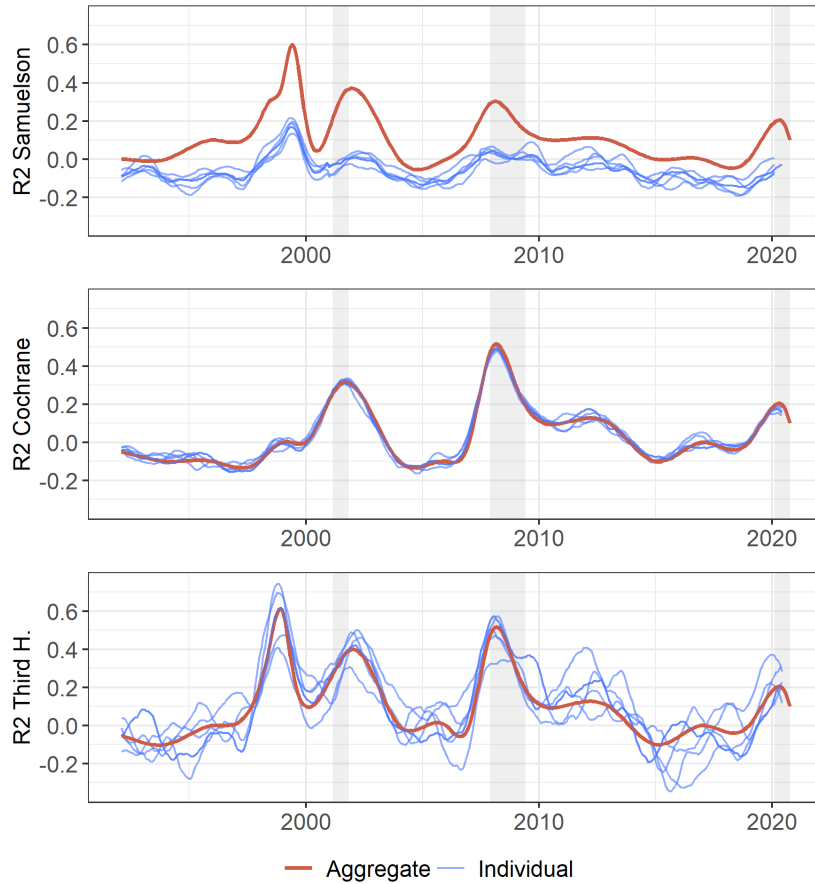
For illustrative purposes, we highlight in the top panel of Figure 1.1 how predictability should behave according to H_1 . In that setting, return predictability only comes from the inefficient component of returns, $\alpha(\mathbf{X}_t)$. As $\alpha(\mathbf{X}_t)$ is mixed with unpredictable news ($\epsilon_{i,t+1}$) at the stock-level, micro-predictability (in blue) should be lower than macro-predictability (in red). Additionally, we consider here that markets are inefficient in times of irrational exuberance (Shiller, 2015) or during downturns as the proportion of noise traders may be especially high in recessions (Veldkamp, 2005). Accordingly, macro-predictability should peak during the late 90s dotcom-bubble, or during the Great Financial Crisis of 2008 (grey bars figure NBER US recessions).

1.3.2 H_2 , Cochrane's view

Our second hypothesis dwells on Cochrane (2008), and assumes return predictability in the absence of market inefficiencies. Consequently, we consider here that $\alpha(\mathbf{X}_t) = \alpha_i(\mathbf{X}_t) = 0$.

²In other words, micro-returns are assumed to be essentially driven by "individual stories" (Jung and Shiller, 2005), whereas macro-returns are more affected by aggregate inefficiencies. More formally it would mean that the variance of the unpredictable factors of micro-returns ($\epsilon_{i,t+1} + \delta_i \epsilon_{t+1} + \beta'_{i,t} \mathbf{u}_{t+1}$) dominates the variance of the predictable part ($\omega_i \alpha(\mathbf{X}_t)$). This is less true for the corresponding factors of macro-returns, respectively $\beta'_t \mathbf{u}_{t+1} + \epsilon_{t+1}$ and $\alpha(\mathbf{X}_t)$.

Figure (1.1) Hypothetical Micro- and Macro-Predictability according to the different views



On the different graphs are represented the hypothetical macro- (in red) and micro- (in blue) predictability according to the three views outlined in Section 1.3. The graph is for illustrative purposes only and is not the result of an econometric estimation. The metric used is the out-of-sample R^2 , later detailed in Section 1.4.2, that can take negative values. The grey vertical bands figure the NBER US recession dates.

On the reverse, return predictability stems from time variation in expected returns, that is from the predictability of the risk factors: $\mathbf{f}_{t+1} = g(\mathbf{X}_t) + \mathbf{u}_{t+1}$. In this setting, expected returns vary with risk aversion along the business cycles, for instance if investors fear to fall short on their consumption targets during downturns (Campbell and Cochrane, 1999). If at time t , a variable like the dividend yield is able to spot changes in contemporaneous risk aversion, and

thus changes in expected returns, it can contain predictive content for future returns³.

We then have for H_2 the following system:

$$\begin{cases} r_{i,t+1} = \beta'_{i,t} \mathbf{f}_{t+1} + \epsilon_{i,t+1} + \delta_i \epsilon_{t+1} \\ r_{t+1} = \beta'_t \mathbf{f}_{t+1} + \epsilon_{t+1} \\ \mathbf{f}_{t+1} = g(\mathbf{X}_t) + \mathbf{u}_{t+1} \end{cases} \quad (1.5)$$

Here, micro- and macro-predictability are influenced by the same phenomenon: the forecastability of \mathbf{f}_{t+1} . As such, they should evolve in similar manners, although some differences may subsist depending on the values of $\beta_{i,t}$ and β_t , and on the realizations of $\epsilon_{i,t+1}$ and ϵ_{t+1} . This point is illustrated by the common trend in micro- (blue lines) and macro-predictability (red line) in the middle panel of Figure 1.1. Additionally, current returns may especially be influenced by expected returns during downturns, since expected returns are more volatile in recessions (Henkel et al., 2011). Therefore, as underlined on Figure 1.1, both micro- and macro-predictability should behave in a counter-cyclical way, and rise in bad times.

1.3.3 H_3 , Third view

Eventually, between the two precedent polar cases, the third view assumes that micro-returns can be *both* influenced by aggregate inefficiencies and by idiosyncratic mispricing, e.g. localized bubbles or specific illiquidity issues.

Leaning back on the previous representation, predictability could therefore emerge from “alpha”-predictability (aggregate or individual inefficiencies, $\alpha_i(\mathbf{X}_t)$ and $\alpha(\mathbf{X}_t)$), or from “beta”-predictability (due to time variation in expected returns, in line with H_2). If we also assume that the $\alpha_i(\mathbf{X}_t)$ and $\epsilon_{i,t+1}$ are diversified away at the aggregate level, H_3 would yield the exact same system of equations as System 1.3.

³Note that, in that case, return predictability is not a “free lunch”, investors have to take extra-risk to benefit from it (Kelly and Pruitt, 2013).

This view entails several implications illustrated on the bottom panel of Figure 1.1. First, depending notably on the relative importance of $\alpha_i(\mathbf{X}_t)$ and $\epsilon_{i,t+1}$, micro-predictability can be higher or lower than macro-predictability. Second, as these two variables are independently distributed, we would expect the average of micro-predictability indices across stocks to be similar to the macro-predictability series. Eventually, as macro-predictability can increase due to aggregate inefficiencies or to time variation in expected returns, it can both peak during speculative bubble periods or during downturns.

1.4. Data and Methodology

We assess the relevance of the three hypotheses with an out-of-sample methodology that tries to encompass the major modelling approaches in the literature. Our analysis is focused on postwar US monthly excess returns (from September 1945 to October 2020), but can easily be extended to other datasets.

1.4.1 Stock Return Data

Throughout this study we investigate the predictability of excess returns, i.e. total returns minus a risk-free rate. We extract monthly postwar US returns from Kenneth French website. This implies that:

1. We evaluate stock return predictability over a market constituted by all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ. We take as a risk-free rate the one-month Treasury bill rate from the same source.
2. We label “aggregate returns” the excess returns of the overall stock market, and “individual returns” the excess returns of the 25 Fama-French portfolios formed on Size and Book-to-Market.

Furthermore, we use supplementary variables as exogenous predictors in Section 1.4.2, or as

covariates in the interpretative regressions of Section 1.5.2. Their collections and their constructions are more thoroughly detailed in Appendix 1.A.2.

1.4.2 Constructing Raw Predictability Metrics

We present here our methodology to gauge the “raw predictability” of stock returns. We call raw predictability our mere ability to forecast future returns compared to a benchmark. This estimate will then be disentangled between alpha- and beta-predictability in Section 1.4.3.

As underlined in Section 1.2, an extensive number of models has been used in the return predictability literature. Besides, return-forecasting suffers from an elevated model instability as the popularity of performing approaches eradicates the predictive pattern they are based upon (Timmermann, 2018). We therefore adopt here an agnostic view, and centre our analysis on the estimation of $K = 23$ model types. These latter cover classic econometric models, forecast averaging methods, factor modelling approaches and Machine Learning tools. They are exhaustively described in Table 1.A.1.

We evaluate return predictability with the out-of-sample R^2 of Campbell and Thompson (2008), a metric widely used in the literature (Welch and Goyal, 2008, Moench and Stein, 2021). This indicator documents how well a model performs compared with the prevailing mean as a benchmark. More formally, given \bar{r}_t the prevailing mean of aggregate or individual returns from $t - L + 1$ to t , r_{t+1}^k the forecast of r_{t+1} of model k based on variables running from $t - L + 1$ to t , the out-of-sample R^2 for model k is defined as:

$$R_{os,k,t}^2 = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^k)^2}{(r_{i+1} - \bar{r}_i)^2} \quad (1.6)$$

In line with Timmermann (2008), we use a rolling window estimation of length $L = 120$ months, and an averaging period for $R_{os,k,t}^2$ of length $n = 36$ months. Our model-selection strategy proceeds as follow:

First, given a specific series of aggregate or individual returns $\{r_{t+1}\}_{t=0}^{T-1}$, we evaluate the different K models on a rolling window of length L . For each model m , we thus obtain a series of out-of-sample forecast: $\{r_{t+1}^m\}_{t=L}^{T-1}$ ⁴.

Second, again for each model k , we compute the corresponding $R_{os,k,t}^2$ at each point in time from $L + n + 1$ to T .

Eventually, as in pseudo-real time strategies, we choose the model with the best average out-of-sample $R_{os,k,t}^2$ over the previous estimation period to perform the next-period forecast⁵. We can therefore build a series of final out-of-sample predictions $\{r_{t+1}^f\}_{t=L+n}^{T-1}$, where, potentially, at each point in time a different model is chosen for the final forecast. From the latter series, we can then construct our final metric of raw R^2 for r_{t+1} : $\{R_{os,t}^2\}_{t=L+n+1}^T$.

1.4.3 Disentangling the Sources of Predictability

Following the different hypotheses outlined in Section 1.3, return predictability can emerge from two different phenomenons: the exposure to predictable risk factors (\mathbf{f}_{t+1}) or to market inefficiencies ($\alpha(\mathbf{X}_t)$ and $\alpha_i(\mathbf{X}_t)$).

For each portfolio returns $r_{i,t+1}$, we compute the series of raw return predictability $R_{i,os,t}^2$ according to the methodology described in Section 1.4.2. In this section, to decompose this metric between the two sources of predictability, we extend the methodology proposed by [Rapach et al. \(2011\)](#).

We first build a “beta-pricing restricted” forecast of r_{t+1} : r_{t+1}^β . To that aim, we define as risk factors \mathbf{f}_{t+1} the factors of the Fama-French three factor model, also extracted from Kenneth French website. We obtain the risk factors forecasts, \mathbf{f}_{t+1}^f , with the exact same prediction algorithm detailed in Section 1.4.2. Then, in line with [Rapach et al. \(2011\)](#), we estimate the risk

⁴In line with [Timmermann \(2008\)](#), we apply a “sanity filter” to our forecasts. If a forecast exceeds any previous return of the estimation period (in absolute value) it is then replaced with a “no change” forecast. This type of filtering is common in the return predictability literature ([Elliott and Timmermann, 2013](#)).

⁵Note that for an estimation period running from $t - L + 1$ to t , we need previous forecasts from $t - L - n + 2$ to $t - L + 1$ so as to build $R_{os,k,t-L+1}^2$. This latter variable will then be used in the model-selection to predict r_{t+1} .

loadings $\hat{\beta}_t$ by regressing, over a rolling window and without constant, $\{r_s\}_{t-L+1}^t$ on $\{\mathbf{f}\}_{t-L+1}^t$.

We can eventually construct:

$$r_{t+1}^\beta = \hat{\beta}_t' \mathbf{f}_{t+1}^f \quad (1.7)$$

In other words, all predictability stemming from the exposure to time varying risk factors should be incorporated in the beta-pricing restricted forecast r_{t+1}^β . Any additional return predictability beyond this beta-predictability reflects the fact that $\alpha_i(\mathbf{X}_t) \neq 0$ or that $\alpha(\mathbf{X}_t) \neq 0$, and is therefore called the alpha-predictability.

We can thus represent the evolution over time of the beta-predictability and the alpha-predictability by decomposing the different $R_{i,os,t}^2$. To do so, we first compute the ‘‘beta- R^2 ’’: $R_{i,\beta,t}^2$. This metric documents the difference in predictive ability between the beta-pricing restricted forecast and the prevailing mean:

$$R_{i,\beta,t}^2 = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^\beta)^2}{(r_{i+1} - \bar{r}_i)^2} \quad (1.8)$$

We then gauge the performance of the unrestricted forecast (r_{t+1}^f) compared to the beta-pricing restricted forecast (r_{t+1}^β) by computing the ‘‘alpha- R^2 ’’: $R_{i,\alpha,t}^2$. This latter assesses the extra-predictability that can be gained beyond the exposition to predictable risk factors:

$$R_{i,\alpha,t}^2 = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^f)^2}{(r_{i+1} - r_{t+1}^\beta)^2} \quad (1.9)$$

In line with [Rapach et al. \(2011\)](#), we can show that:

$$R_{i,os,t}^2 = R_{i,\alpha,t}^2 + R_{i,\beta,t}^2 - R_{i,\alpha,t}^2 * R_{i,\beta,t}^2 \quad (1.10)$$

Given that levels out-of-sample R^2 are particularly low in return forecasting exercises, we can therefore omit the cross-product and write:

$$R_{i,os,t}^2 \sim R_{i,\alpha,t}^2 + R_{i,\beta,t}^2 \quad (1.11)$$

In other words, looking at raw macro- and micro-predictability, $R_{os,t}^2$ and $R_{i,os,t}^2$ is helpful to discriminate between the three different hypotheses of Section 1.3. But analyzing more closely the behaviours of $R_{i,\alpha,t}^2$ and $R_{i,\beta,t}^2$ enables to evaluate whether the two sources of predictability are indeed at play in the sample⁶.

1.5. Empirical Results

This section first describes the raw predictability results over time, from both a micro- and a macro-perspective. It then outlines the decomposition of the raw predictability series between the alpha- and the beta-predictability, as well as the interpretation of the latter.

1.5.1 Micro- and Macro- Raw Predictability

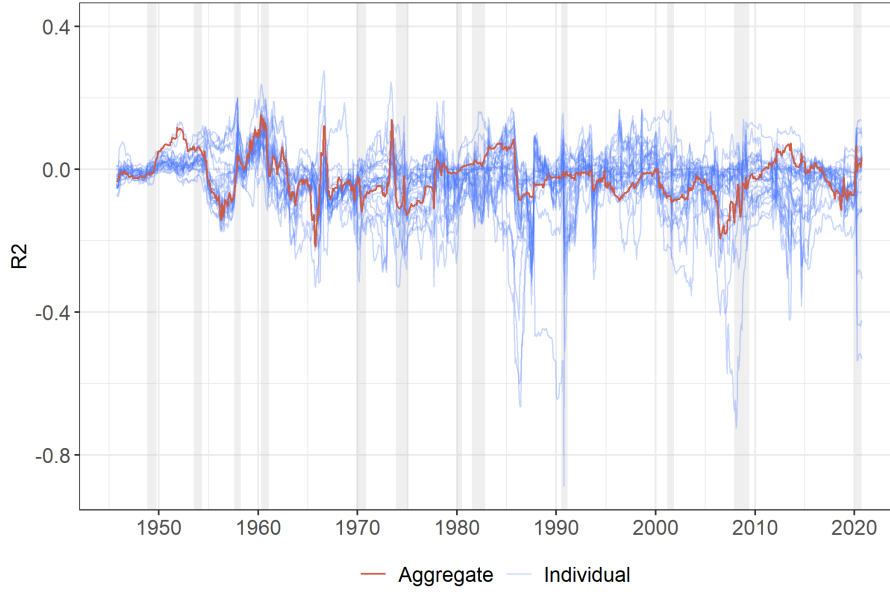
We represent on Figure 1.2 the raw predictability metrics for portfolio-returns ($R_{i,os,t}^2$, in blue) and for aggregate returns ($R_{os,t}^2$, in red). The 25 $R_{i,os,t}^2$ series are also plotted separately on Figure 1.A.1 of Appendix 1.A.3.

Several findings emerge from Figure 1.2 that help to discriminate between the three hypotheses of Section 1.3. First micro-predictability is not structurally lower than macro-predictability. This result invalidates the main assumption of H_1 , Samuelson's view, that macro-returns are more affected by market inefficiencies compared to micro-returns.

Second, micro-predictability does not seem to follow the exact same behaviour as the macro-

⁶Note that $R_{i,\alpha,t}^2$ is not necessarily positive. Theory-driven forecasts (such as r_{t+1}^β) may perform better than unrestricted forecasts (here r_{t+1}^f) in out-of-sample comparisons (Rapach et al., 2011).

Figure (1.2) Micro- and Macro-Raw Predictability series, over time



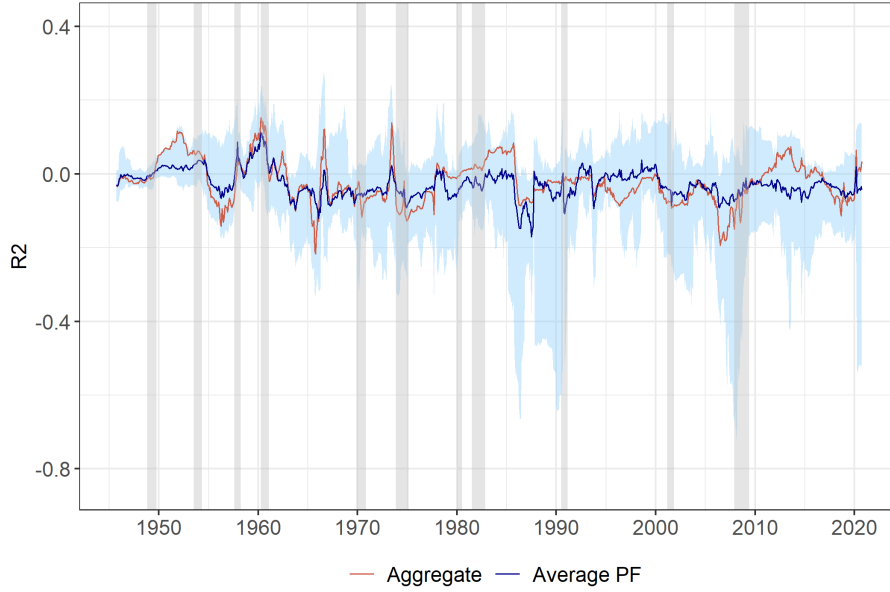
On the different graphs are represented the macro- ($R_{os,t}^2$, in red) and micro- ($R_{i,os,t}^2$, in blue) raw predictability indices according to the methodology outlined in Section 1.4.2. The metric used is the out-of-sample R^2 , also detailed in Section 1.4.2, that can take negative values. The grey vertical bands figure the NBER US recession dates.

predictability series. Although common factors are present in the micro-predictability series (as analyzed in Section 1.5.2), we notice that $R_{i,os,t}^2$ is sometimes significantly lower or higher than $R_{os,t}^2$. This finding contradicts H_2 (Cochrane's view) according to which micro- and macro-predictability should behave similarly.

Eventually, two observations appear to corroborate the last hypothesis (H_3 , the "third view"). First, we remark on Figure 1.2 that the variances of $R_{i,os,t}^2$ are considerably higher than for $R_{os,t}^2$. Second we plot on Figure 1.3 the average of the micro-predictability series over the I different portfolios ($\overline{R}_{i,os,t}^2 \equiv I^{-1} \sum_{1 \leq i \leq I} R_{i,os,t}^2$, in dark blue) along the macro-predictability series ($R_{os,t}^2$, in red).

We observe that pooling the different $R_{i,os,t}^2$ results in a time series that is significantly more correlated with $R_{os,t}^2$ than the individual micro-predictability series. Both of these findings are in line with the implications of H_3 . In this setting, micro-predictability is affected upward by

Figure (1.3) Macro-Raw Predictability series and averaged Micro-Raw Predictability, over time



On the graph are represented the macro-raw predictability ($R_{os,t}^2$, in red) and the average micro-raw predictability across portfolios ($\overline{R}_{i,os,t}^2$, in blue). The details of the methodology are outlined in Section 1.4.2. The light blue area represents the gap between the minimum and the maximum values taken by the different portfolio-raw predictability series ($R_{i,os,t}^2$). The metric used is the out-of-sample R^2 , also detailed in Section 1.4.2, that can take negative values. The grey vertical bands figure the NBER US recession dates.

idiosyncratic inefficiencies ($\alpha_i(\mathbf{X}_t)$) and downward by idiosyncratic news ($\epsilon_{i,t+1}$). As these two components are centered around 0, they do not translate to macro-returns. Accordingly, macro-predictability should be less volatile than micro-predictability, whereas the average of the micro-predictability series should mimic the evolution of the macro-predictability series. We find both of these results on Figures 1.2 and 1.3. The results outlined in this section, regarding the means and the variances of the macro- and micro-predictability series, as well as regarding the strong correlation of $\overline{R}_{i,os,t}^2$ with respect to $R_{os,t}^2$, are detailed in Figure 1.A.2 of Appendix 1.A.4^{7 8}.

⁷Note that all the aforementioned results concerning $R_{i,os,t}^2$ and $R_{os,t}^2$ cannot be explained by the variances of the input returns $r_{i,t+1}$ and r_{t+1} . We plot on Figures 1.A.5 and 1.A.6 of Appendix 1.A.4 the standard deviations of stock returns against either the level or the variance of their corresponding raw predictability indices. For both graphs the relationships between these variables appear weak at best.

⁸In Appendix 1.A.5, we take into account the uncertainty regarding the coefficients with bootstrapping techniques.

On this specific exercise, we underline two additional results. First, a natural question regarding Figures 1.2 and 1.3 is whether or not an investor would have been able to make money out of these forecasting exercises. Similar to Timmermann (2008), we find that both $R_{os,t}^2$ and $R_{i,os,t}^2$ (for all portfolios) are negative on average. This means that, on the overall estimation period, an investor would not have been able to build a profitable strategy based on our forecasts. However, in line with Timmermann (2008) and Farmer et al. (2022), returns appear predictable at specific time periods. In our case, at the macro-level, $R_{os,t}^2$ is positive on average during two decades: the 50s and the 80s (amounting to, respectively, 2.0% and 0.3%). Although relatively small, Campbell and Thompson (2008) showed that R^2 of small magnitudes may translate to a substantial gain improvements for an investor with mean-variance preferences. Following their rule of thumb for the macro-returns, we find that an investor with a coefficient of risk aversion of 3 could have improved the returns of his portfolio by 80 bp in the 50s and by 10 bp in the 80s. At the micro-level, evidence of return predictability appear more mixed, with most of $R_{i,os,t}^2$ being negative on average on these two decades. However, for the portfolios with the highest $R_{i,os,t}^2$, the same calculations imply that a similar investor would have been able to increase his returns by 230 bp and 74 bp over these two periods.

Second, we investigate whether the results of Figures 1.2 and 1.3 remain the same if we consider individual stock returns instead of portfolio returns to assess micro-predictability. To do so, we retrieve more than 100 individual stock returns from Refinitiv starting from January 1986 to October 2020⁹. The results of this exercise are depicted on Figure 1.A.4 in Appendix 1.A.6. It can be seen that the main results remain unchanged when we gauge micro-predictability with individual stock returns. Here also we find that the variances of $R_{i,os,t}^2$ are considerably higher than for $R_{os,t}^2$, whereas pooling the different $R_{i,os,t}^2$ results in a time series that is significantly

We can notice on Figure 1.A.3 that the mean of $R_{os,t}^2$ is similar to the means of $R_{i,os,t}^2$, while the standard deviation of $R_{os,t}^2$ appears significantly lower than the standard deviations of $R_{i,os,t}^2$.

⁹To select the stocks, based on Refinitiv data, we retrieve all the companies that belonged to the S&P 500 for at least a month, from January 2008 until October 2020. We then try to strike a balance between the number of individual stocks that we consider and the availability of their returns over a long period. Overall our samples of individual stock returns covers 110 companies from January 1986 to October 2020

more correlated with $R_{os,t}^2$ than the individual micro-predictability series.

1.5.2 Alpha- and Beta-Predictability

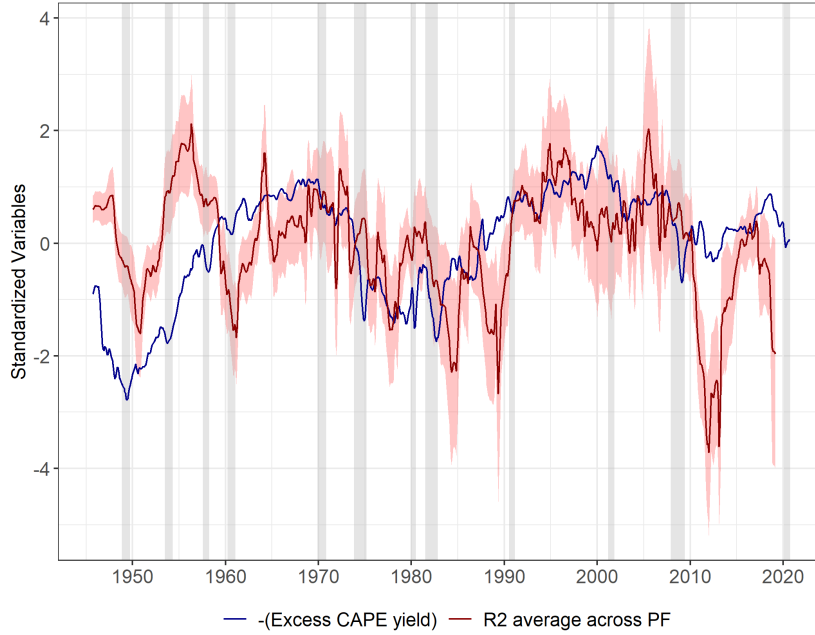
Building Alpha- and Beta-Predictability

The findings highlighted with Figures 1.2 and 1.3 enabled to discard the first two hypotheses: Samuelson's and Cochrane's views. On the reverse, the third view, H_3 , seems to fit well with the behaviours of the micro- and macro- raw predictability series outlined above. However, H_3 has also implications regarding the time variation of macro-predictability. Since macro-predictability is influenced by alpha- and beta-predictability, it should be significant both in times of elevated market inefficiencies and during economic downturns. As such, Figures 1.2 and 1.3 do not help disentangling these two potential factors, since drops in alpha-predictability may counterbalance rises in beta-predictability (and the reverse).

We therefore attempt in this section to better understand the sources of variation of macro-predictability over time. To do so, we take as a starting point the individual portfolio returns ($r_{i,t+1}$) that we use to estimate the individual series of alpha-predictability ($R_{i,\alpha,t}^2$) and beta-predictability ($R_{i,\beta,t}^2$) with the methodology detailed in Section 1.4.3. Eventually, we represent on Figures 1.4 and 1.5 the behaviours of, respectively, the pooled series $\overline{R}_{i,\alpha,t}^2 \equiv I^{-1} \sum_{1 \leq i \leq I} R_{i,\alpha,t}^2$ and $\overline{R}_{i,\beta,t}^2 \equiv I^{-1} \sum_{1 \leq i \leq I} R_{i,\beta,t}^2$ ¹⁰.

We draw several conclusions from these figures. First remember that, in line with H_3 , we expect $\overline{R}_{i,\alpha,t}^2$ to rise in periods of market exuberance, and $\overline{R}_{i,\beta,t}^2$ to increase during recessions. In order to better visualize their time variations, we plot along $\overline{R}_{i,\alpha,t}^2$ and $\overline{R}_{i,\beta,t}^2$ the opposite of the "Excess CAPE yield" (ECY, built as the inverse of the CAPE ratio minus a risk-free rate) and

¹⁰On Figures 1.4 and 1.5 we center the R^2 metrics around the mid-point of their estimation periods. In other words, whereas in Section 1.4.2 we had $R_{os,m,t}^2 = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^m)^2}{(r_{i+1} - \bar{r}_i)^2}$, here we consider that $R_{os,m,t}^2 = 1 - \sum_{i=t-n/2}^{t-1+n/2} \frac{(r_{i+1} - r_{i+1}^m)^2}{(r_{i+1} - \bar{r}_i)^2}$, with n an even number. We do this as, for the out-of-sample predictive algorithm, we need all the previous forecasting errors to perform our model selection. However, for interpretative purposes, building the R^2 metrics with only past data will tend to artificially shift the series with respect to the other external variables.

Figure (1.4) $\overline{R^2}_{i,\alpha,t}$ and US ECY, over time

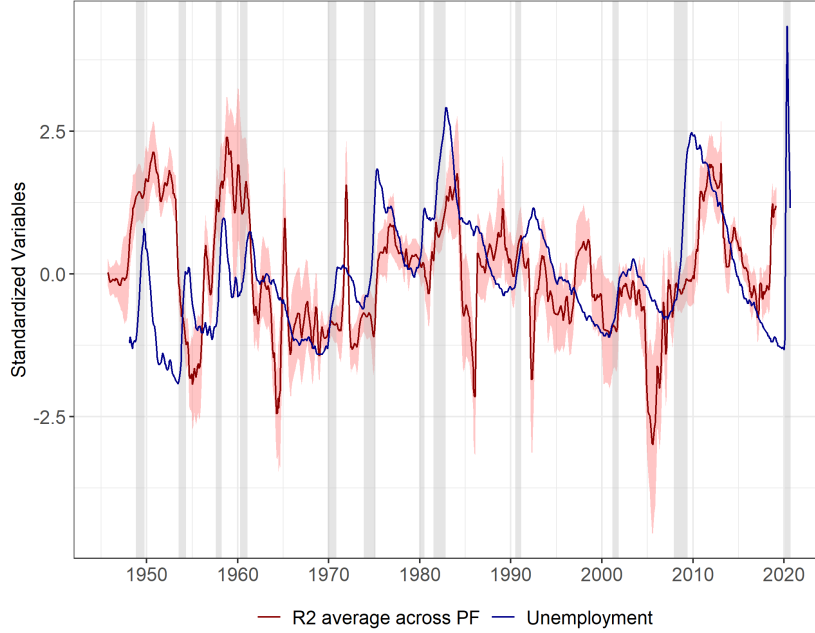
On the graph are represented the average across portfolios of the alpha-predictability series ($\overline{R^2}_{i,\alpha,t}$, in red) and the US Excess CAPE yield multiplied by -1 (in blue). These monthly series have been standardized to fit in the same graph, and, for visual purposes, they have been smoothed over a 3-month period. Raw series of $\overline{R^2}_{i,\alpha,t}$ are yet available in the Figure 1.A.7 of Appendix 1.A.7. The red area figures the cross-sectional dispersion around $\overline{R^2}_{i,\alpha,t}$ (+/-1 standard deviation). The metric used is the out-of-sample alpha-predictability $R^2_{i,\alpha,t}$, detailed in Section 1.4.3, that can take negative values. The grey vertical bands figure the NBER US recession dates.

the Unemployment rate. The former has been advocated to be a good metric of market effervescence¹¹ (Shiller et al., 2020), while the latter stands as an intuitive variable to spot changes in the business cycle. Regarding the behaviour of $\overline{R^2}_{i,\alpha,t}$ on Figure 1.4, the series appears positively correlated with the opposite of the US ECY. As expected, $\overline{R^2}_{i,\alpha,t}$ is relatively high in periods of market booms. These periods include notably the “Kennedy-Johnson peak” (Shiller, 2015) around 1966, the dotcom bubble of the late 90s and finally the period preceding the Great Financial Crisis of 2007. As for $\overline{R^2}_{i,\beta,t}$, the series appears also positively associated with the US

¹¹Adjusting likewise the CAPE ratio enables to take into account the role of the fall in risk-free rates for stock valuations in the recent years. In line with Chatelais and Stalla-Bourdillon (2020) we multiply the ECY by -1 throughout the rest of this paper, so that an increase in this metric reflects stronger stock valuations (with respect to bonds).

Unemployment rate. It rises during economic downturns, for example throughout the 1960-61 recession, in the neighbourhood of the 1973- oil shock, along the Great Financial Crisis or during the recent Covid crisis¹².

Figure (1.5) $\overline{R^2}_{i,\beta,t}$ and US Unemployment rate, over time



On the graph are represented the average across portfolios of the beta-predictability series ($\overline{R^2}_{i,\beta,t}$, in red) and the US Unemployment rate (in blue). These monthly series have been standardized to fit in the same graph, and, for visual purposes, they have been smoothed over a 3-month period. Raw series of $\overline{R^2}_{i,\beta,t}$ are yet available in the Figure 1.A.8 of Appendix 1.A.7. The red area figures the cross-sectional dispersion around $\overline{R^2}_{i,\beta,t}$ (± 1 standard deviation). The metric used is the out-of-sample beta-predictability $R^2_{i,\beta,t}$, detailed in Section 1.4.3, that can take negative values. The grey vertical bands figure the NBER US recession dates.

Second, the red areas surrounding $\overline{R^2}_{i,\alpha,t}$ and $\overline{R^2}_{i,\beta,t}$ figure the cross-sectional dispersion of alpha- and beta-predictability across portfolios. We thus notice that the series of $R^2_{i,\alpha,t}$ are way more dispersed than the series of $R^2_{i,\beta,t}$. This result is quite intuitive as well: in line with H_3 , alpha-predictability depends on the importance of both idiosyncratic and aggregate factors,

¹²Note that, following the Covid-shock, all predictability appears to stem from the beta-predictability. This finding is in line with other recent studies, such as Gormsen and Koijen (2020). This latter argue that the apparent disconnection between the macroeconomic situations and the US stock market wasn't due to irrational investors' behaviours, but could be rationalized through the fall in long-term sovereign rates.

$\alpha_i(\mathbf{X}_t)$ and $\alpha(\mathbf{X}_t)$. On the reverse, beta-predictability should reflect a single phenomenon, the predictability of f_{t+1} . Therefore we should indeed observe more dispersion among the different $R_{i,\alpha,t}^2$ than for the different $R_{i,\beta,t}^2$.

These two findings appear in accordance with the implications of H_3 regarding either the timing of alpha-predictability and beta-predictability peaks, or the dispersion among portfolio returns for these series. However, to better assess the drivers of $\overline{R}_{i,\alpha,t}^2$ and $\overline{R}_{i,\beta,t}^2$ beyond pure visual examination, we turn to regression analysis in the next section.

Interpreting Alpha- and Beta-Predictability

According to the different implications of H_3 , three variable types may affect $\overline{R}_{i,\alpha,t}^2$ and $\overline{R}_{i,\beta,t}^2$. First, $\overline{R}_{i,\alpha,t}^2$ is supposed to increase during periods of either elevated market frictions, or of irrational exuberance. Conversely, following [Henkel et al. \(2011\)](#), $\overline{R}_{i,\beta,t}^2$ should especially be high during economic downturns. Thus, let $j \in \{\alpha, \beta\}$, we look at regressions of the form:

$$\overline{R}_{i,j,t}^2 = c_j + \gamma'_{IE,j} \mathbf{X}_{IE,j,t} + \gamma'_{FC,j} \mathbf{X}_{FC,t} + \gamma'_{RA,j} \mathbf{X}_{RA,t} + \epsilon_{j,t} \quad (1.12)$$

With $\mathbf{X}_{IE,t}$ spotting periods of irrational exuberance (valuation ratios or speculative bubble indicators), $\mathbf{X}_{FC,t}$ indicating financial constraints which prevent arbitrageurs from exploiting potential mispricings (stock return volatility, financial intermediary leverage) and $\mathbf{X}_{RA,t}$ following closely the business cycles (unemployment level).

H_3 has several implications for the signs of the different coefficients. If we assume that increases in $\mathbf{X}_{IE,t}$, $\mathbf{X}_{FC,t}$ and $\mathbf{X}_{RA,t}$ reflect an increase in market effervescence, an aggravation of financial constraints and a strengthening of economic activity, respectively, we would expect, in line with [Section 1.3](#), that $\gamma_{IE,\alpha} > 0$, $\gamma_{FC,\alpha} > 0$ and that $\gamma_{RA,\beta} < 0$. Furthermore, we would also expect that a tightening of financial conditions leaves beta-predictability unaffected as the latter shouldn't be influenced by market frictions. Eventually, we remain agnostic regarding the

link between economic expansions and alpha-predictability. Alpha-predictability can either be positively influenced by the former (if an improvement in macroeconomic conditions triggers investor's excessive enthusiasm) or negatively (if noise traders are especially present during recessions, [Veldkamp, 2005](#)). Therefore, we expect $\gamma_{FC,\beta}$ to be non-significant while we do not form any expectation regarding the sign of $\gamma_{RA,\alpha}$.

To test these implications on the US stock market, we first use for $\mathbf{X}_{IE,t}$ two different valuation ratios: the Excess CAPE yield, already described in Section 1.5.2, and the S&P 500 Price Earning Ratio. Additionally, we also look at survey variables to gauge market exuberance in the form of the U.S. One-Year Confidence Index of Yale university. Second, we consider for $\mathbf{X}_{FC,t}$ three different metrics to reflect funding constraints. The first one is stock return volatility, the second one the Baa-Aaa corporate bond spread and the third one the seasonally adjusted changes in U.S. broker-dealer leverage (LF_t , [Adrian et al., 2014](#)). Following [Farmer et al. \(2022\)](#), we take LF_t as proxy of funding constraints, since lower leverage is associated with a reduced availability of arbitrage capital. Eventually, for the business cycles variables, $\mathbf{X}_{RA,t}$, we take as a main proxy the US unemployment rate, but we also use the Consumer Sentiment Index from the University of Michigan in the robustness checks of Section 1.6. These different covariates, as well as their originating sources are more precisely detailed in Table 1.A.3 of Appendix 1.A.2.

The regression results are presented in Tables 1.1 and 1.2. For the alpha-predictability, we notice in Table 1.1 that whatever the proxy for $\mathbf{X}_{IE,t}$, the associated coefficient $\gamma_{IE,\alpha}$ is significantly positive in the nine specifications outlined here. This finding suggests that alpha-predictability is particularly at play in times of elevated market effervescence. As for the business cycles variables, we observe that the corresponding slopes $\gamma_{RA,\alpha}$ are always significant and positive. This last result indicates that alpha-predictability tends to be especially high in times of bullish stock market combined with sound macroeconomic conditions. Conversely, the mechanism outlined by [Veldkamp \(2005\)](#) does not seem to play any role here. Eventually, for all the different regressions, the coefficients $\gamma_{FC,\alpha}$ are either non-significant or (significantly) positive. Thus, although

financial constraints' coefficients have most of the time the expected sign, these variables appear to have only a secondary importance in the drivers of alpha-predictability.

Table (1.1) Regression results for the Alpha-Predictability

	<i>Dependent variable:</i>								
	Alpha-predictability: $\overline{R^2}_{i,\alpha,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$-ecy_t$	0.240*** (0.068)	0.443*** (0.130)	0.243*** (0.069)						
pe_t				0.001*** (0.0002)	0.001*** (0.0001)	0.001*** (0.0002)			
$Yale_t$							0.002*** (0.0004)	0.002*** (0.0003)	0.002*** (0.0004)
$-unemp_t$	0.007*** (0.002)	0.006** (0.003)	0.007*** (0.002)	0.008*** (0.002)	0.010*** (0.002)	0.007*** (0.002)	0.013*** (0.004)	0.015*** (0.004)	0.014*** (0.004)
$vol_{1,t}$	-0.00001 (0.001)			-0.001 (0.001)			0.003*** (0.001)		
$-LF_t$		0.0001 (0.0002)			-0.00002 (0.0002)			0.001* (0.001)	
Baa_t			0.003 (0.008)			-0.004 (0.008)			0.021*** (0.007)
Const.	0.032*** (0.010)	0.032*** (0.012)	0.032*** (0.010)	0.018* (0.010)	0.033*** (0.011)	0.017* (0.010)	-0.101*** (0.033)	-0.050* (0.027)	-0.113*** (0.034)
Obs.	856	597	856	856	597	856	214	214	214
R ²	0.175	0.265	0.175	0.166	0.250	0.166	0.363	0.427	0.380
Adj. R ²	0.172	0.261	0.172	0.163	0.246	0.163	0.354	0.418	0.371

Note:

*p<0.1; **p<0.05; ***p<0.01

On the table are represented the different regression results with $\overline{R^2}_{i,\alpha,t}$ as a predicted variable. t -statistics have been computed using Newey-West standard errors. Variables are rearranged so that an increase in $X_{IE,t}$, $X_{FC,t}$ and $X_{RA,t}$ reflects, respectively, a surge in market effervescence, an aggravation of financial constraints and a strengthening of economic activity.

Regarding the beta-predictability, we remark in Table 1.2 that, as expected, a decrease in economic activity is related to an increase in beta-predictability ($\gamma_{RA,\beta} < 0$) for all nine regressions, in line with Henkel et al. (2011). Similarly, beta-predictability seems to coincide with bearish financial markets, as the coefficients $\gamma_{IE,\beta}$ are significantly negative irrespective of the cho-

sen metric. Eventually, again as expected, financial constraints do not seem to play a role in determining the level of beta-predictability, as coefficients $\gamma_{FC,\beta}$ are non-significant across all specifications of Table 1.2.

Table (1.2) Regression results for the Beta-Predictability

	<i>Dependent variable:</i>								
	Beta-predictability: $\overline{R^2}_{i,\beta,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$-ecy_t$	-0.582*** (0.101)	-0.214* (0.125)	-0.598*** (0.102)						
pe_t				-0.001*** (0.0003)	-0.0004*** (0.0001)	-0.001*** (0.0003)			
$Yale_t$							-0.002*** (0.0005)	-0.002*** (0.0004)	-0.002*** (0.0005)
$-unemp_t$	-0.005*** (0.002)	-0.011*** (0.002)	-0.007*** (0.002)	-0.008*** (0.002)	-0.013*** (0.002)	-0.008*** (0.003)	-0.014*** (0.003)	-0.017*** (0.003)	-0.015*** (0.003)
$vol_{1,t}$	-0.0002 (0.001)			0.001 (0.001)			-0.001 (0.001)		
$-LF_t$		-0.0002 (0.0002)			-0.0001 (0.0002)			-0.001 (0.001)	
Baa_t			-0.013 (0.008)			-0.0003 (0.008)			-0.006 (0.009)
Const.	-0.067*** (0.011)	-0.094*** (0.012)	-0.067*** (0.011)	-0.042*** (0.016)	-0.093*** (0.011)	-0.042*** (0.016)	0.073* (0.038)	0.026 (0.032)	0.076** (0.038)
Obs.	856	597	856	856	597	856	214	214	214
R ²	0.234	0.302	0.242	0.109	0.305	0.108	0.395	0.468	0.395
Adj. R ²	0.231	0.298	0.240	0.106	0.301	0.105	0.386	0.460	0.387

Note:

*p<0.1; **p<0.05; ***p<0.01

On the table are represented the different regression results with $\overline{R^2}_{i,\beta,t}$ as a predicted variable. t -statistics have been computed using Newey-West standard errors. Variables are rearranged so that an increase in $X_{IE,t}$, $X_{FC,t}$ and $X_{RA,t}$ reflects, respectively, a surge in market effervescence, an aggravation of financial constraints and a strengthening of economic activity

The results of Tables 1.1 and 1.2 bring new additional evidence in favor of H_3 : all the different coefficients exhibited the expected signs according to this hypothesis. The findings highlighted in this section as well as in Section 1.5.1 corroborate the two main ideas of this paper: First,

that there is indeed a diversification effect of efficient and inefficient individual factors when we compare micro-returns to macro-returns. Second, that regarding more specifically the drivers of macro-predictability, both types of return predictability, alpha- and beta-predictability, seem to play at the same time in our dataset. This last finding contrasts with the return predictability literature, where previous studies tended to oppose these two mechanisms.

1.6. Robustness checks

We provide here different robustness checks for the results outlined in Section 1.5.2.

First, to build the alpha- and the beta-predictability indices, we relied on the 3 factor-model of Fama and French (1993) as proxies for the risk factors f_{t+1} , namely the excess return on the market, the size factor and the value factor. On Figures 1.A.7 and 1.A.8 of Appendix 1.A.7, we also plotted the resulting $\overline{R^2}_{i,\alpha,t}$ and $\overline{R^2}_{i,\beta,t}$ whether we rely on the 1-factor (in green), the 3-factor (in red) or the 5-factor (in blue) Fama-French models¹³. We notice on both figures that, despite some discrepancies for the 1-factor model indices, the different metrics behave in a very similar way. These similarities are noticeable whether we look at the pooled series ($\overline{R^2}_{i,\alpha,t}$ and $\overline{R^2}_{i,\beta,t}$) or at the dispersion around the latter (the shaded areas on Figures 1.A.7 and 1.A.8).

Second, we provide on Table 1.A.4 of Appendix 1.A.8 additional regression results in line with our analysis of Section 1.5.2. We use as an alternative business cycle variable the Consumer Sentiment Index from the University of Michigan, and as a supplementary financial friction proxy a different metric of stock market volatility (computed by estimating a GARCH(1,1) on daily stock returns instead of taking the monthly average of squared returns). We thus notice in Table 1.A.4 that these modifications leave the main results unchanged: $\gamma_{IE,\alpha}$ and $\gamma_{RA,\beta}$ are still significantly positive and negative, $\gamma_{FC,\alpha}$, $\gamma_{FC,\beta}$ non-significant, and $\gamma_{RA,\alpha}$ positive (although

¹³The two last factors "Robust Minus Weak" and "Conservative Minus Aggressive" are also extracted from Kenneth French website. Due to their limited availability, the $\overline{R^2}_{i,\alpha,t}$ and $\overline{R^2}_{i,\beta,t}$ for the 5-factor model start later than for the 1-factor or 3-factor models.

not significantly).

1.7. Conclusion

Based on US postwar data, we manage in this paper to discriminate between three opposite hypotheses regarding the behaviours of micro- and macro-stock return predictability. Overall, by looking at raw predictability metrics, we find that our results are consistent with a model (H_3) that lies in-between Samuelson's and Cochrane's views (H_1 and H_2). Indeed, micro-predictability series do not appear to be structurally higher or lower than macro-predictability indices, but tend to "bounce" around the latter. Furthermore, pooling micro-predictability series across portfolios yields an index that is significantly more correlated with the macro-predictability metric. All these observations corroborate an hypothesis where individual returns are mostly affected by idiosyncratic efficient and inefficient components, but also by common factors. If the former are diversified away at the index-level, we should indeed observe more variability in micro-predictability series, but also an averaged micro-predictability index that mimic the macro-predictability series.

Additionally, by extending over time the framework by [Rapach et al. \(2011\)](#), we are able to disentangle the two sources of return predictability, the alpha- and the beta-predictability. Here again, our results underpin an intermediate view where return predictability is both affected by these two mechanisms. As a matter of fact, our two estimated indices match the expected theoretical patterns: alpha-predictability rises in period of market effervescence whereas beta-predictability increases during downturns. This last finding enables to reconcile two opposite blocks of the literature: whereas previous papers tend to stress a specific source of predictability ([Farmer et al., 2022](#), [Dangl and Halling, 2012](#)), our results suggest that the two phenomena are at play in our sample.

Eventually, we argue that our estimated alpha-predictability index ($\overline{R^2}_{i,\alpha,t}$) constitutes a theo-

retically based and easily updatable series to assess periods of irrational exuberance in real time. Along with other metrics of speculative bubbles (Shiller et al., 2020, Blot et al., 2018), it can be used for financial stability purposes to gauge potential overvaluations on the stock market.

Bibliography

- Adrian, T., E. Etula, and T. Muir. 2014. Financial Intermediaries and the Cross-Section of Asset Returns. *Journal of Finance* 69:2557–2596.
- Aiolfi, M., and A. Timmermann. 2006. Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics* 135:31–53.
- Avramov, D. 2004. Stock return predictability and asset pricing models. *Review of Financial Studies* 17:699–738.
- Avramov, D., and T. Chordia. 2006. Predicting stock returns. *Journal of Financial Economics* 82:387–415.
- Baetje, F. 2018. Does a lot help a lot? Forecasting stock returns with pooling strategies in a data-rich environment. *Journal of Forecasting* 37:37–63.
- Bai, J., T. Philippon, and A. Savov. 2016. Have financial markets become more informative? *Journal of Financial Economics* 122:625–654.
- Binsbergen, J. H. V., and R. S. Koijen. 2010. Predictive regressions: A present-value approach. *Journal of Finance* 65:1439–1471.
- Blot, C., P. Hubert, and F. Labondance. 2018. Monetary Policy and Asset Price Bubbles. *Sciences Po OFCE Working Paper* .
- Campbell, J., and J. Ammer. 1993. What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns. *The Journal of Finance* 48:3–37.

- Campbell, J. Y. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18:373–399.
- Campbell, J. Y. 1991. A Variance Decomposition for Stock Returns. *Economic Journal* 101:157–79.
- Campbell, J. Y., and J. H. Cochrane. 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107:205–251.
- Campbell, J. Y., and R. J. Shiller. 1988. Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance* 43:661–676.
- Campbell, J. Y., and S. B. Thompson. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–1531.
- Chatelais, N., and A. Stalla-Bourdillon. 2020. What are the factors behind current high stock market valuations? *Banque de France Blog* .
- Chinco, A., A. D. Clark-Joseph, and M. Ye. 2019. Sparse Signals in the Cross-Section of Returns. *Journal of Finance* 74:449–492.
- Cochrane, J. H. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21:1533–1575.
- Cohen, R. B., C. Polk, and T. Vuolteenaho. 2003. The Value Spread. *Journal of Finance* 58:609–642.
- Dangl, T., and M. Halling. 2012. Predictive regressions with time-varying coefficients. *Journal of Financial Economics* 106:157–181.
- Davila, E., and C. Parlatore. 2018. Identifying Price Informativeness. *National Bureau of Economic Research Working Paper Series* .
- Elliott, G., and A. Timmermann. 2013. *Handbook of economic forecasting*. Elsevier.

- Fama, E. F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* 25:383.
- Fama, E. F., and K. R. French. 1988. Permanent and Temporary Components of Stock Prices. *Journal of Political Economy* 96:246–273.
- Fama, E. F., and K. R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.
- Farboodi, M., A. Matray, L. Veldkamp, and V. Venkateswaran. 2020. Where Has All the Data Gone? *National Bureau of Economic Research Working Paper Series* .
- Farmer, L., L. Schmidt, and A. Timmermann. 2022. Pockets of predictability. *Journal of Finance, forthcoming* .
- Gormsen, N. J., and R. S. Kojen. 2020. Coronavirus: Impact on stock prices and growth expectations. *Review of Asset Pricing Studies* 10:574–597.
- Guidolin, M., D. G. McMillan, and M. E. Wohar. 2013. Time varying stock return predictability: Evidence from US sectors. *Finance Research Letters* 10:34–40.
- Henkel, S. J., J. S. Martin, and F. Nardari. 2011. Time-varying short-horizon predictability. *Journal of Financial Economics* 99:560–580.
- Hirshleifer, D., K. Hou, and S. H. Teoh. 2009. Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics* 91:389–406.
- Jung, J., and R. J. Shiller. 2005. Samuelson’s dictum and the stock market. *Economic Inquiry* 43:221–228.
- Kelly, B., and S. Pruitt. 2013. Market expectations in the cross-section of present values. *Journal of Finance* 68:1721–1756.

- Kong, A., D. Rapach, J. Strauss, and G. Zhou. 2011. Predicting market components out of sample: asset allocation implications. *Journal of Portfolio Management* .
- Kothari, S. P., J. Lewellen, and J. B. Warner. 2006. Stock returns, aggregate earnings surprises, and behavioral finance. *Journal of Financial Economics* 79:537–568.
- Lee, C. M., J. Myers, and B. Swaminathan. 1999. What is the intrinsic value of the dow? *Journal of Finance* 54:1693–1741.
- Lettau, M., and S. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56:815–849.
- Moench, E., and T. Stein. 2021. Equity Premium Predictability Over the Business Cycle. *Bundesbank Discussion Paper* .
- Pesaran, M. H. 2010. Predictability of asset returns and the efficient market hypothesis. In *Handbook of Empirical Economics and Finance*, pp. 281–312.
- Rapach, D. E., J. K. Strauss, J. Tu, and G. Zhou. 2011. Out-of-Sample Industry Return Predictability: Evidence from A Large Number of Predictors. *SMU Working Paper* 2-2011.
- Rapach, D. E., J. K. Strauss, J. Tu, and G. Zhou. 2019. Industry Return Predictability: A Machine Learning Approach. *The Journal of Financial Data Science* 1:9–28.
- Rapach, D. E., J. K. Strauss, and G. Zhou. 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23:821–862.
- Sadka, G., and R. Sadka. 2009. Predictability and the earnings-returns relation. *Journal of Financial Economics* 94:87–106.
- Samuelson, P. A. 1975. Lifetime portfolio selection by dynamic stochastic programming.

- Shiller, R. J. 2015. *Irrational exuberance: Revised and expanded third edition*. Princeton university press.
- Shiller, R. J., L. Black, and F. Jivraj. 2020. CAPE and the COVID-19 Pandemic Effect. *SSRN Electronic Journal* 3714737.
- Timmermann, A. . 2008. Elusive return predictability. *International Journal of Forecasting* .
- Timmermann, A. 2018. Forecasting methods in finance. *Annual Review of Financial Economics* 10:449–479.
- Veldkamp, L. L. 2005. Slow boom, sudden crash. *Journal of Economic Theory* 124:230–257.
- Vuolteenaho, T. 2002. What drives firm-level stock returns? *Journal of Finance* 57:233–264.
- Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455–1508.

Appendix

1.A. Appendix

1.A.1 List of estimated Models

With r_{t+1} the predicted variable (index or portfolio excess returns) and r_{t+1}^f the model-forecast.

Table (1.A.1) Estimated Models

Name	Model description	References
Model 1	<i>Simple Exponential Smoothing</i> $p_{t+1} = \alpha p_t + (1 - \alpha)r_t$ With $p_1 = r_1$	Timmermann
Model 2	<i>Double Exponential Smoothing</i> $p_{t+1} = \alpha(p_t + \lambda_{t-1}) + (1 - \alpha)r_t$ $\alpha_t = \beta(p_{t+1} - p_t) + (1 - \beta)\lambda_{t-1}$ With $p_1 = 0, f_2 = r_2$ and $\lambda_2 = r_2 - r_1$	Timmermann
Model 3	<i>Autoregressive Model (BIC)</i> $r_{t+1} = \alpha + \beta(L)r_t + u_t$ Number of lags chosen with the Bayesian Information Criterion	Timmermann
Model 4	<i>Autoregressive Model (AIC)</i> $r_{t+1} = \alpha + \beta(L)r_t + u_t$ Number of lags chosen with the Aikake Information Criterion	Timmermann
Model 5	<i>Smooth Transition Autoregressive Model 1</i> $r_{t+1} = \theta'_0 \eta_t d_t + \theta'_1 \eta_t + u_{t+1}$ $d_t = 1/(1 + \exp(\gamma_0 + \gamma_1(r_t - r_{t-6})))$ With $\eta_t = (1, r_t)'$	Timmermann
Model 6	<i>Smooth Transition Autoregressive Model 2</i> $r_{t+1} = \theta'_0 \eta_t d_t + \theta'_1 \eta_t + u_{t+1}$ $d_t = 1/(1 + \exp(\gamma_0 + \gamma_1 r_{t-3}))$ With $\eta_t = (1, r_t)'$	Timmermann
Model 7	<i>Neural net model 1</i> $r_{t+1} = \theta_0 + \sum_{i=1}^n \theta_i g(\beta'_i \eta_t) + u_{t+1}$ With g the logistic function, $\eta_t = (1, r_t, r_{t-1}, r_{t-2})'$ and $n = 2$	Timmermann

Table (1.A.1) Estimated Models

Name	Model description	References
Model 8	<p><i>Neural net model 2</i></p> $r_{t+1} = \theta_0 + \sum_{i=1}^{n_1} \theta_i g(\sum_{j=1}^{n_2} \beta_j g(\alpha'_j \boldsymbol{\eta}_t)) + u_{t+1}$ <p>With g the logistic function, $\boldsymbol{\eta}_t = (1, r_t, r_{t-1}, r_{t-2})'$, $n_1 = 2$ and $n_2 = 1$</p>	Timmermann
Model 9 to Model 18	<p><i>Univariate regressions</i></p> $r_{t+1} = \theta_0 + \theta_1 x_t + u_{t+1}$ <p>With x_t (univariate) exogenous regressors from the list detailed in Table 1.A.2</p>	Welch and Goyal (2008)
Model 19	<p><i>"Kitchen sink" regression</i></p> $r_{t+1} = \theta_0 + \boldsymbol{\theta}'_1 \mathbf{X}_t + u_{t+1}$ <p>With \mathbf{X}_t the exogenous regressors from the list detailed in Table 1.A.2</p>	Welch and Goyal (2008)
Model 20	<p><i>"Model selection" from Goyal and Welch (2008)</i></p> <p>With all the potential combinations $\mathbf{X}_{i,t}$ from the list detailed in Table 1.A.2, we evaluate:</p> $r_{t+1} = \theta_{i,0} + \boldsymbol{\theta}'_{i,1} \mathbf{X}_{i,t} + u_{i,t+1}$ <p>At each point in time, we choose the model with the smallest out-of-sample R^2</p>	Welch and Goyal (2008)
Model 21	<p><i>Factor model from Kelly and Pruitt (2013)</i></p> <p>Only for aggregate return predictions</p> <p>With $bm_{i,t}$ the book-to-market ratio of portfolio i and F_t the estimated factor, we run the following three regressions:</p> $bm_{i,t} = \theta_{i,0} + \theta_{i,1} r_{t+1} + e_{i,t} \text{ (time series)}$ $bm_{i,t} = c_t + F_t \hat{\theta}_{i,1} + u_{i,t} \text{ (cross section)}$ $r_{t+1} = \gamma_1 + \gamma_2 \hat{F}_t + \epsilon_{i,t+1} \text{ (time series)}$	Kelly and Pruitt (2013)
Model 22	<p><i>Forecast averaging - equally weighted</i></p> <p>Let $p_{j,t+1}$ the forecasts from the J precedent models, we use a simple equally-weighted forecast averaging of the form:</p> $p_{t+1} = \sum_{j=1}^J p_{j,t+1}$	Timmermann
Model 23	<p><i>Model selection - in-sample</i></p> <p>From the J precedent models (apart from Model 22), we evaluate the in-sample RMSE for each single model and take as a prediction the forecast of the model with the lowest RMSE.</p>	Timmermann

1.A.2 Datasets

Table (1.A.2) External regressors used in Model 9 to 20 in Table 1.A.1

Variable	Description	Sources
tms_t	<i>Term spread</i> 10 Year Treasury rate minus the 3-Month T-Bill rate	Amit Goyal website (before January 2020), FRED (after January 2020)
$cape_{1,t}$	<i>Cyclically-adjusted PE (CAPE) ratio 1</i> Real S&P 500 Prices divided by the 10-year moving average of the corresponding real Earnings	Robert Shiller website
$cape_{2,t}$	<i>Cyclically-adjusted PE (CAPE) ratio 2</i> CAPE ratio with scaled Earnings (i.e. adjusted to account for changes in corporate payout policy)	Robert Shiller website
pe_t	<i>PE ratio</i> Nominal S&P 500 prices divided by corresponding nominal Earnings	Robert Shiller website
bm_t	<i>Book-to-Market ratio</i> Median Book-to-Market ratio of Fama-French 100 portfolios	Kenneth French website
ecy_t	<i>Excess CAPE yield</i> Inverse of $cape_{1,t}$ minus the 10-year real sovereign rate	Robert Shiller website
dp_t	<i>Dividend-Price ratio</i> Log of S&P 500 nominal dividends minus log of S&P 500 contemporaneous nominal dividends (as in Goyal and Welch (2008))	Robert Shiller website
dy_t	<i>Dividend Yield</i> Log of S&P 500 nominal dividends minus log of S&P 500 previous nominal dividends (as in Goyal and Welch (2008))	Robert Shiller website
$vol_{1,t}$	<i>Return Volatility 1</i> Monthly average of daily squared aggregate returns, as in Goyal and Welch (2008)	Kenneth French website
$vol_{2,t}$	<i>Return Volatility 2</i> Monthly average of daily aggregate return volatility estimated with a GARCH(1,1)	Kenneth French website
$index_t$	<i>Index level</i> S&P 500 index level	Robert Shiller website
IP_t	<i>Industrial Production</i> US Industrial Production	FRED

Above listed variables are available over the all estimation period (September 1945-October 2020).

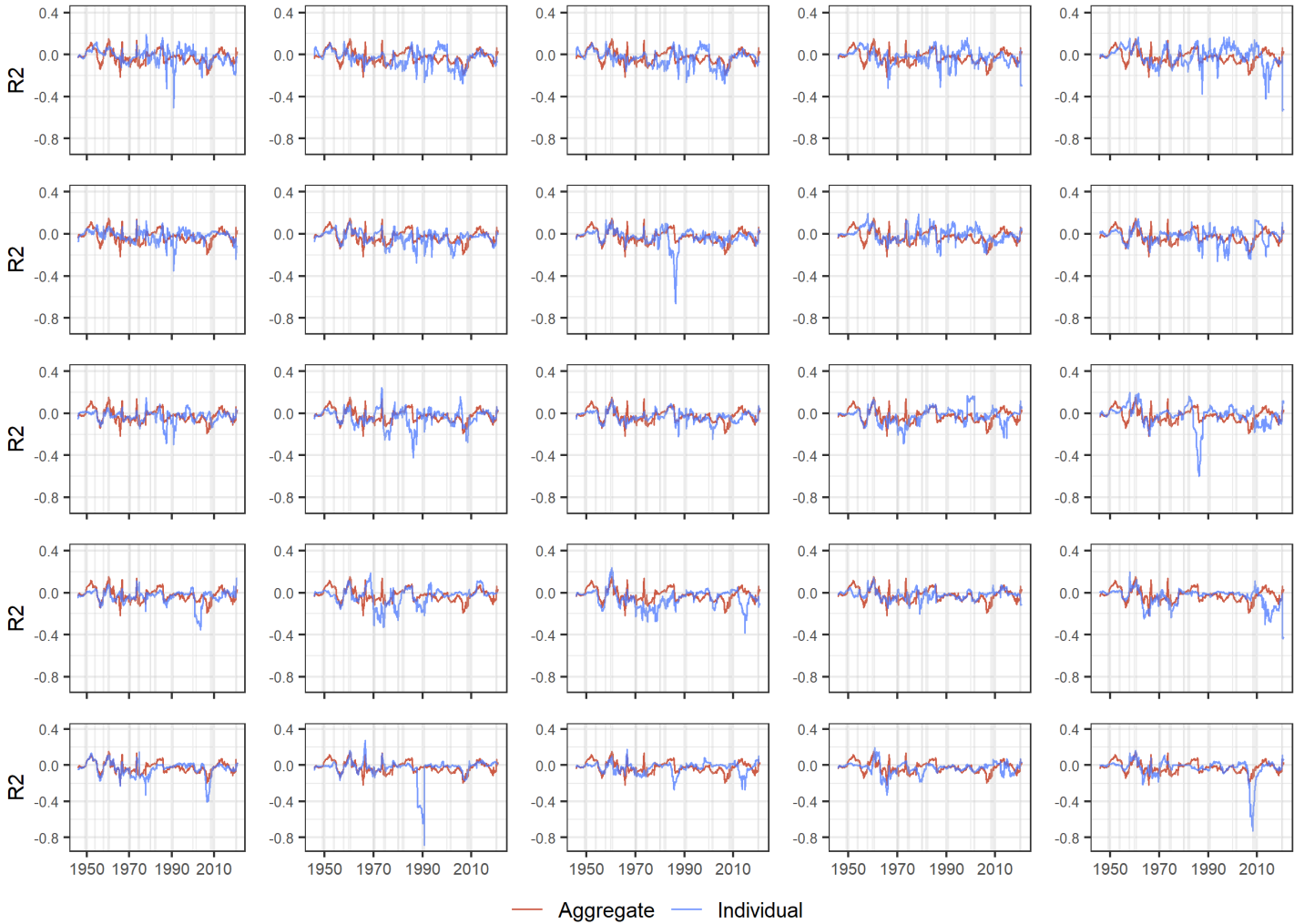
Table (1.A.3) Additional external regressors used in Sections 1.5.2 and 1.6

Variable	Description	Sources
$Michigan_t$	<i>Consumer Sentiment</i> Consumer Sentiment Index from the University of Michigan	FRED
$unemp_t$	<i>Unemployment rate</i> US Unemployment rate	FRED
Baa_t	<i>Baa-Aaa spread</i> Moody's Seasoned Baa Corporate Bond Yield minus Aaa Corporate Bond Yield	FRED
LF_t	<i>U.S. broker-dealer leverage</i> Seasonally adjusted changes in U.S. broker-dealer leverage (Adrian, Etula and Muir (2014))	Tyler Muir website
$Yale_t$	<i>Confidence Index of Yale University</i> Seasonally adjusted changes in U.S. broker-dealer leverage (Adrian, Etula and Muir (2014))	Yale University website

Various variables also used in the regressions of Section 1.5.2 and 1.6 are already detailed in Table 1.A.2: ecy_t , pe_t , $vol_{1,t}$ and $vol_{2,t}$.

1.A.3 Raw Predictability series: individual graphs

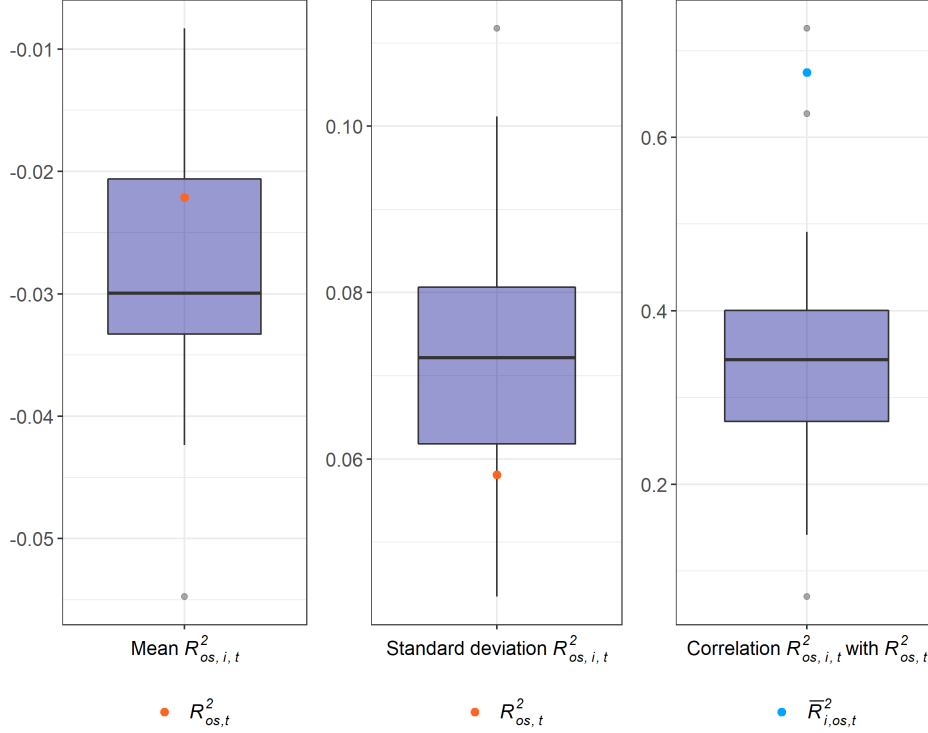
Figure (1.A.1) Individual Micro- and Macro-Raw Predictability series, over time



On the different graphs are represented the macro- ($R_{os,t}^2$, in red) and micro- ($R_{i,os,t}^2$, in blue) raw predictability indices according to the methodology outlined in Section 1.4.2. The metric used is the out-of-sample R^2 , also detailed in Section 1.4.2, that can take negative values. The grey vertical bands figure the NBER US recession dates.

1.A.4 Moments of the raw return predictability series

Figure (1.A.2) Distribution of Micro-Raw Predictability Statistics



On the different graphs are represented the distributions (in blue) of different statistics of the I series $R^2_{i,os,t}$: their means, their standard deviations and their correlations with respect to $R^2_{os,t}$. The grey points represent the outliers of the aforementioned distributions. The coloured point represent the corresponding statistics either for $R^2_{os,t}$ or for $\bar{R}^2_{i,os,t}$. We thus notice, along Section 1.5.1: first that the mean of $R^2_{os,t}$ is in line with the means of the different $R^2_{i,os,t}$, second that the standard deviation of $R^2_{os,t}$ stands below the first quartile of $R^2_{i,os,t}$, eventually that pooling the different series $R^2_{i,os,t}$ into $\bar{R}^2_{i,os,t}$ sharply increases the correlation with $R^2_{os,t}$. The metric used is the out-of-sample R^2 , also detailed in Section 1.4.2, that can take negative values.

1.A.5 Standard errors, mean and standard deviations of raw predictability series

To assess the difference in means and standard deviations of $R_{os,t}^2$ with respect to $R_{i,os,t}^2$, we fit an ARMA(1,1) on each series. More precisely, with Y_t being either $R_{os,t}^2$ or $R_{i,os,t}^2$, we estimate:

$$Y_t = c + \gamma Y_{t-1} + \theta \epsilon_{t-1} + \epsilon_t \text{ and } E(\epsilon_t^2) = \sigma_\epsilon^2 \quad (1.13)$$

For each series, we then compute their estimated unconditional means m as:

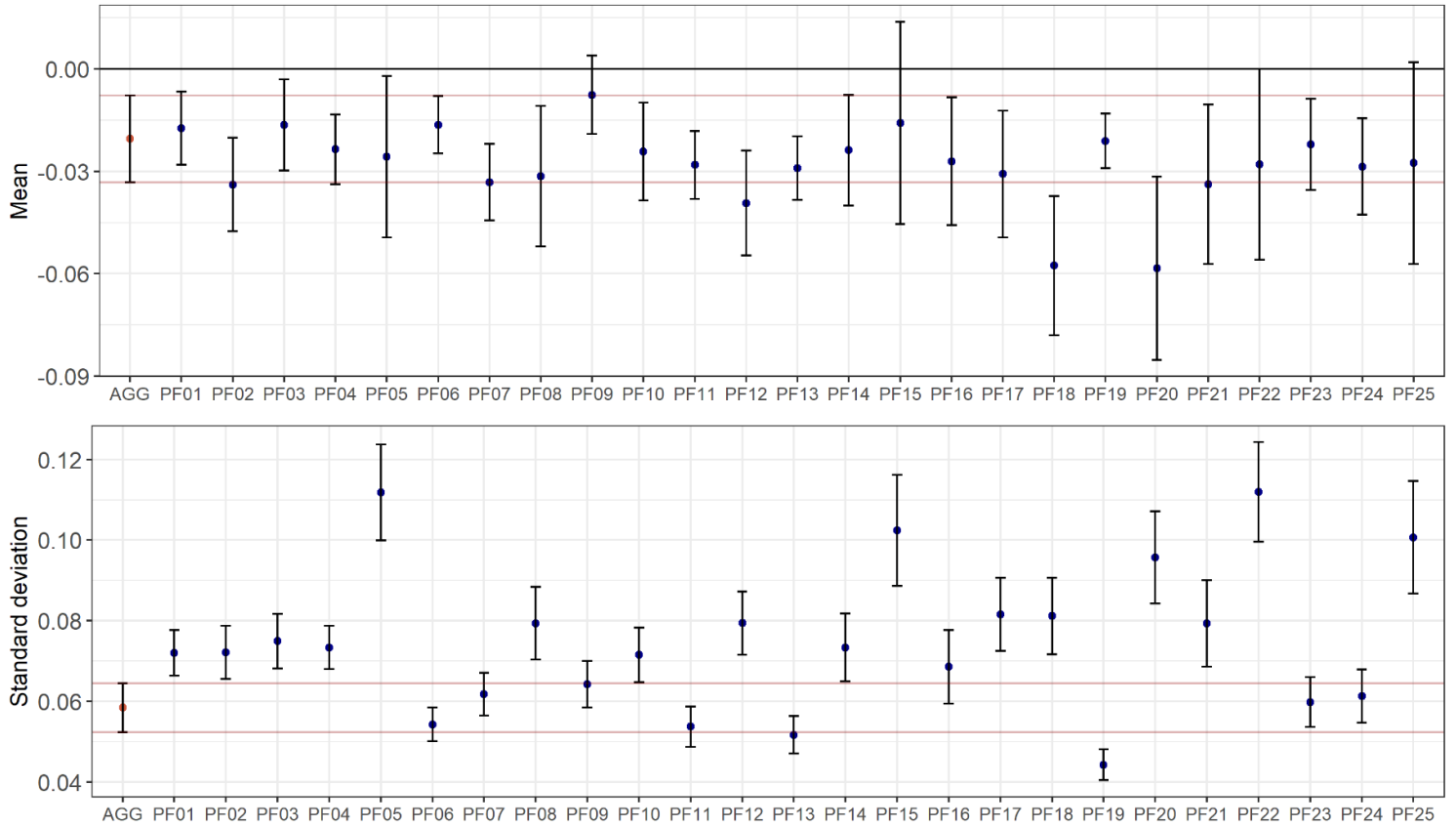
$$\hat{m} = \frac{\hat{c}}{1 - \hat{\gamma}} \quad (1.14)$$

And their variances σ^2 as:

$$\hat{\sigma}^2 = \frac{(1 + 2\hat{\gamma}\hat{\theta} + \hat{\gamma}^2)\hat{\sigma}_\epsilon^2}{1 - \hat{\gamma}^2} \quad (1.15)$$

Standard errors for these two estimates are obtained with 500 bootstrap simulations. Mean and standard deviation for $R_{os,t}^2$ are depicted in red in Figure 1.A.3, and in blue for the 25 $R_{i,os,t}^2$. Black error bands figure +/- 1 standard error confidence intervals along the estimates. We thus notice on Figure 1.A.3, in line with Section 1.5.1, that, although the means of $R_{os,t}^2$ and $R_{i,os,t}^2$ appear indistinguishable from each other, the standard deviation of $R_{os,t}^2$ is significantly lower than for $R_{i,os,t}^2$.

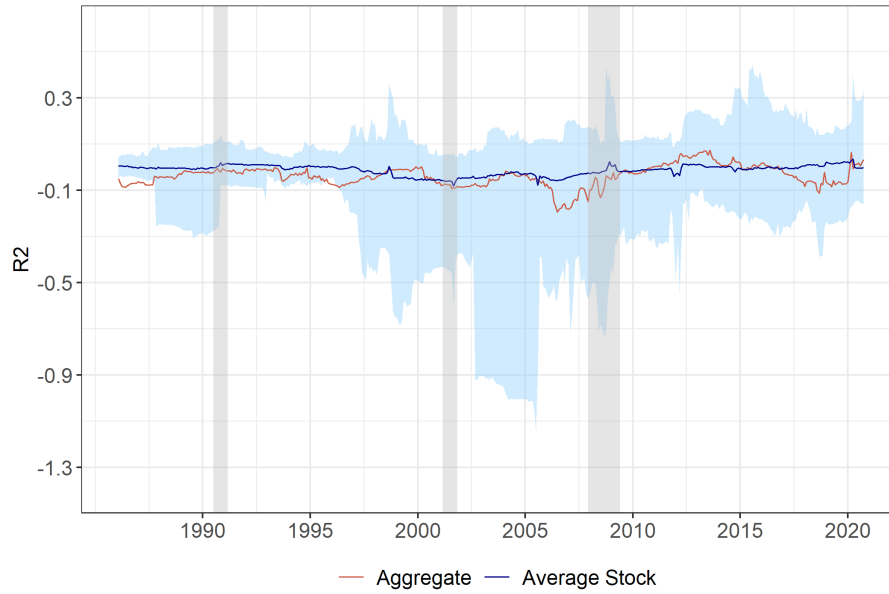
Figure (1.A.3) Mean and standard deviations of raw predictability series: confidence intervals



On the graph are represented the unconditional means (upper panel) and the standard deviations (lower panel) of $R_{os,t}^2$ (in red) $R_{i,os,t}^2$ (in blue). The coefficients are obtained by fitting the series with ARMA(1,1) processes, as described in Equations 1.14 and 1.15. The standard errors are obtained with 500 bootstrap simulations. Black error bands figure +/- 1 standard error confidence intervals.

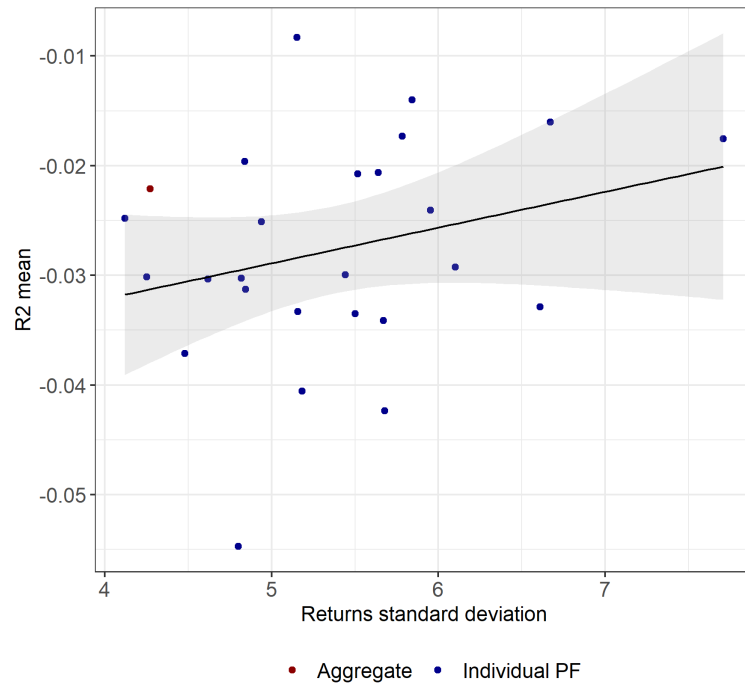
1.A.6 Individual stock return predictability

Figure (1.A.4) Macro-Raw Predictability series and averaged Individual Raw Predictability



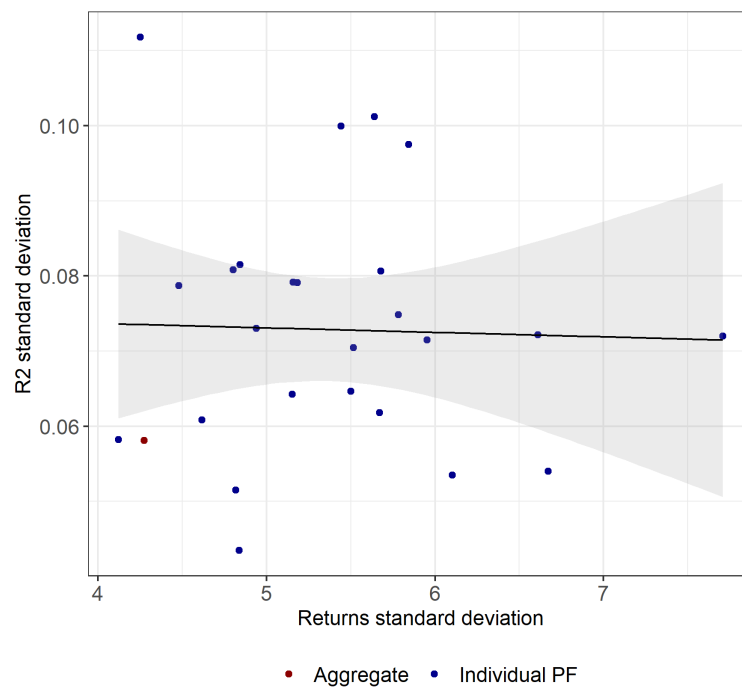
On the graph are represented the macro-raw predictability ($R_{os,t}^2$, in red) and the average micro-raw predictability across individual stocks ($\overline{R}_{i,os,t}^2$, in blue). The details of the methodology are outlined in Section 1.4.2. The light blue area represents the gap between the minimum and the maximum values taken by the different individual-raw predictability series ($R_{i,os,t}^2$). The metric used is the out-of-sample R^2 , also detailed in Section 1.4.2, that can take negative values. The grey vertical bands figure the NBER US recession dates. The average standard deviation of $R_{i,os,t}^2$ is 0.06 whereas the standard deviation of $R_{os,t}^2$ is 0.04. The correlation of $\overline{R}_{i,os,t}^2$ with $R_{os,t}^2$ is 0.3 whereas the average correlation of $R_{i,os,t}^2$ with $R_{os,t}^2$ is 0.09.

Figure (1.A.5) Raw Predictability levels vs. Returns standard deviations



On the scatter plot are represented, for $r_{i,t+1}$ (in blue) and r_{t+1} (in red), the standard deviations of the returns series on the x-axis, and the mean of their raw predictability, $R_{i,os,t}^2$ or $R_{os,t}^2$, on the y-axis. The metric used for the y-axis is the out-of-sample R^2 , detailed in Section 1.4.2, that can take negative values.

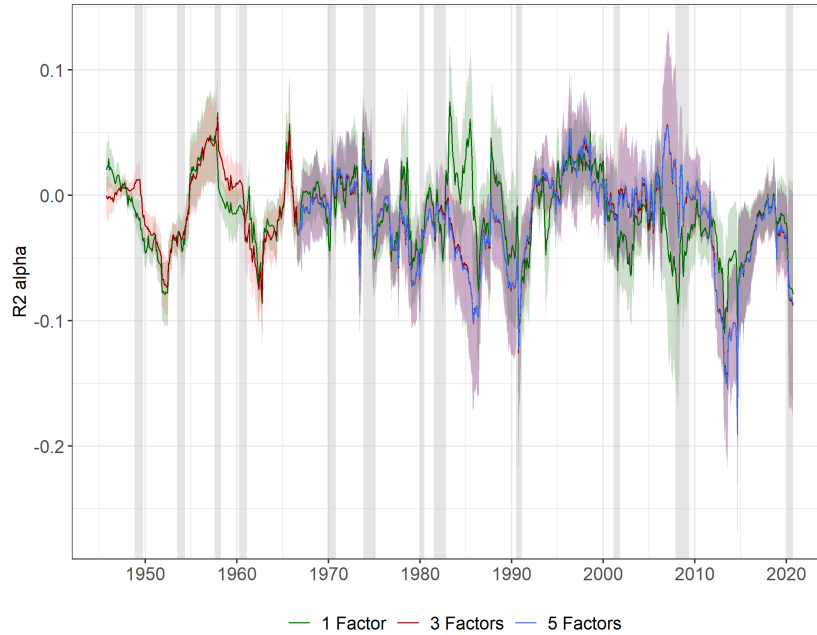
Figure (1.A.6) Raw Predictability standard deviations vs. Returns standard deviations



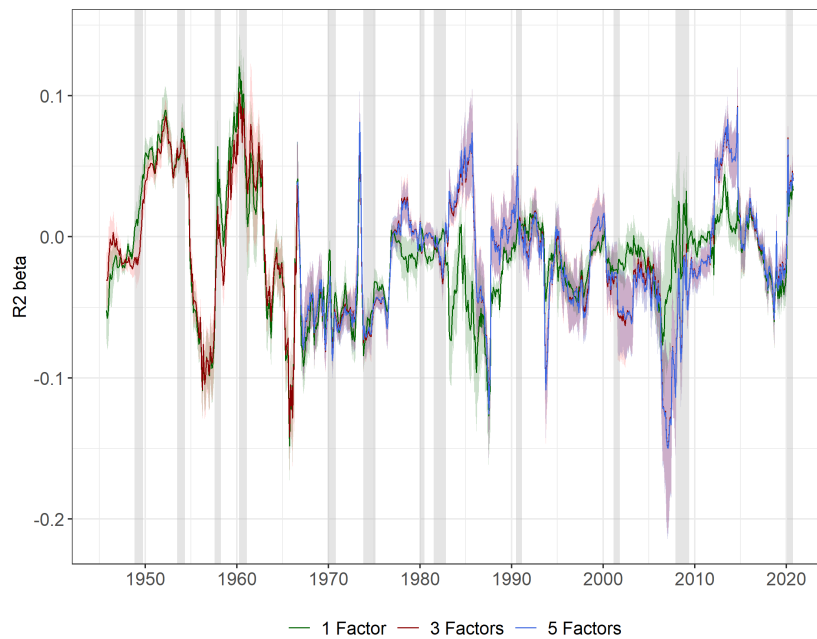
On the scatter plot are represented, for $r_{i,t+1}$ (in blue) and r_{t+1} (in red), the standard deviations of the returns series on the x-axis, and the standard deviations of their raw predictability, $R_{i,os,t}^2$ or $R_{os,t}^2$, on the y-axis. The metric used for the y-axis is the out-of-sample R^2 , detailed in Section 1.4.2, that can take negative values.

1.A.7 Robustness checks: alternative risk factors

Figure (1.A.7) $\overline{R^2}_{i,\alpha,t}$ with different Factor Specifications



On the graph are represented the average across portfolios of the alpha-predictability series ($\overline{R^2}_{i,\alpha,t}$) computed using the 1-factor (in green), the 3-factor (in red) or the 5-factor (in blue) Fama-French models. The coloured areas figure the corresponding cross-sectional dispersion around the different $\overline{R^2}_{i,\alpha,t}$ (± 0.5 standard deviation). The metric used is the out-of-sample alpha-predictability $R^2_{i,\alpha,t}$, detailed in Section 1.4.3, that can take negative values. The grey vertical bands figure the NBER US recession dates.

Figure (1.A.8) $\overline{R^2}_{i,\beta,t}$ with different Factor Specifications

On the graph are represented the average across portfolios of the beta-predictability series ($\overline{R^2}_{i,\alpha,t}$) computed using the 1-factor (in green), the 3-factor (in red) or the 5-factor (in blue) Fama-French models. The coloured areas figure the corresponding cross-sectional dispersion around the different $\overline{R^2}_{i,\beta,t}$ (± 0.5 standard deviation). The metric used is the out-of-sample beta-predictability $R^2_{i,\beta,t}$ detailed in Section 1.4.3, that can take negative values. The grey vertical bands figure the NBER US recession dates.

1.A.8 Robustness checks: regression results

Table (1.A.4) Additional Regression Results for the Alpha- and Beta-Predictability

	<i>Dependent variable:</i>			
	Alpha-pred.: $\overline{R^2}_{i,\alpha,t}$		Beta-pred.: $\overline{R^2}_{i,\beta,t}$	
	(1)	(2)	(1)	(2)
pe_t	0.001** (0.0003)	0.001*** (0.0002)	-0.001*** (0.0002)	-0.001*** (0.0003)
$Michigan_t$	0.0004 (0.0003)		-0.001*** (0.0002)	
$-unemp_t$		0.008*** (0.002)		-0.008*** (0.002)
vol_t	-0.0004 (0.001)		0.0004 (0.001)	
$vol_{2,t}$		-0.002 (0.002)		0.002 (0.001)
Const.	-0.074*** (0.026)	0.017* (0.010)	0.063*** (0.022)	-0.041*** (0.016)
Obs.	496	856	496	856
R ²	0.077	0.168	0.086	0.107
Adj. R ²	0.071	0.165	0.086	0.107

Note: *p<0.1; **p<0.05; ***p<0.01

On the table are represented the different regression results with $\overline{R^2}_{i,\alpha,t}$ and $\overline{R^2}_{i,\beta,t}$ as a predicted variables. t -statistics have been computed using Newey-West standard errors. Variables are rearranged so that an increase in $X_{IE,t}$, $X_{FC,t}$ and $X_{RA,t}$ reflects, respectively, a surge in market effervescence, an aggravation of financial constraints and a strengthening of economic activity

2. Forecasting Real Activity using Cross-Sectoral Stock Market Information

(with Nicolas Chatelais¹ and Menzie Chinn²)

Abstract

Stock prices declined abruptly in the wake of the Covid-19, reflecting both the deterioration of investors' expectations of economic activity as well as the surge in risk aversion. In the following months however, economic activity remained sluggish while equity markets bounced back. This disconnect between equity values and macro-variables can be partially explained by other factors, namely the decline in risk-free interest rates, and -for the US- the strong profitability of the IT sector. As a result, an econometrician forecasting economic activity with aggregate stock market variables during the Covid-crisis is likely to get poor results. Our main contribution is thus to rely on sectorally disaggregated equity variables within a factor model in order to predict US economic activity. We find, first, that the factor model better predicts future economic activity compared to aggregate equity variables, or to conventional benchmarks used in the literature, both in-sample and out-of-sample. Second, we show that the strong performance of the factor model comes from the fact that it filters out the "expected returns" component of the sectoral equity variables as well as the foreign component of aggregate future cash flows. The constructed factor overweights upstream and "value" sectors that are found to be closely linked to the future state of the business cycle.

Lors du choc Covid de 2020, les prix d'actions chutèrent brutalement en lien avec la détérioration des perspectives d'activité économique et avec la hausse de l'aversion au risque. Dans les mois qui suivirent toutefois, alors que le niveau d'activité restait morose, les marchés action rebondirent rapidement. Cette déconnexion entre variables boursières et variables macros peut en partie être expliquée par d'autres facteurs, notamment par le déclin des taux d'intérêt et, pour les États-Unis, par la forte rentabilité du secteur du numérique. Par conséquent, un économètre essayant de prédire l'activité économique pendant la crise du Covid à l'aide de données action agrégées obtiendrait des résultats médiocres. La principale idée de ce chapitre est ainsi d'utiliser des données action sectorielles, dans le cadre d'un modèle à facteurs, pour prédire la future production industrielle. Côté résultats, nous trouvons, en premier lieu, que notre modèle prédit mieux l'activité économique par rapport aux variables action agrégées mais également par rapport aux variables de référence utilisées en prédiction macroéconomique. Deuxièmement, nous montrons que la bonne performance de notre modèle vient du fait qu'il filtre, parmi les variables action sectorielles, la composante liée au taux d'actualisation ainsi que la composante liée aux dividendes tirés de l'activité étrangère des firmes. Nous soulignons également que notre modèle surpondère les secteurs en amont des chaînes de valeur ainsi que les secteurs *value* qui apparaissent intimement liés au futur cycle des affaires.

Keywords: *factor model, forecasting, equity markets.*

JEL: C22, C53, G12, G14, G17

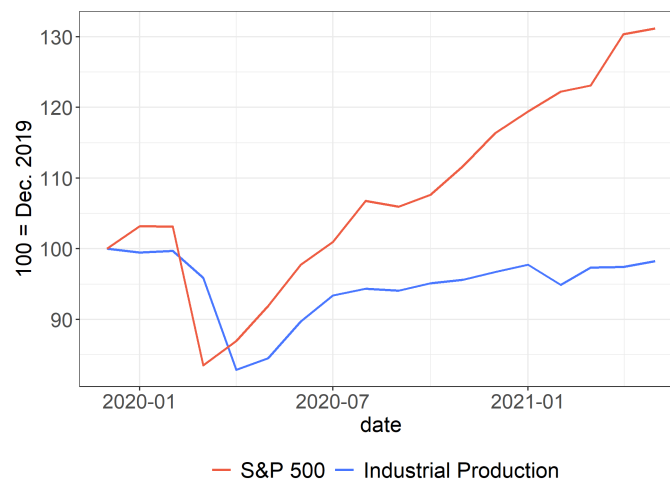
¹Banque de France, email: nicolas.chatelais@banque-france.fr

²University of Wisconsin-Madison and NBER, email: mchinn@lafollette.wisc.edu

2.1. Introduction

Forecasting macroeconomic variables using financial indicators has proven a challenging task, a surprising outcome given the fact that financial variables like bond yields and stock prices should impound expectations of future economic activity. The recent divergence between developments in equity markets and the economic activity has only highlighted the apparent disconnect between finance and the real economy. After the Covid shock in March 2020, stock prices declined abruptly, reflecting both the deterioration of investors expectations of future economic activity as well as the surge in aggregate risk aversion. In the following months however, and to the surprise of many, whereas economic activity remained relatively sluggish, equity markets bounced back sharply, as illustrated in Figure 2.1.1.

Figure (2.1.1) S&P 500 and US Industrial Production (100 = Dec.2019)



Note: The graph represents the evolution of the US Industrial Production and of the S&P 500 Index. Both indices are set to 100 in December 2019. Sources: Federal Reserve Economic Data, Refinitiv Datastream.

A simple, but incomplete, explanation is that not only do stock prices reflect expected future cash flows and investors risk aversion, but also the level of risk free interest rates. Focusing on the American example, US 10 year sovereign rates declined from March to August 2020 and can therefore explain part of the equity rebound ([Chatelais and Stalla-Bourdillon, 2020](#)).

We think that this seeming disconnect between finance and the real economy can be more fully reconciled with the data by recognizing that relying on a given aggregate stock price index discards a lot of information that might be of particular importance, especially during business cycle turning points. Specifically, US aggregate stock indices can be influenced by other forces that do not entirely reflect the state of the US economy. For example the S&P 500 was driven up in 2020 by IT sector companies whose valuations either largely depend on foreign activity or are orthogonal to US economic performance. As their profitability derived tremendously from Covid19 lockdown policies. As a result, an econometrician trying to forecast economic activity with aggregate stock variables during the Covid-crisis is likely to get poor results.

In this paper, we address this problem by building a factor model constructed using sectorally disaggregated equity variables to predict future US economic activity. Hence, this study constitutes one of the rare instances where stock market variables specifically are used to perform macroeconomic forecasting. Furthermore, this study adds to a surprisingly small literature relying on sectoral equity variables. To our knowledge, this paper is the first to use factor models to extract the predictive content from disaggregated sectoral stock prices. Even papers employing factor models based on large sets of variables seldom go beyond using aggregate stock indices ([Barhoumi et al., 2010](#), [Jardet and Meunier, 2022](#)).

We obtain three main results, relating to forecasting performance, and the sectoral sources of forecasting business cycle activity.

First, we find that a factor based on sectoral dividend yields (DYs) better predicts industrial production (IP) growth, as compared to the same variable measured as an aggregate. That factor model also typically outperforms conventional benchmark models, such as the term spread or the lagged IP growth, particularly during times of negative IP growth. In our baseline specification, we forecast future IP growth over a 12-month horizon, but these results hold at the 18-month and the 24-month horizons, both in-sample and out-of-sample. We also find that our factor model helps to improve the forecasting accuracy of a widely used factor model à la [Stock et al. \(2002\)](#)

that relies on a vast number of macro-financial variables (but not on sectoral equity indices). Interestingly, our finding generalizes to a number of other countries³.

Second, relying upon the present value formula of [Campbell and Shiller \(1988\)](#), we find that our model improves forecasting accuracy because it filters out the expected returns/discount rate component of the sectoral equity variables, as well as the foreign component of aggregate future cash flows. We attribute the elevated outperformance of our factor model, especially during periods of negative IP growth such as during the Covid pandemic or during the Global Financial Crisis, to this filtering out of extraneous information. As expected returns are more volatile in recessionary states ([Henkel et al., 2011](#)) they tend to particularly affect the forecasting accuracy of the aggregate DY during these periods, but not of our factor model.

Third, we are able to identify the specific sectors that provide additional forecasting power. Specifically, we find that our factor model overweights upstream sectors (primary industry and other industrial inputs) and “value” sectors, as the latter are found to be closely linked to the US business cycle ([Zhang, 2005](#), [Kojen et al., 2017](#), [Xu, 2018](#)).

For those who are particularly concerned about the trajectory of economic activity during economic downturn, our forecasting model should be of special interest, given the economically and statistically significant outperformance relative to conventional benchmarks. The identification of key sectoral indicators also provides an appealing economic intuition for our findings.

In the following section, we present the basic theory placed in the context of the literature. In [Section 2.3](#) we present the empirical model and detail the data used in the analysis. [Section 2.4](#) provides a set of in-sample results, and [Section 2.5](#) a corresponding set of out-of-sample results. We draw out the economic implications of those results in [Section 2.6](#). Concluding remarks are contained in [Section 2.7](#).

³The outperformance also extends to specifications including some measure of volatility, such as the VIX. This point, as well as the results regarding other countries industrial production growth, are discussed in the [Section 2.5](#)

2.2. Background

2.2.1 Theoretical Framework

When using aggregate financial measures to predict economic activity, one wants the factors influencing the financial variables to correspond to the appropriate macroeconomic variable. Since our objective is to forecast US economic activity, we want our financial predictor to reflect solely US activity. In order to extract the US component, we rely upon the present value formula of [Campbell and Shiller \(1988\)](#), a decomposition that has been widely used to model equity returns (see [Campbell and Ammer, 1993](#), [Vuolteenaho, 2002](#) and [Binsbergen and Koijen, 2010](#)).

More precisely, DYs (x_t) can be decomposed into two factors: expected returns (or discount rates) and expected cash flows growth likewise:

$$x_t = \frac{\kappa}{(1 - \rho)} + \sum_{j=1} \rho^{j-1} E_t[r_{t+j} - \Delta cf_{t+j}] \quad (2.1)$$

Where $E_t[r_{t+j}]$ represents expected returns and $E_t[\Delta cf_{t+j}]$ expected cash flows (κ and ρ are constant parameters). One could also decompose the cash flow component into two sub-components: one depending on the domestic activity of the firm, $E_t[\Delta cf_{D,t+j}]$, and the other one stemming from its foreign activity, $E_t[\Delta cf_{F,t+j}]$, such that we would get:

$$x_t = \frac{\kappa}{(1 - \rho)} + \sum_{j=1} \rho^{j-1} E_t[r_{t+j} - \Delta cf_{D,t+j} - \Delta cf_{F,t+j}] \quad (2.2)$$

Note eventually that a similar decomposition can be applied to other equity variables, such as price-earnings or book-to-market ratios.

In order to forecast future aggregate returns, [Kelly and Pruitt \(2013\)](#) underline that the usual predictive regressions of aggregate future returns and aggregate dividend growth on aggregate

DY:

$$r_{t+h} = \alpha_1 + \beta_1 x_t + u_{1,t+h} \quad (2.3)$$

$$\Delta c f_{t+h} = \alpha_2 + \beta_2 x_t + u_{2,t+h} \quad (2.4)$$

are misspecified, since the DY both reflects expected returns and expected cash flows, while they would like this variable only to reflect the former (when predicting aggregate returns), or the latter (when predicting aggregate dividend growth).

Relying on disaggregated book-to-market ratios, which can also be decomposed with the [Campbell and Shiller \(1988\)](#) formula, [Kelly and Pruitt \(2013\)](#) estimate a factor model via Partial Least Squares on that appears to predict accurately future aggregate returns and future aggregate dividends. They explain the improved accuracy by the fact that the factor model, by overweighting or underweighting certain sectoral book-to-markets, filters out the expected cash flow component while predicting future aggregate returns (and vice versa when predicting future aggregate dividends).

In an approach similar to theirs, we implement the same filtering to extract a factor to predict future economic activity. In our case we want the factor model to not only filter out the expected returns component, but also the foreign cash flow component. Implicitly, we assume that the domestic cash flow component represents a good proxy for domestic US economic activity. We also assume that this filtering is possible because sectoral DYs are informative about future aggregate cash flows. We return to this point more formally in Section [2.A.1](#) of the Appendix.

2.2.2 Selected Literature Review

There are three strands of the literature relevant to our contribution. The first is the literature using stock prices to predict economic activity. The second is the use of factor modeling for

forecasting purposes. The third focuses on how expectations regarding future economic activity affect the cross section of returns.

Turning to the first strand, the theoretical arguments underlining the predictive power of stock prices are twofold (Croux and Reusens, 2013). On one hand equity prices are inherently forward looking and should therefore reflect investors expectations of future economic activity. On the other hand, stock prices can have a causal effect on the business cycle: if stock prices go up, households should consume more through the induced wealth effect. Hence, stock prices should lead aggregate activity. Consequently, various papers try to predict future GDP or industrial production with equity variables, typically with aggregate stock indices (Binswanger, 2000, Ólan Henry et al., 2004, Croux and Reusens, 2013, McMillan, 2021, Chen and Ranciere, 2019, Lan et al., 2020).

Some papers, however, rely on *disaggregated* stock price data and can be further divided into two subcategories. In the first subcategory are papers that first build an aggregate variable from sectoral equity data and then forecast future activity with the former. Loungani et al. (1990) for example use industry-level equity prices to build a metric of price dispersion. They reason that if stock prices are increasing in some industries but declining in others, in subsequent years capital and labor will have to be reallocated from the contracting industries to the expanding ones, which will be costly in the aggregate. Liew and Vassalou (2000) rely on the Fama-French factors, built from disaggregated portfolio returns, to forecast future GDP. Their rationale is that, before a recession, investors should be able to anticipate that small stocks and value stocks will perform badly. Indeed, small-sized firms and value companies, i.e. firms with low price-earnings ratios and typically elevated fixed capital as in the automobile industry, are usually deemed as less resilient to strong negative shocks (Zhang, 2005, Xu, 2018). As a result, small minus big (SMB) returns and high minus low (HML) book-to-market returns should decrease ahead of recessions. In the second subcategory are other papers that directly use the sectoral equity variables in their estimation, most of the time by evaluating the predictive power of specific sector variables in

isolation from the other (Browne and Doran, 2005, Andersson and Agostino, 2008, Zalgiryte et al., 2014).

We depart from the approach adopted in these papers first by estimating a factor model based on sectoral equity variables. We therefore make use of the *entire cross section* of stock market variables at the same time (in contrast to Browne and Doran, 2005, Andersson and Agostino, 2008, Zalgiryte et al., 2014). Moreover, we do not constrain the predictive content of disaggregated stock variables into a specific aggregate predictor, like the dispersion of stock prices or the Fama-French factors. Second, in contrast to all the papers cited above, we also investigate the over- and under-weights of the different sectors in our factor model.

In the end, our approach comes closest to two papers that also rely on the Kelly and Pruitt (2013) factor model to predict macroeconomic activity on the basis of equity variables. However, unlike our approach, they either use aggregate and not sectoral – indices to build their factor, i.e., the number of IPOs or the share turnover in the US (Huang et al., 2015), or they only perform their analysis in-sample and do not analyze what is filtered out in their factor modelling (Jagannathan and Marakani, 2015).

Second, we also contribute to the literature on factor modelling that does not specifically focus on the predictive content of equity variables. Surprisingly enough, whereas disaggregated equity data is easily available and is accessible without lags, to our knowledge the literature on factor models for forecasting exercises rarely relies on sectoral stock data, even when using large datasets (Bessec et al., 2012, Hopenstrick and Marcellino, 2016, Fan et al., 2017, Ferrara and Marsilli, 2019, Jardet and Meunier, 2022) or when using other types of sectoral variables, like surveys (Barhoumi et al., 2010).

Finally, we also contribute to the financial literature that takes perspective inverse of the standard, by evaluating how future economic activity affect cross-sectional stock returns (Kojen et al., 2017, Zhu et al., 2020). By analyzing how the factor model over/underweights certain equity sectors we shed a new light on the pro- and counter-cyclicality of specific portfolios.

2.3. Model Specification and Data

2.3.1 A Factor Model

We follow [Kelly and Pruitt \(2013\)](#), who utilize the Partial Least Square (PLS) methodology estimated using disaggregated equity variables. The approach resembles Principal Components Analysis (PCA), but instead of reducing the dimensionality according to the covariance of the sectoral variables between themselves, we implement the reduction according to the covariance between the predicted variable and the sectoral variables.

Starting with y_{t+h} the predicted variable (in our case, the growth rate of Industrial Production) and x_{it} the different sectoral equity variables (here the sectoral DYs), the PLS is estimated in three steps.

First, for each sector i , a univariate time series regression is estimated:

$$x_{it} = \phi_{i0} + \phi_i y_{t+h} + e_{it} \quad (2.5)$$

Second, for each time period t , the sectoral DYs x_{it} are regressed on the coefficients $\hat{\phi}_i$ estimated above. Note that this regression is a cross-sectional one, and that the estimated coefficient will be the value of the factor F_t at time t :

$$y_{t+h} = \beta_0 + \beta_1 \hat{F}_t + u_{t+h} \quad (2.6)$$

Finally, we use the estimated factor in a (time series) predictive regression:

$$y_{t+h} = \beta_0 + \beta_1 \hat{F}_t + u_{t+h} \quad (2.7)$$

The estimated factor \hat{F}_t can be seen as a weighted sum of the different x_{it} since:

$$\hat{\phi}_i = \frac{\sum_t (x_{it} - \bar{x}_i)(y_{t+h} - \bar{y})}{\sum_t (y_{t+h} - \bar{y})^2} \quad (2.8)$$

With $\bar{x}_i = \frac{1}{T} \sum_t x_{it}$ and $\bar{y} = \frac{1}{T} \sum_t y_{t+h}$. And since:

$$F_t = \frac{\sum_i (x_{it} - \bar{x}_i)(\phi_i - \bar{\phi})}{\sum_i (\phi_i - \bar{\phi})^2} \quad (2.9)$$

With $\bar{\phi} = \frac{1}{I} \sum_i \phi_i$ and $\bar{x}_i = \frac{1}{T} \sum_t x_{it}$. We can therefore write:

$$\hat{F}_t = \frac{1}{C} \sum_i x_{it}(\phi_i - \bar{\phi}) \quad (2.10)$$

With $C = \sum_i (\phi_i - \bar{\phi})^2$. In other words, the more x_{it} is correlated with y_{t+h} the more it will influence \hat{F}_t through the coefficients $(\phi_i - \bar{\phi})$.

2.3.2 Data

Throughout the paper we focus mainly on the United States. In our main specification, we predict future Industrial Production growth. Depending on the forecast horizon h , and with IP_t the Industrial Production index, we forecast at time t the variable:

$$y_{t+h} = \frac{IP_{t+h}}{IP_t} - 1 \quad (2.11)$$

The DYs are drawn from Refinitiv Datastream indices either collected to reflect the overall US equity market or sectoral portfolios. The sectoral indices are based on the Industry Classification Benchmark (IBC), and are available at different granularity: either 11, 20 or 44 sectors. We rely on the most detailed breakdown available (44 sectors), although we retrieve from it 4 sectors

for which the DY series were incomplete: Alternative Energy, Closed end Investments, Precious Metals and Mining and Mortgage Real Estate Investment Trusts. Thus in our main exercise we forecast IP growth with a factor model based on 40 different DY series. In the paper we also consider the aggregate DY, which corresponds to the average DY of the US stock market, also collected by Refinitiv Datastream.

The other macroeconomic and financial data are from sources detailed in Table 2.A.3 of the Appendix. The data is at a monthly frequency, spanning the period from 02-1973 to 05-2021. We define the term spread as the spread between the Treasury 10 year and 3 month yields, in line with Chinn and Kucko (2015).

2.4. In-Sample Results

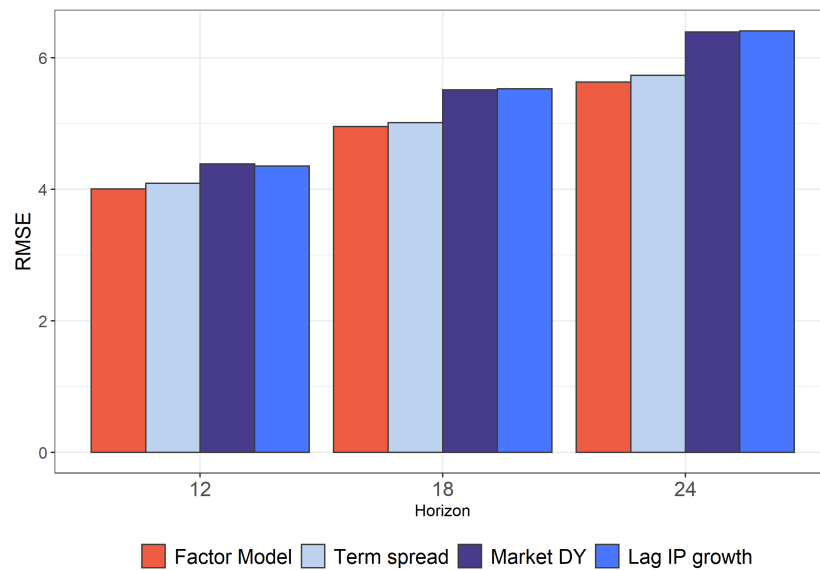
In order to determine whether our disaggregated equity variable based factor model exhibits greater predictive power than models based on aggregate DY, or conventional benchmark models, we conduct both in-sample and out-of-sample analyses. In this section, we present the former set of results, reserving the latter for Section 2.4.

To summarize the prediction results, in Figure 2.4.1 we present the in-sample RMSE of different predictive models at various horizons. In light blue, purple and dark blue bars are represented, respectively, simple forecasting models based either on the term spread, on the aggregate DY or on the lagged IP growth. The in-sample RMSE based on the factor model is shown as the red bar.

Several findings are readily apparent. First, irrespective of the horizon, the factor model constantly beats the conventional benchmarks, that is the lagged IP growth or the term spread, although the term spread appears as the second best performing model.

Second, the factor model outperforms the simple predictive regression based on aggregate equity data (here the aggregate DY), thus highlighting the additional accuracy that can be gained from

Figure (2.4.1) In-Sample RMSE from the different estimated models



Note: On the graph are represented the In-Sample RMSE of different models (the factor model or univariate regressions relying on the aggregate DY, on the lagged IP growth or on the term spread). The predicted variable is the IP growth over 12, 18 and 24 months.

working with sectoral stock market variables. For this last result, it should however be borne in mind that, in an in-sample setting, our factor model should in any case outperform the aggregate DY given that it overweights the sectoral DYs which are the most correlated with future IP growth.

Focusing on the 12-month horizon, we show on Figure 2.A.1 of the Appendix that the same in-sample results hold when we look at alternate proxies of economic activity, although the out-performance with respect to the term spread appears more mixed. We considered manufacturing sales, the number of house permits delivered, the OECD indicator of monthly US GDP, the US unemployment rate or total nonfarm payroll employment.

We perform a second simple in-sample evaluation by determining whether or not the estimated factor brings additional information as compared to our main benchmark (here the aggregate DY, x_t). To do so, we run the following predictive regression:

$$y_{t+h} = \beta_0 + \beta_1 x_t + \beta_2 \hat{F}_t + u_{t+h} \quad (2.12)$$

And evaluate the significance of the coefficient β_2 . Table 2.4.1 below reports the results of these in-sample regressions at horizon 12, 18 and 24 months. To account for the serial correlation of the error terms, we conduct our statistical inference using Newey-West standard errors. Notice in Table 2.4.1 that the coefficient associated with the factors built on sectoral equity variables is significant for all different horizons. This result thus suggests that the factor model has forecasting value even with the inclusion of the aggregate DY in the regression.

Table (2.4.1) Predictive coefficients of the estimated factor (In-sample estimates)

	<i>Dependent variable:</i>		
	12 months (1)	IP growth 18 months (2)	24 months (3)
Market DY	-0.014* (0.008)	-0.015 (0.012)	-0.014 (0.014)
Factor	0.282** (0.125)	0.297*** (0.112)	0.313*** (0.116)
Constant	0.038* (0.022)	0.039 (0.035)	0.035 (0.046)
Observations	532	526	520
R ²	0.265	0.270	0.285
Adjusted R ²	0.262	0.267	0.282
Residual Std. Error	0.038 (df = 529)	0.047 (df = 523)	0.054 (df = 517)
F Statistic	95.295*** (df = 2; 529)	96.723*** (df = 2; 523)	103.059*** (df = 2; 517)

Note: The reported regressions are made using Newey-West heteroskedasticity and serial correlation robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

2.5. Out-of-Sample Results

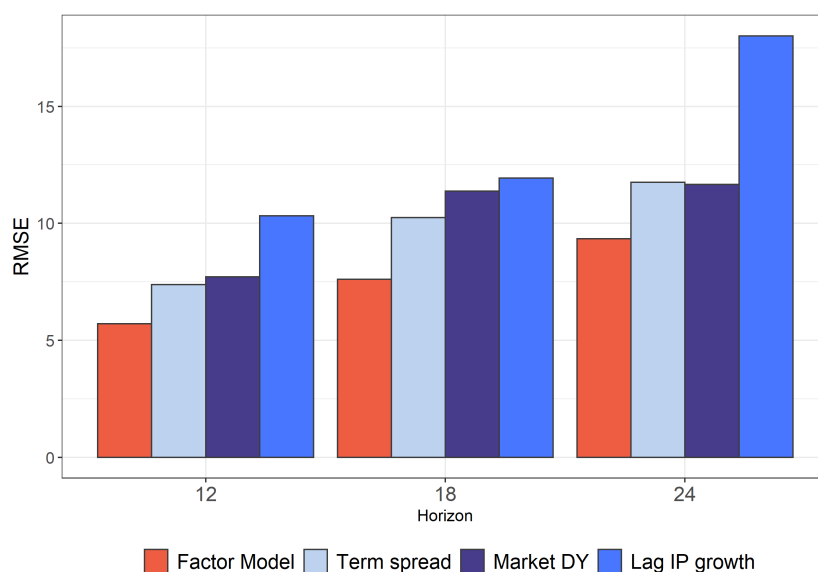
2.5.1 Out-of-Sample Performance

We conduct an out-of-sample forecasting exercise in order to guard against overfitting. Following the same procedure outlined in Section 2.4, we set the rolling window used for estimation to 36 months (3 years). This means that for a 12-month horizon, the first observation to be predicted is January 1977. Our results are robust to consideration of shorter or longer rolling windows. Note that for the out-of-sample exercise, we closely follow the procedure described in Kelly and Pruitt (2013), so that, when predicting IP growth at time $t + h$ based with variables at time t , all the regressions outlined in Section 2.5 are based on training samples that exclude observations posterior to time t .

Figure 2.5.1 indicates, in a format similar to that in Figure 2.4.1, the out-of-sample RMSE estimated for the different models. In line with the in-sample analysis, relying on disaggregated – rather than on aggregate – equity variables strongly improves the forecasting accuracy of our model. Again, this improvement is noticeable through all the different considered horizons. Regarding the relative performance of the other benchmarks, here also the factor model appears to outperform the term spread or the lagged IP growth. Finally, we run the same robustness check as in the in-sample exercise and assess the predictive accuracy of the different models for the other proxies of economic activity. As shown in Figure 2.A.2 in the Appendix, the factor model strongly improves our forecasting accuracy for virtually all the different predicted variables, sometimes decreasing the out-of-Sample RMSE by close to 20%, relative to the best performing benchmark.

We further assess the outperformance of the factor model with respect to the different benchmarks by conducting Diebold-Mariano tests for statistical significance (Diebold and Mariano, 2002, West, 1996). Table 2.5.1 reports the difference in RMSE between the factor model and the different benchmarks, along with the Diebold-Mariano p-values under the null hypothesis that

Figure (2.5.1) Out-of-Sample RMSE from the different estimated models



Note: On the graph are represented the Out-of-Sample RMSE of different models (the factor model or univariate regressions relying either on the aggregate DY, on the lagged IP growth or on the term spread). The predicted variable is the IP growth over 12, 18 and 24 months.

the factor model performs worse than the corresponding benchmarks.

Overall, in line with Figure 2.5.1 and at the notable exception of the term spread at the 12-month horizon, we find that the factor model improves significantly the prediction of future IP growth compared to the three different benchmarks, and at the three different horizons⁴.

We eventually run two out-of-sample exercises to underline the performance of our factor model. First, we evaluate the accuracy of our model compared to forecasting regressions using different metrics of market volatility. Either we rely only on the volatility variables alone in univariate regressions, or we augment the models with the term spread given that recent papers underlined that market volatility may prove useful to extract the forecasting signal out of the term spread (Kumar et al., 2022, Venditti and Veronese, 2020). Table 2.A.1 in the Appendix reports the

⁴The performances of our factor model appear more mixed at shorter horizons. Compared to an univariate model based on the aggregate DY, our factor model does not improve the forecasting accuracy at the 1-month horizon, but exhibits a lower RMSE at the 3- and 6-month horizons, although the difference in RMSE is not significant in lights of Diebold-Mariano tests.

Table (2.5.1) Difference in RMSE with the main benchmark models (Factor model Corresponding benchmark, Out-of-sample estimates)

<i>Benchmark:</i>	<i>Horizon</i>		
	12 months	18 months	24 months
Market DY	-2.01*	-3.76*	-2.33***
Term spread	-1.68	-2.63*	-2.41***
Lagged IP growth	-4.62**	-4.32***	8.67***

*Note: The table reports the difference in RMSE of the factor model compared to the different benchmarks (a negative value means that the factor model outperforms the corresponding benchmark in terms of RMSE). Stars represent the Diebold-Mariano test p-values under the null hypothesis that the factor model performs worse than the benchmark models indicated in the first column. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$*

differences in RMSE between these benchmarks and our factor model. As could be seen on the Table, it appears that our model significantly outperforms the aforementioned benchmarks, at various horizons and for different proxies of market volatility.

Second, we vet whether our results remain robust for other advanced economies. To do so, we collect data for 5 additional countries: Canada, France, Germany, Switzerland and the United Kingdom. We report on Table 2.A.2 of the Appendix the differences in RMSE, for each country, between the same benchmark models⁵ as in Figure 2.5.1 and our factor model for a 12-month horizon forecasting exercise. As can be seen on the Table, on the 15 different specifications considered here, our factor model appears to outperform the benchmarks in 12 cases. For France and the United Kingdom our factor model exhibits a lower RMSE compared to a univariate regression based on the lagged IP growth, but the difference does not appear significant. Only with respect to French term spread does our factor model display a higher RMSE when it comes to forecasting IP growth.

⁵For each country, the Market DY, the IP growth and the term spread are all collected from Refinitiv Datastream.

2.5.2 Comparison with traditional factor models

In addition, we investigate whether our factor, based on sectoral equity variables, can be used to improve more conventional factor models that rely on macroeconomic variables and on aggregated financial indicators. Indeed, whereas sectoral equity variables are easily available and published without lags, they seem to be rarely used in the forecasting literature relying on large datasets (Barhoumi et al., 2010, Hepenstrick and Marcellino, 2016, Jardet and Meunier, 2022).

To do so, we build a large dataset of 147 variables that includes aggregate macroeconomic indicators (CPI, unemployment rates), disaggregated macroeconomic variables (sectoral retail sales, sectoral industrial production indices) and aggregate financial indicators (exchange rates, interest rates and equity variables). A detailed list of the variables used is available in Table 2.A.3 of the Appendix. In the spirit of Stock et al. (2002), we then extract factors \mathbf{H}_t from this dataset with a simple Principal Component Analysis⁶. The question is then whether our factor, based on disaggregated equity variables, F_t , helps to improve the out-of-sample forecasts made with PCA-factors \mathbf{H}_t , without the use of these precise variables.

To do so, based on the same rolling window length, we compare the forecasts made by estimating a model relying on the PCA-factors:

$$y_{t+h} = \beta_0 + \beta_1' \mathbf{H}_t + u_{t+h} \quad (2.13)$$

And a model relying on the PCA-factors along with the lag of the predicted variable:

$$y_{t+h} = \beta_0 + \beta_1' \mathbf{H}_t + \beta_2 y_t + u_{t+h} \quad (2.14)$$

⁶We applied Dickey-Fuller tests to all the variables and transform them into growth rates in cases where we could not reject the null hypothesis of a unit root. We make several exceptions to that rule though, in the sense that we also include the benchmark variables of Section 2.5.1 in levels and we also incorporate several financial variables in log returns.

With the same models augmented with our factor, that is:

$$y_{t+h} = \beta_0 + \beta_1' \mathbf{H}_t + \beta_2 F_t + u_{t+h} \quad (2.15)$$

And:

$$y_{t+h} = \beta_0 + \beta_1' \mathbf{H}_t + \beta_2 y_t + \beta_3 F_t + u_{t+h} \quad (2.16)$$

We are agnostic regarding the number of relevant PCA-factors and therefore include in our regressions 1 to 3 PCA-factors. Table 2.5.2 below summarizes the differences in RMSE of the aforementioned models, augmented or not with our factor stemming from the sectoral equity variables. As the models that we compare are nested, the reported p-values in Table 2.5.2 stem from Clark and West (2007) tests.

Table (2.5.2) Difference in RMSE with alternative factor models (Factor model Corresponding PCA-factor benchmark, Out-of-sample estimates)

Benchmark:	Horizon		
	12 months	18 months	24 months
1 PCA-Factor	-5.72*	-13.31*	-8.16*
1 PCA-Factor & lag IP growth	-3.43**	-8.49	-7.79**
2 PCA-Factors	-0.17**	-0.16**	-0.58*
2 PCA-Factors & lag IP growth	-0.85	-1.29	-7.54
3 PCA-Factors	-0.18***	-0.36*	-0.76*
3 PCA-Factors & lag IP growth	-5.62**	-0.69	-6.55

*Note: The table reports the difference in RMSE of the models indicated in the first columns (augmented with the factor F_t stemming from the sectoral equity variables) with respect to the same models without this specific factor. A negative value means that augmenting the model with the factor F_t improves the RMSE. Stars represent the Clark and West (2007) test p-values under the null hypothesis of equal MSPE. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$*

In Table 2.5.2, notice that augmenting the PCA-factors with the factor built with the sectoral DYs

improves the RMSE in virtually all cases, with RMSE gains being significant in two thirds of the considered cases. This highlights the extra information that can be gained with disaggregated equity variables.

2.5.3 Performance by Sample Period

In the Introduction, we outlined that the gains of relying on sectoral rather than on aggregate equity variables may especially be strong in times of negative economic growth, such as during the pandemic. This may be the case if, for example, in these periods aggregate DY is driven mostly by sectors which are only loosely linked to the future economic activity, or if variations in aggregate DY reflect more changes in investors discount rates/expected returns rather than changes in earnings expectations.

Although we return to more formally discuss these economic mechanisms in Section 2.6, in this section, we investigate whether the forecasting performance of our factor model differs between periods of contraction and of expansion. In Table 2.5.3, we define periods of contraction as months during which the annual IP growth is negative (and the reverse for periods of expansion). In line with Moench and Stein (2021), the Table reports the difference in RMSE between our factor model based on sectoral equity variables and the same univariate model benchmarks outlined in Section 2.5.1 (along with the p-values of Diebold Mariano tests). Note that we segment here our estimation according to the dates in which the forecasts are made. In other words, if we consider here a forecast horizon of 12 months, the “Negative IP growth” period refers to predictions made when the annual IP growth was negative (and not predictions made 12 months before the contraction in economic activity).

Note that in Table 2.5.3, although our factor model outperforms other benchmarks both in periods of negative and positive IP growth, the gain in forecast accuracy of our factor model appears to be strongly concentrated in negative IP growth period. The difference between the two periods can be substantial: looking at the 12-month horizon for example, relying on our factor based

Table (2.5.3) Difference in RMSE by Period (Factor model Corresponding benchmark, Out-of-sample estimates)

<i>Benchmark:</i>	<i>Period:</i>	<i>Horizon</i>		
		12 months	18 months	24 months
Market DY	Negative IP growth	-3.8*	-7.06**	-6.82***
Market DY	Positive IP growth	-1.05**	-0.82***	-0.33
Term spread	Negative IP growth	-3.47*	-8.05**	-6.62***
Term spread	Positive IP growth	-0.67**	0.17	-0.58
Lagged IP growth	Negative IP growth	-8.71**	-9.59***	-21.71***
Lagged IP growth	Positive IP growth	-2.31**	-3.21***	-1.5***

*Note: The table reports the difference in RMSE of the factor model compared to the different benchmarks (a negative value means that the factor model outperforms the corresponding benchmark in terms of RMSE). Stars represent the Diebold-Mariano test p-values under the null hypothesis that the factor model performs worse than the benchmark models indicated in the first column. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$*

on sectoral DYs rather than on the aggregate DY can yield a RMSE-gain close to 4 times higher in negative IP growth period than in positive growth period.

One potential interpretation is that expected returns/discount rates are more volatile during recessions (Henkel et al., 2011), and can therefore blur the forecasting ability of the aggregate DY in those times. In contrast, as outlined in next section, given that our factor model filters out the expected returns component of sectoral DYs, it can yield strong forecasting accuracy gains in periods of contracting economic activity. As an example, in 2009, close to end of the Great Recession, aggregate DY was still very high, notably because investors risk aversion, and thus investors discount rates, were very high as well. As a result, the 12-month ahead IP growth forecast from the aggregate DY was still very pessimistic (-29.1% in May 2009 for the next year IP growth). On the reverse, the forecast from the factor model was much closer to the realized IP growth at the same time (+6.2% against a realized value, in May 2010, of +7.9%), probably

because the forecasting ability of our factor model was not blurred by this elevated discount rate component.

2.6. Economic Interpretation

2.6.1 Filtering the “return” and the “foreign cash flow” components

In some ways, it should be unsurprising that predictions based on factors extracted from the cross section of sectoral portfolio variables should outperform predictions based on an aggregate variable, given that aggregate measures average out important information, and at the same time include information not directly relevant to the variable being forecasted. The question is whether one can estimate the factors with sufficient precision that one outperforms a simple model using an aggregate index. In our case, the economically important information gleaned using our approach yields a substantial gain in prediction.

In this section, we further investigate how the results can be interpreted in economic terms. [Kelly and Pruitt \(2013\)](#) show that, while trying to predict future aggregate returns with disaggregated book-to-market ratios, their factor model puts *positive* weights on all sectoral book-to-market ratios, especially for “growth” portfolios (i.e. portfolios with low book-to-market ratios) which are known to be very much affected by future aggregate returns. However, some of these sectoral book-to-market ratios are positively correlated with future aggregate dividends, whereas others are negatively correlated with future aggregate dividends. Consequently, the factor, which is a weighted sum of the sectoral portfolios book-to-market ratios, will be very positively correlated with future aggregate returns but little exposed to future aggregate dividends. Similarly, when they try to forecast future aggregate dividends, they show that their factor is very positively correlated with future aggregate dividends but little exposed to future aggregate returns.

In our analysis, we replicate this exercise to identify what is filtered out in our factor model based on disaggregated DYs. To show how we do this, we display on [Figure 2.6.1](#) three variables. In

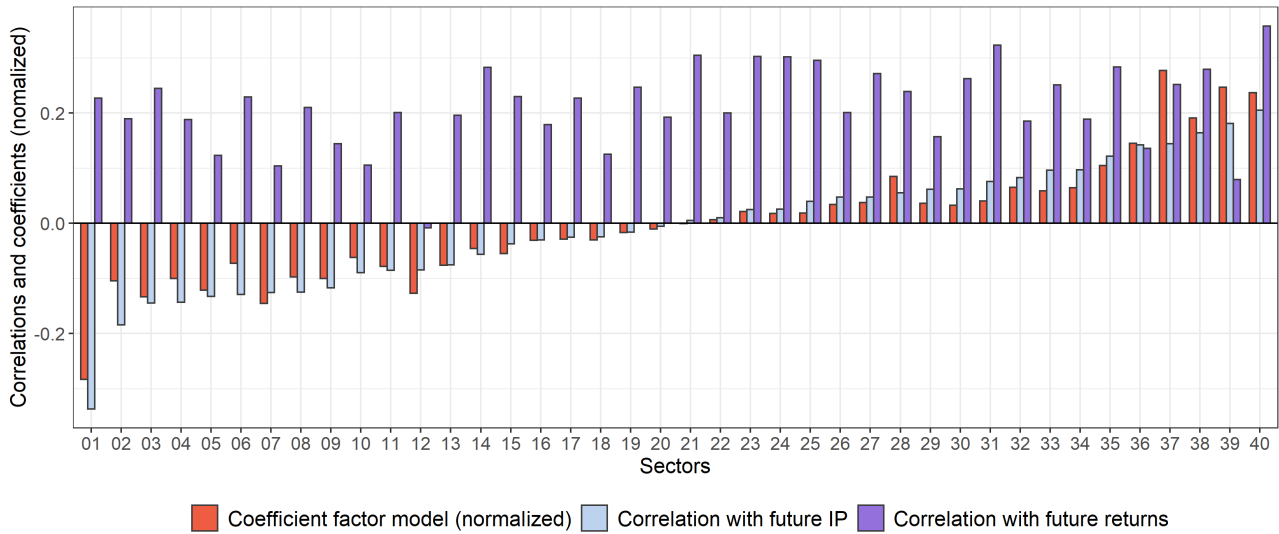
red are represented, for each of the sectors, the weights $(\hat{\phi}_i - \bar{\phi})$ that correspond to the relative importance of each sector in the factor estimation as outlined in Section 2.3.1⁷. In blue are represented the correlations of each sectoral DY with the predicted variable (IP growth, y_{t+h}) that is $\text{corr}(y_{t+h}, x_{it})$. Displayed in purple are the correlations of each sectoral DY with the aggregate equity returns compounded over the forecasting horizon r_{t+h} , that is $\text{corr}(r_{t+h}, x_{it})$. As in Kelly and Pruitt (2013) and throughout Section 2.6, we perform this analysis by examining in-sample estimates of the weights $(\hat{\phi}_i - \bar{\phi})$, while the different correlations are computed on the overall sample. We consider here, and also for the remaining of Section 2.6, a forecasting exercise over the a 12-month horizon. Finally, for visual purposes, we normalized the sector weights so that their cross-sectional standard deviation equals the standard deviations of the correlations between sectoral DYs and future IP growth.

Figure 2.6.1 clearly highlights the fact that positive weights tend to be associated with positive correlation of the sectoral DYs with future IP growth, whereas negative portfolio weights tend to be associated with negative correlation of the sectoral DYs with future IP growth. In contrast, both positive and negative portfolio weights are associated with the positive correlations of the sectoral DYs with future aggregate returns. As a result, the estimated factor –which equals the weighted sum of the sectoral DYs– is strongly exposed to future IP growth, but little exposed to future aggregate returns, in a fashion similar to what Kelly and Pruitt (2013) found.

A visual way to notice this filtering can be done by representing our factor, estimated in-sample, over time. We therefore depict on Figure 2.A.3 in the Appendix our factor along with the aggregate Market DY and the IP growth lead by 12 month. We can thus see on the Figure that, during the 90s, our factor appears to track relatively well the future IP growth. On the reverse, the (opposite of the) aggregate DY exhibits an upward trend over the period, probably linked with the fact that, amidst the so-called “irrational exuberance” (Shiller, 2015) of the dotcom bubble,

⁷Unlike Kelly and Pruitt (2013), for this analysis we rely on the centered weights $(\hat{\phi}_i - \bar{\phi})$, whereas they rely on the *uncentered* weights $\hat{\phi}_i$. Our approach seems more appropriate to us, given that the relationship between the sectoral DYs and the estimated factors is given precisely by the centered weights: $\hat{F}_t = \frac{1}{C} \sum_i x_{it}(\phi_i - \bar{\phi})$.

Figure (2.6.1) Factor weights and DYs correlations with future IP growth and future aggregate returns (In-sample estimates, forecasting over a 12-month horizon)



Note: The Figure represents the estimate factor weights (in red), the correlation of sectoral DYs with future IP growth (in blue) or with future aggregate returns (in purple). For visual purposes, the sector weights are normalized so that their cross-sectional standard deviation equals the standard deviations of the correlations between sectoral DYs and future IP growth. Correlations are computed on the overall dataset, while the coefficients stem from an in-sample estimation of the factor model based on a forecast horizon of 12 months.

investors were requiring very low discount rates which tended to push stock prices significantly high. As our factor model purges the discount rates/expected returns component of aggregate DY, it is less affected by this trend, and therefore spots more accurately movements in future IP growth.

Additionally, we want our factor model to not only to filter out the “expected returns” component of the sectoral DYs, but to also filter out the “foreign cash flow” component. In other words, relying on the notations of Section 2.2.1, we would like $corr(\hat{F}_t, \Delta cf_{D,t+j})$ to be high and $corr(\hat{F}_t, r_{t+j})$ and $corr(\hat{F}_t, \Delta cf_{F,t+j})$ to be low.

However, whereas we can directly observe the levels of future aggregate returns, we need to rely on a proxy to assess the correlation between our estimated factor and the aggregate foreign cash flow component. Since the latter theoretically represents the component of the sectoral DYs

that reflect the foreign profitability of the US firms, we rely on the foreign industrial production indices of [Grossman et al. \(2014\)](#). The index that we consider here, $IP_{F,t}$, corresponds to the level of industrial activity of advanced economies, excluding the US.

Note that US IP and $IP_{F,t}$ are of course strongly correlated. Therefore, a direct assessment whether the factor model filters out adequately the future foreign activity component of sectoral DYs with $IP_{F,t}$ is likely to give biased results precisely because the estimated factor is itself positively correlated with US IP growth. On the other hand, we would like our factor model to filter out the part of foreign activity that is orthogonal to US economic activity. To do so we first regress foreign IP growth ($IP_{F,t}$) on US IP growth (y_t):

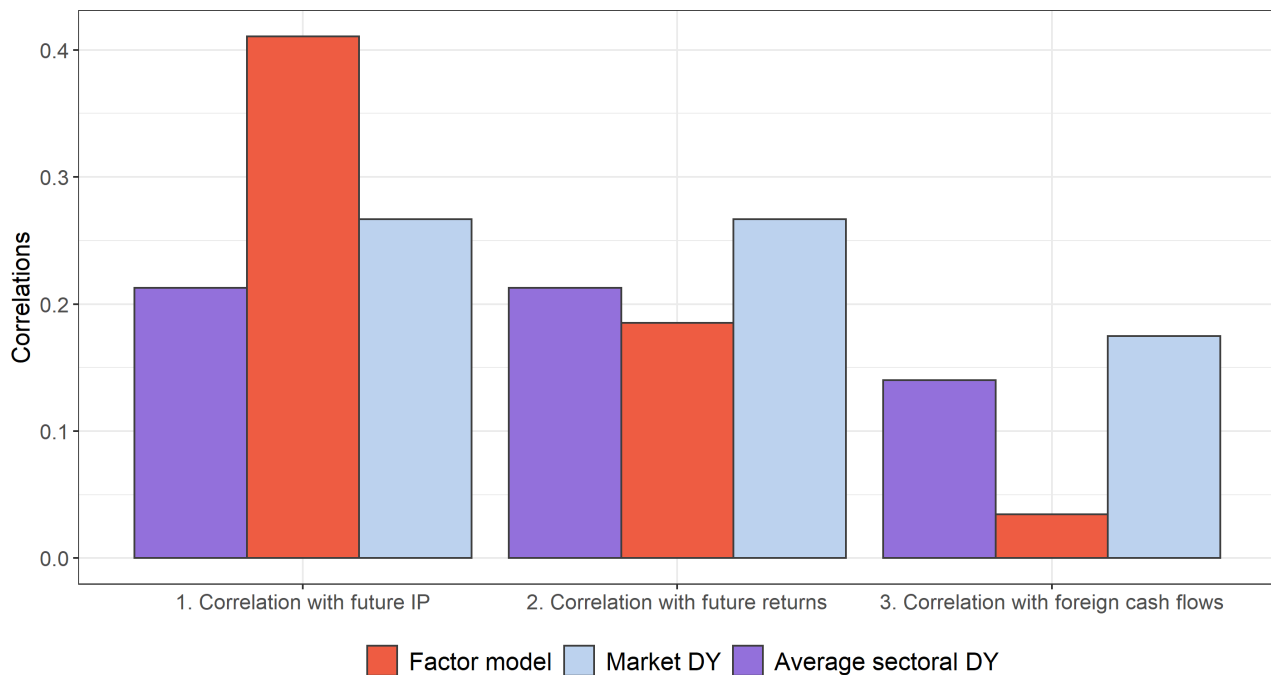
$$IP_{F,t} = \alpha + \beta y_t + u_t \quad (2.17)$$

And rely on the estimated error terms (\hat{u}_t) to conduct our analysis.

Figure 2.6.2 summarizes the different filterings that we consider in this section. Again, the analysis is performed here on an in-sample basis and for the 12-month prediction exercise. In red are represented the correlations of the estimated factor (\hat{F}_t) with future US IP growth (y_{t+h}), with future aggregate US returns (r_{t+h}) or with the component of future foreign IP growth that is orthogonal to future US IP growth (\hat{u}_{t+h}). In light blue are represented the same quantities but for the aggregate DY instead of the estimated factor. Finally, in purple are pictured the average correlation of the sectoral DYs with the aforementioned variables (that is $\frac{1}{I} \sum_i \text{corr}(x_{it}, y_{t+h})$, $\frac{1}{I} \sum_i \text{corr}(x_{it}, r_{t+h})$ and $\frac{1}{I} \sum_i \text{corr}(x_{it}, \hat{u}_{t+h})$).

In line with Figure 2.6.1, we can see in Figure 2.6.2 that the estimated factor is more correlated to future IP growth, and less correlated to future aggregate returns than the Market DY or than the sectoral DYs (on average). Additionally, Figure 2.6.2 also highlights that the estimated factor is clearly less correlated with the future foreign cash flow component, here proxied by our estimates \hat{u}_{t+h} . In other words, our factor model appears to play this role: by over/underweight-

Figure (2.6.2) Factor correlations along with Sectoral and Aggregate DY correlations (In-sample estimates, forecasting over a 12-month horizon)



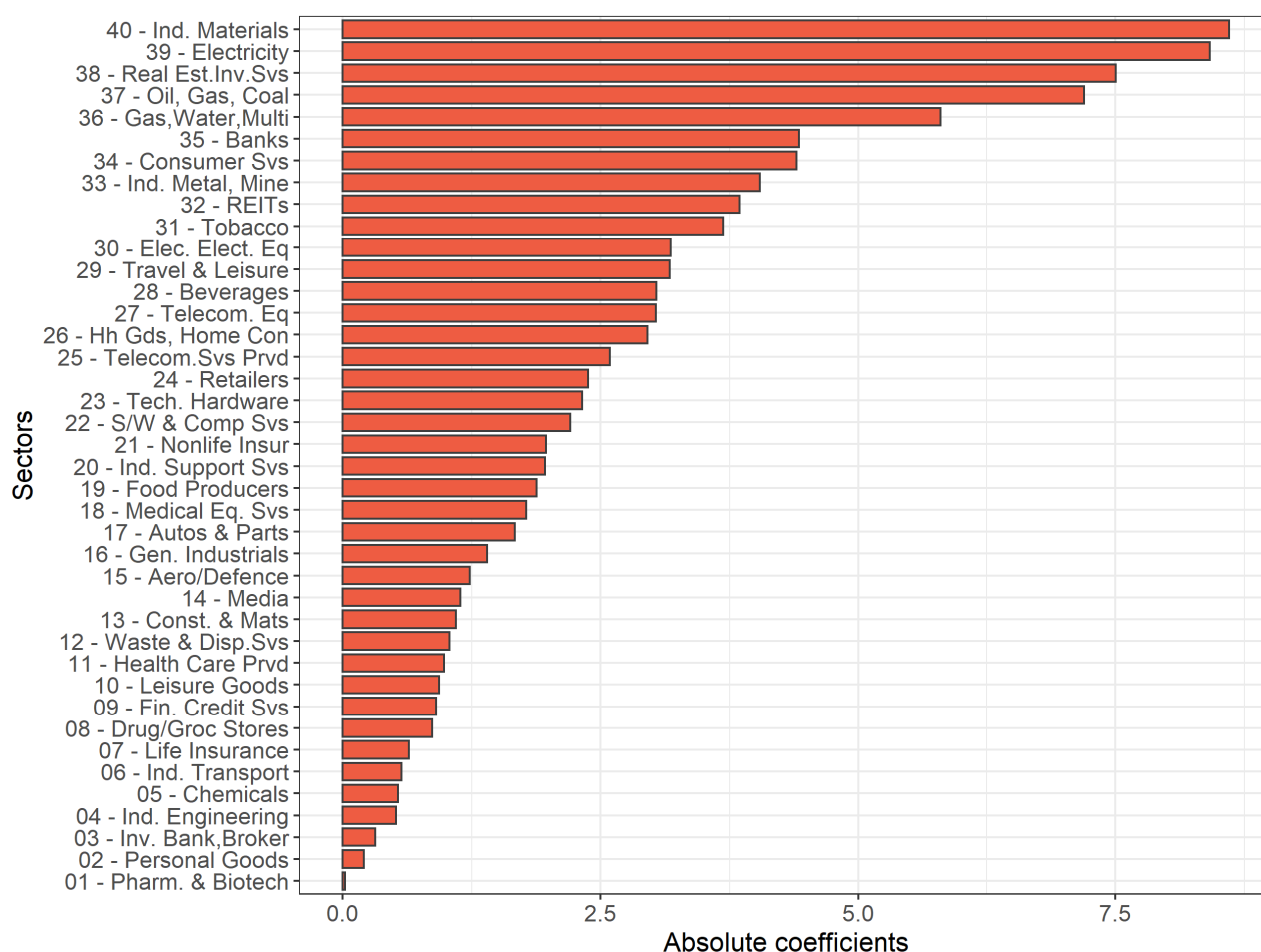
Note: On the Figure above are represented in red the correlations of the estimated factor with future US IP growth, with future aggregate US returns or with the component of future foreign IP growth that is orthogonal to future US IP growth. In light blue are represented the same quantities but for the aggregate DY instead of the estimated factor. Eventually in purple are pictured the average correlation of the sectoral DYs with the aforementioned variables. Correlations are computed on the overall dataset, while the estimated factor stems from an in-sample estimation of the factor model based on a forecast horizon of 12 months.

ing certain sectors it increases the correlation with our predicted variable while filtering out the noisy components of the sectoral DYs.

2.6.2 sector overweighting

We investigate further the economic analysis of the outperformance of our factor model by identifying more precisely which sectors are overweighted in this exercise. To do so, in Figure 2.6.3, we depict the (absolute) weights ($\phi_i - \bar{\phi}$) to understand which sectoral DYs affect the most the estimated factor. Here also we conduct this analysis on an in-sample basis, with a forecast horizon of 12 months.

Figure (2.6.3) Absolute factor weights (In-sample estimates, forecasting over a 12-month horizon)



Note: the graph represents the absolute factor weights $|(\hat{\phi}_i - \bar{\phi})|$, estimated in an in-sample forecasting exercise over a 12-month horizon.

Several findings emerge from inspecting Figure 2.6.3. First we notice that the factor model overweights strongly *upstream* sectors, i.e. sectors that mainly produce inputs for manufacturing and services (Oil, Gas and Coal; Industrial Materials; Electricity, Gas and Water; Industrial Metals). Second, the factor model appears also to put more weights on industries related to the real estate sector, like Real Estate Investment Trusts (REITs) or Real Estate Investment and Services, probably due the strong link between property price dynamics and the business cycle (Leamer, 2015, Borio et al., 2020).

We further investigate which sectors appear to have the more importance in our factor model by testing two additional hypotheses:

- Are “value” sectors, i.e. sectors that are little valued by equity investors and therefore exhibit low Price-Earnings Ratios (PER), overweighted compared to “growth” sectors, which, in contrast, have elevated PER. Value sector equities, like the automobile sector, are sometimes deemed to be more closely linked to the future business cycle as investors may estimate that they are less able to downsize their activity in case of an incoming recession (Kojien et al., 2017, Xu, 2018).
- To what extent does our factor model overweight sectors whose DYs are correlated with future domestic IP growth compared to sectors with a high exposure on foreign economic activity.

To do so, we estimate the following cross-sectional regression:

$$|(\hat{\phi}_i - \bar{\phi})| = \alpha + \beta_1 |corr(y_{t+h}, x_{it})| + \beta_2 PER_i + \beta_3 |corr(E_t, x_{it})| + \alpha_i + u_i \quad (2.18)$$

Where $|corr(y_{t+h}, x_{it})|$ represents, for the sector i , the absolute correlation of the sectoral DY with future IP growth, PER_i stands for the average PER of the sector i on the overall period, $|corr(E_t, x_{it})|$ represents the absolute correlation of the sectoral DY with either the US real effective exchange rate, REER, retrieved from the BIS website, or with our metric of future foreign IP growth that is orthogonal to future US IP growth (\hat{u}_{t+h}). Finally, α_i stands for the industry fixed effects (where the 40 sectors that we are relying on are regrouped in 11 different industries in the IBC classification).

Table 2.6.1 presents the regression results. Here again, the coefficients $|(\phi_i - \bar{\phi})|$ are from an in-sample estimation of the factor based on a 12-month horizon.

Table (2.6.1) Absolute factor weights regressions (In-sample estimates, forecasting over a 12-month horizon)

	<i>Dependent variable:</i>			
	Abs. Factor coefficients			
	(1)	(2)	(3)	(4)
Abs. corr. Future IP with sectoral DYs	25.791*** (2.175)	27.352*** (1.827)	27.202*** (1.594)	27.288*** (1.577)
Average PER		-0.130*** (0.042)	-0.142*** (0.039)	-0.121*** (0.035)
Abs. corr. Exchange rate with sectoral DYs			-1.757** (0.732)	
Abs. corr. Foreign IP with sectoral DYs				-1.792** (0.873)
Constant	0.034 (0.388)	2.611*** (0.905)	3.150*** (0.838)	2.697*** (0.800)
Observations	40	40	40	40
R ²	0.889	0.925	0.936	0.932
Adjusted R ²	0.846	0.892	0.904	0.899
F Statistic	20.470*** (df = 11; 28)	27.832*** (df = 12; 27)	29.325*** (df = 13; 26)	27.622*** (df = 13; 26)

*Note: All regressions include industry-level fixed effects. The reported regressions are made using White heteroscedasticity-robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$*

We can see first in Table 2.6.1 that, by construction and in absolute terms, factor weights are strongly and positively related with the correlation between sectoral DYs and future IP growth. Second, Table 2.6.1 underlines that, in line with the hypothesis formulated above, the DYs from the value sectors seem to contain relatively more information regarding future IP growth given that lower PERs are positively associated with the factor weights in our regressions. Third, it appears that our factor significantly underweights sectors whose DYs are strongly correlated, in absolute terms, with the US REER or with our metric of foreign IP growth. This would mean that our estimated factor puts less weight on sectors with a strong exposure on foreign economic activity, so as to better spot changes in future domestic IP growth.

2.7. Conclusion

We show that a factor model based on sectorally disaggregated stock market variables can significantly outperform other extant macroeconomic forecasting models. Previous approaches either relied on aggregate equity variables, on disaggregated equity variables taken in isolation, or on indices built from disaggregated equity variables but in a constrained manner (for example by using the Fama-French factors). We show that our factor model outperforms –over several different horizons– both in-sample and out-of-sample, the usual macroeconomic benchmarks. We attribute this out-performance to two characteristics of our factor model. First, we show that our model over/underweights certain sectors so that the resulting factor is strongly associated with future IP growth, but is, conversely, relatively less associated with the noisy components of the sectoral DYs, namely expected returns and the foreign component of future cash flows. Second, we argue that the superior performance of our model is related to the fact that it overweights both upstream sectors (Oil and Gas, Industrial Materials etc.) and value sectors that are deemed relatively more informative regarding future IP growth. As a consequence, we are able to better predict activity overall, but also particularly during periods of negative growth.

Bibliography

- Andersson, B. and Agostino, A. D. . (2008). Are sectoral stock prices useful for predicting euro area gdp? *ECB Working Paper*.
- Barhoumi, K., Darné, O., Ferrara, L., Barhoumi, K., Darné, O., and Ferrara, L. (2010). Are disaggregate data useful for factor analysis in forecasting french gdp? *Journal of Forecasting*, 29:132–144.
- Bessec, M., Doz, C., Bessec, M., and Doz, C. (2012). Prévission à court terme de la croissance du pib français à laide de modèles à facteurs dynamiques. *Économie et Prévission*, 199:1–30.
- Binsbergen, J. H. V. and Kojen, R. S. (2010). Predictive regressions: A present-value approach. *Journal of Finance*, 65:1439–1471.
- Binswanger, M. (2000). Stock returns and real activity: Is there still a connection? *Applied Financial Economics*, 10:379–387.
- Borio, C., Drehmann, M., and Xia, F. D. (2020). Forecasting recessions: the importance of the financial cycle. *Journal of Macroeconomics*, 66.
- Browne, F. and Doran, D. (2005). Do equity index industry groups improve forecasts of inflation and production? a us analysis. *Applied Economics*, 37:1801–1812.
- Campbell, J. and Ammer, J. (1993). What moves the stock and bond markets? a variance decomposition for longterm asset returns. *The Journal of Finance*, 48:3–37.
- Campbell, J. Y. and Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance*, 43:661–676.
- Chatelais, N. and Stalla-Bourdillon, A. (2020). What are the factors behind current high stock market valuations? *Banque de France Blog*.

- Chen, S. and Ranciere, R. (2019). Financial information and macroeconomic forecasts. *International Journal of Forecasting*, 35:1160–1174.
- Chinn, M. and Kucko, K. (2015). The predictive power of the yield curve across countries and time. *International Finance*, 18(2):129–156.
- Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138:291–311.
- Croux, C. and Reusens, P. (2013). Do stock prices contain predictive power for the future economic activity? a granger causality analysis in the frequency domain. *Journal of Macroeconomics*, 35:93–103.
- Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 20:134–144.
- Fan, J., Xue, L., and Yao, J. (2017). Sufficient forecasting using factor models. *Journal of Econometrics*, 201:292–306.
- Ferrara, L. and Marsilli, C. (2019). Nowcasting global economic growth: A factor-augmented mixed-frequency approach. *The World Economy*, 42:846–875.
- Grossman, V., Mack, A., Martínez-García, E., Grossman, V., Mack, A., and Martínez-García, E. (2014). A new database of global economic indicators. *Journal of Economic and Social Measurement*, pages 163–197.
- Henkel, S. J., Martin, J. S., and Nardari, F. (2011). Time-varying short-horizon predictability. *Journal of Financial Economics*, 99:560–580.
- Hepenstrick, C. and Marcellino, M. (2016). Forecasting with large unbalanced datasets: The mixed-frequency three-pass regression filter. *SNB Working Papers*.
- Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, 28:791–837.

- Jagannathan, R. and Marakani, S. (2015). Price-dividend ratio factor proxies for long-run risks. *The Review of Asset Pricing Studies*, 5:1–47.
- Jardet, C. and Meunier, B. (2022). Nowcasting world gdp growth with high-frequency data. *Journal of Forecasting*, 41:1181–1200.
- Kelly, B. and Pruitt, S. (2013). Market expectations in the cross-section of present values. *Journal of Finance*, 68:1721–1756.
- Koijen, R. S., Lustig, H., and Nieuwerburgh, S. V. (2017). The cross-section and time series of stock and bond returns. *Journal of Monetary Economics*, 88:50–69.
- Kumar, A., Mallick, S., Mohanty, M., and Zampolli, F. (2022). Market volatility, monetary policy and the term premium. *Oxford Bulletin of Economics and Statistics*.
- Lan, C., Roussanov, N., Lan, C., and Roussanov, N. (2020). Stock price movements: Business-cycle and low-frequency perspectives. *The Review of Asset Pricing Studies*, 10:335–395.
- Leamer, E. E. (2015). Housing really is the business cycle: What survives the lessons of 200809? *Journal of Money, Credit and Banking*, 47:43–50.
- Liew, J. and Vassalou, M. (2000). Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics*, 57:221–245.
- Loungani, P., Rush, M., and Tave, W. (1990). Stock market dispersion and unemployment. *Journal of Monetary Economics*, 25:367–388.
- McMillan, D. G. (2021). Predicting gdp growth with stock and bond markets: Do they contain different information? *International Journal of Finance Economics*, 26:3651–3675.
- Moench, E. and Stein, T. (2021). Equity premium predictability over the business cycle. *Bundesbank Discussion Paper*.

- Shiller, R. J. (2015). *Irrational exuberance: Revised and expanded third edition*. Princeton university press.
- Stock, J., Watson, M., Stock, J., and Watson, M. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business Economic Statistics*, 20:147–62.
- Venditti, F. and Veronese, G. (2020). Global financial markets and oil price shocks in real time. *ECB Working Paper*.
- Vuolteenaho, T. (2002). What drives firm-level stock returns? *Journal of Finance*, 57:233–264.
- West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica*, 64:1067.
- Xu, J. (2018). Essays on the value effect in the time series and cross section of stock returns. *Hong Kong Polytechnic University - Thesis*.
- Zalgiryte, L., Guzavicius, A., and Tamulis, V. (2014). Stock market and economic growth in the u.s. france: Evidence from stock market sector indices. *Engineering Economics*, 25:47–53.
- Zhang, L. (2005). The value premium. *The Journal of Finance*, 60:67–103.
- Zhu, S., Gao, J., and Sherman, M. (2020). The role of future economic conditions in the cross-section of stock returns: Evidence from the us and uk. *Research in International Business and Finance*, 52:101193.
- Ólan Henry, Olekalns, N., Thong, J., Ólan Henry, Olekalns, N., and Thong, J. (2004). Do stock market returns predict changes to output? evidence from a nonlinear panel data model. *Empirical Economics*, 29:527–540.

Appendix

2.A. Appendix

2.A.1 The Factor model for sectoral and aggregate DYs

We use the sectoral DYs x_{it} in a factor model instead of the aggregate DY x_t to predict future IP growth. By doing so, we are implicitly assuming that the sectoral DYs are indicative of future aggregate domestic cash flows, which are themselves a proxy for the future US economic activity. We are also assuming that the factor model is able to isolate this information while filtering the remaining noisy components in sectoral DYs.

More precisely, in line with [Kelly and Pruitt \(2013\)](#), we are assuming that the expectation of sectoral returns, sectoral domestic cash flow growth and sectoral foreign cash flow growth are linearly determined by a set of common factors \mathbf{F}_t :

$$\begin{aligned} E_t(r_{i,t+1}) &= \alpha_{i,0} + \boldsymbol{\alpha}'_{i,1} \mathbf{F}_t + u_{i,t} \\ E_t(\Delta cf_{D,i,t+1}) &= \beta_{i,0} + \boldsymbol{\beta}'_{i,1} \mathbf{F}_t + e_{i,t} \\ E_t(\Delta cf_{F,i,t+1}) &= \gamma_{i,0} + \boldsymbol{\gamma}'_{i,1} \mathbf{F}_t + \epsilon_{i,t} \end{aligned}$$

Where $u_{i,t}$, $e_{i,t}$ and $\epsilon_{i,t}$ are idiosyncratic and independently distributed components with $E_t(u_{i,t+1}) = E_t(e_{i,t+1}) = E_t(\epsilon_{i,t+1}) = 0$.

The expectations of aggregate variables follow similar processes, that is:

$$\begin{aligned} E_t(r_{t+1}) &= \alpha_0 + \boldsymbol{\alpha}'_1 \mathbf{F}_t + u_t \\ E_t(\Delta cf_{D,t+1}) &= \beta_0 + \boldsymbol{\beta}'_1 \mathbf{F}_t + e_t \\ E_t(\Delta cf_{F,t+1}) &= \gamma_0 + \boldsymbol{\gamma}'_1 \mathbf{F}_t + \epsilon_t \end{aligned}$$

Finally, we assume that the factors follow an autoregressive process:

$$\mathbf{F}_{t+1} = \boldsymbol{\Theta} \mathbf{F}_t + \boldsymbol{\nu}_{t+1}$$

Therefore, in line with Section 2.2.1, we can use the Campbell and Shiller (1988) formula for sectoral DYs:

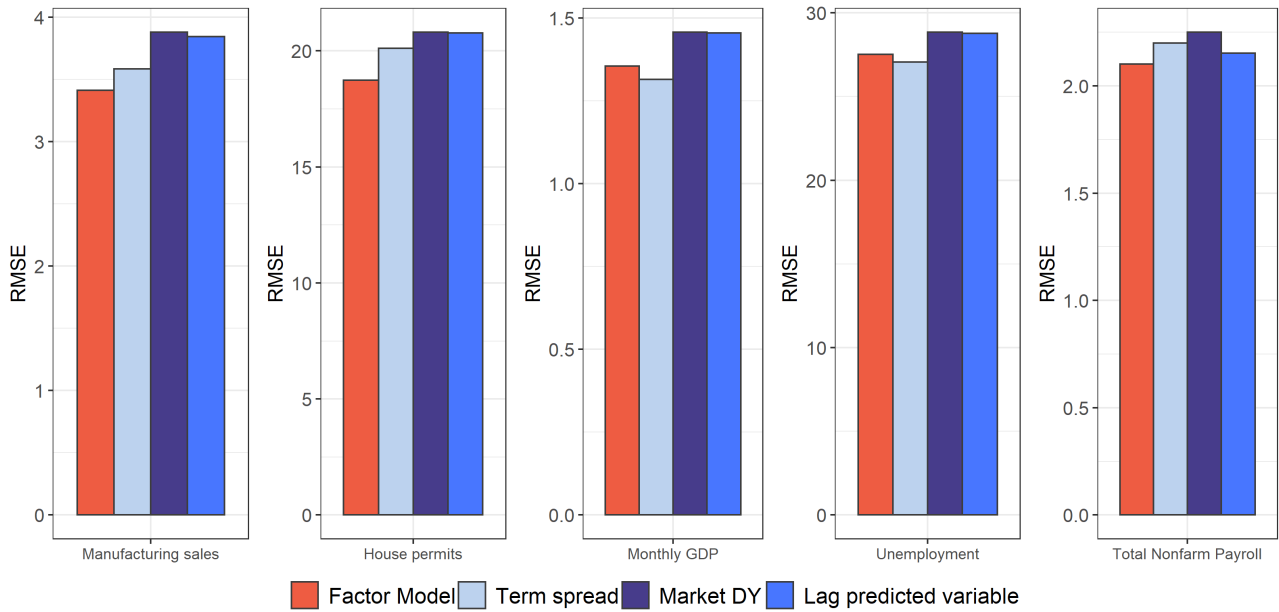
$$\begin{aligned}
x_{it} &= \frac{\kappa_i}{(1 - \rho_i)} + \sum_{j=1} \rho_i^{j-1} E_t[r_{i,t+j} - \Delta cf_{D,i,t+j} - \Delta cf_{F,i,t+j}] \\
&= \frac{\kappa_i}{(1 - \rho_i)} + \sum_{j=1} \rho_i^{j-1} E_t[(\alpha_{i,0} + \boldsymbol{\alpha}'_{i,1} \mathbf{F}_t + u_{i,t}) - (\beta_{i,0} + \boldsymbol{\beta}'_{i,1} \mathbf{F}_t + e_{i,t}) - (\gamma_{i,0} + \boldsymbol{\gamma}'_{i,1} \mathbf{F}_t + \epsilon_{i,t})] \\
&= \frac{\kappa_i + \alpha_{i,0} - \beta_{i,0} - \gamma_{i,0}}{1 - \rho_i} + \sum_{j=1} \rho_i^{j-1} E_t[\mathbf{i}' \boldsymbol{\Gamma}'_i \mathbf{F}_{t+j-1} + u_{i,t+j-1} - e_{i,t+j-1} - \epsilon_{i,t+j-1}] \\
&= \frac{\kappa_i + \alpha_{i,0} - \beta_{i,0} - \gamma_{i,0}}{1 - \rho_i} + \mathbf{i}' \boldsymbol{\Gamma}'_i (\mathbf{I} - \rho_i \boldsymbol{\Theta})^{-1} \mathbf{F}_t + u_{it} - e_{it} - \epsilon_{it} \\
&= \phi_{i,0} + \boldsymbol{\phi}'_{i,1} \mathbf{F}_t + \nu_{i,t}
\end{aligned}$$

With $\phi_{i,0} = \frac{\kappa_i + \alpha_{i,0} - \beta_{i,0} - \gamma_{i,0}}{1 - \rho_i}$, $\boldsymbol{\phi}'_{i,1} = \mathbf{i}' \boldsymbol{\Gamma}'_i (\mathbf{I} - \rho_i \boldsymbol{\Theta})^{-1}$, $\nu_{i,t} = u_{it} - e_{it} - \epsilon_{it}$, $\mathbf{i} = (1, -1, -1)'$ and $\boldsymbol{\Gamma}_i = (\boldsymbol{\alpha}_{i,1}, \boldsymbol{\beta}_{i,1}, \boldsymbol{\gamma}_{i,1})$.

In other words, the calculus above underlines how, by assuming that common factors affects both the expectations of sectoral and aggregate returns and cash flows, we can show that sectoral DYs are linearly related to these factors. Since the latter also affect linearly future aggregate domestic cash flows, it is therefore attractive, in this framework, to rely on the cross-section of sectoral DYs to extract a predictive signal for the future domestic cash flows.

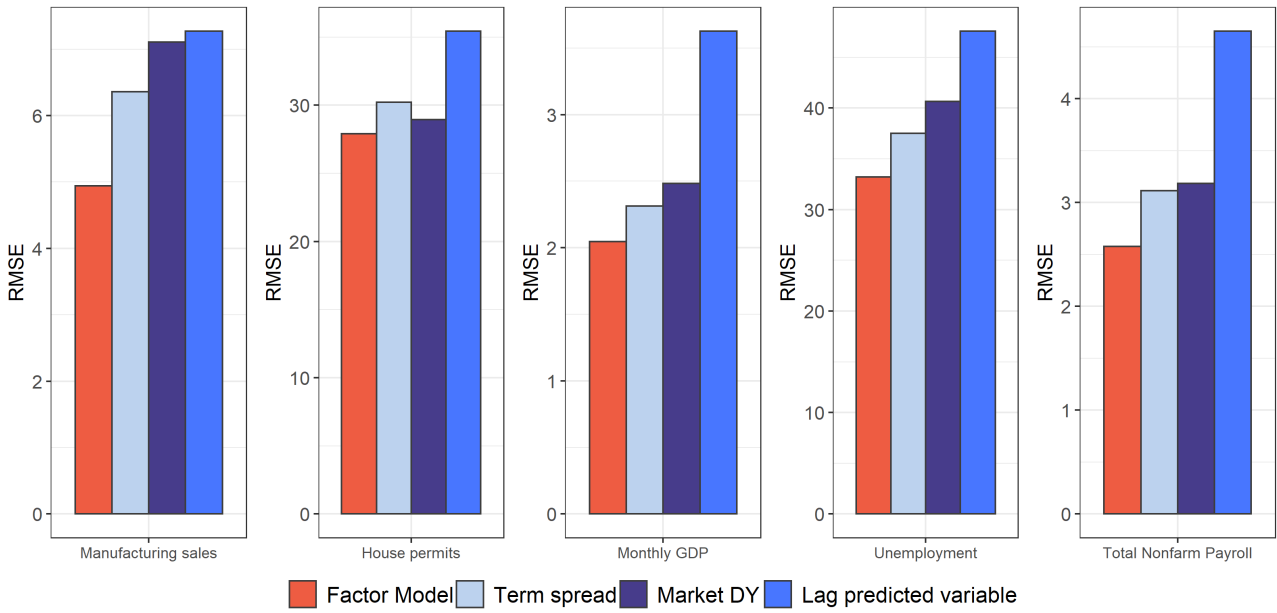
2.A.2 Additional forecasting results

Figure (2.A.1) Robustness check, In-Sample RMSE from the different estimated models



Note: On the graph are represented the In-Sample RMSE of different models (the factor model, or univariate regressions relying either on the aggregate DY, on the lagged IP growth or on the term spread). The predicted variables (Manufacturing sales, House permits etc.) are all defined as growth rates, similarly to the IP growth, before conducting the forecasting exercise.

Figure (2.A.2) Robustness check, Out-of-Sample RMSE from the different estimated models



Note: On the graph are represented the Out-of-Sample RMSE of different models (the factor model, or univariate regressions relying either on the aggregate DY, on the lagged IP growth or on the term spread). The predicted variables (Manufacturing sales, House permits etc.) are all defined as growth rates, similarly to the IP growth, before conducting the forecasting exercise.

Table (2.A.1) Difference in RMSE with volatility models (Factor model Corresponding benchmark, Out-of-sample estimates)

<i>Benchmark:</i>	<i>Horizon</i>		
	12 months	18 months	24 months
Volatility 1	-4.76*	-1.57*	-9.97*
Volatility 2	-4.82*	-2.22**	-7.28**
VIX	-2.17*	-2.09**	-0.98**
MOVE	-1.69*	-2.26*	-2.81**
Volatility 1 + Term spread	-3.69**	-3.27*	-10.46**
Volatility 2 + Term spread	-3.76**	-2.94*	-8.07**
VIX + Term spread	-3.67**	-3.27**	-2.89**
MOVE + Term spread	-2.74**	-4.06*	-5.09**

*Note: The table reports the difference in RMSE of the factor model compared to the different benchmarks (a negative value means that the factor model outperforms the corresponding benchmark in terms of RMSE). The benchmarks used in this exercise are univariate or bivariate regressions relying on a market volatility variable augmented with the term spread for the last four models. The volatility metrics are: the monthly variance of daily log returns on the US stock market, Volatility 1, the monthly sum of daily squared returns on the US stock market, à la Goyal and Welch (2008), Volatility 2, the VIX and the Merrill Lynch Option Volatility Expectations, or MOVE, a metric of bond market volatility. Stars represent the Diebold-Mariano test p-values under the null hypothesis that the factor model performs worse than the benchmark models indicated in the first column. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$*

Table (2.A.2) Difference in RMSE by country (Factor model Corresponding benchmark, Out-of-sample estimates, 12-month horizon)

<i>Benchmark:</i>	Canada	France	Germany	Switzerland	UK
Market DY	-2.13*	-3.14*	-1.27***	-5.17*	-1.5***
Term spread	-3.03*	0.16	-1.69***	-1.46***	-1.26*
Lagged IP growth	-4.2**	-0.08	-7.62*	-4.44**	-0.11
Number of sectors	21	28	24	30	38

*Note: The table reports the difference in RMSE of the factor model compared to the different benchmarks (a negative value means that the factor model outperforms the corresponding benchmark in terms of RMSE). In the same line as for our main specification (for the United States), we filter from this exercise IBC sectoral DY series that were incomplete over the time period. As a result, the number of sectors used in this analysis may differ between the different countries. Stars represent the Diebold-Mariano test p-values under the null hypothesis that the factor model performs worse than the benchmark models indicated in the first column. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$*

2.A.3 Dataset - traditional factor model

Table (2.A.3) List of the variables used to estimate PCA-factors

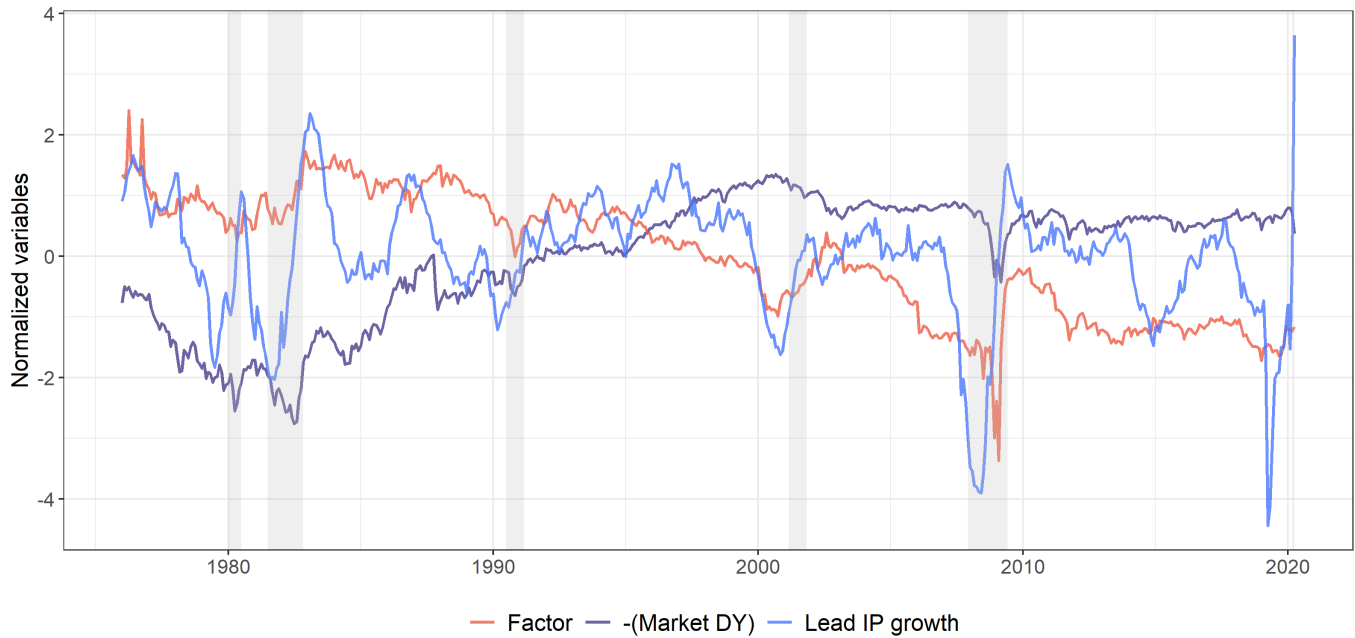
Group	Variable	Source
Consumer Price Index	CPI: All items	US BLS
Consumer Price Index	CPI: Food	US BLS
Consumer Price Index	CPI: Food at home	US BLS
Consumer Price Index	CPI: Cereals and bakery products	US BLS
Consumer Price Index	CPI: Meats, poultry, fish, and eggs	US BLS
Consumer Price Index	CPI: Dairy and related products	US BLS
Consumer Price Index	CPI: Fruits and vegetables	US BLS
Consumer Price Index	CPI: Nonalcoholic beverages and beverage materials	US BLS
Consumer Price Index	CPI: Other food at home	US BLS
Consumer Price Index	CPI: Food away from home	US BLS
Consumer Price Index	CPI: Energy	US BLS
Consumer Price Index	CPI: Energy commodities	US BLS
Consumer Price Index	CPI: Fuel oil	US BLS
Consumer Price Index	CPI: Motor fuel	US BLS
Consumer Price Index	CPI: Gasoline (all types)	US BLS
Consumer Price Index	CPI: Energy services	US BLS
Consumer Price Index	CPI: Electricity	US BLS
Consumer Price Index	CPI: Utility (piped) gas service	US BLS
Consumer Price Index	CPI: All items less food and energy	US BLS
Consumer Price Index	CPI: Commodities less food and energy commodities	US BLS
Consumer Price Index	CPI: Apparel	US BLS
Consumer Price Index	CPI: New vehicles	US BLS
Consumer Price Index	CPI: Used cars and trucks	US BLS
Consumer Price Index	CPI: Medical care commodities	US BLS
Consumer Price Index	CPI: Alcoholic beverages	US BLS
Consumer Price Index	CPI: Tobacco and smoking products	US BLS
Consumer Price Index	CPI: Services less energy services	US BLS
Consumer Price Index	CPI: Shelter	US BLS
Consumer Price Index	CPI: Rent of primary residence	US BLS
Consumer Price Index	CPI: Owners equivalent rent of residences	US BLS
Consumer Price Index	CPI: Medical care services	US BLS
Consumer Price Index	CPI: Physicians' services	US BLS
Consumer Price Index	CPI: Hospital services	US BLS
Consumer Price Index	CPI: Transportation services	US BLS
Consumer Price Index	CPI: Motor vehicle maintenance and repair	US BLS
Consumer Price Index	CPI: Motor vehicle insurance	US BLS
Consumer Price Index	CPI: Airline fares	US BLS
Equity market	S&P 500 Dividend yield	S&P Dow Jones Indices LLC
Equity market	Dow Jones Dividend yield	S&P Dow Jones Indices LLC
Equity market	US stock market Dividend yield	Refinitiv Datastream
Equity market	US stock market Price earnings ratio	Refinitiv Datastream
Equity market	US stock market Earnings	Refinitiv Datastream
Equity market	US stock market Volatility	Refinitiv Datastream
Equity market	US stock market Log-returns	Refinitiv Datastream
Equity market	S&P 500 Excess CAPE yield	Robert Shiller website
Equity market	S&P 500 Price Index	Refinitiv Datastream
Equity market	S&P 500 Cyclically Adjusted Price earnings ratio	Robert Shiller website
Equity market	S&P 500 CAPE Ratio	Refinitiv Datastream
Equity market	Fama-French Small-minus-Big Factor	Kenneth French website
Equity market	Fama-French High-minus-Low Factor	Kenneth French website
Exchange rate	Real Effective Exchange Rates Based on Manufacturing	
Exchange rate	Consumer Price Index for the US	OECD
Exchange rate	Nominal Effective Exchange Rates Based on Manufacturing	
Exchange rate	Consumer Price Index for the US	OECD
Exchange rate	Exchange rate EURUSD	Federal Reserve Board
Exchange rate	Exchange rate JPYUSD	Federal Reserve Board
Exchange rate	Exchange rate CHFUSD	Federal Reserve Board
Exchange rate	Exchange rate GBPUSD	Federal Reserve Board
Exchange rate	Exchange rate Australian dollar USD	Federal Reserve Board
Exchange rate	Exchange rate Swiss FRanc USD	Federal Reserve Board
Household statistics	US Real Disposable Personal Income	US BEA
Household statistics	US Personal Saving Rate	US BEA
Housing statistics	Revolving Home Equity Loans, All Commercial Banks	Federal Reserve Board
Housing statistics	Revolving Home Equity Loans, Small Domestically Commercial Banks	Federal Reserve Board
Housing statistics	Housing Starts: Total: New Privately Owned Housing Units Started	U.S. Census Bureau
Housing statistics	S&P/Case-Shiller U.S. National Home Price Index	S&P Dow Jones Indices LLC
Housing statistics	Housing Starts: Total: New Privately Owned Housing Units Started	U.S. Census Bureau
Housing statistics	Supply of Houses in the United State	U.S. Census Bureau
Housing statistics	New Private Housing Units Authorized by Building Permits	U.S. Census Bureau
Housing statistics	New One Family Houses Sold: United States	U.S. Census Bureau
Housing statistics	Median Sales Price for New Houses Sold in the United State	U.S. Census Bureau
Interest rate	Ted Spread	FED Saint Louis
Interest rate	10 Year US government rate	Federal Reserve Board
Interest rate	US Bank Prime Loan Rate	Federal Reserve Board
Interest rate	Federal funds rate	Federal Reserve Board
Interest rate	Term Spread	Refinitiv Datastream
Interest rate	Moody's Seasoned Aaa Corporate Bond Yield	FED Saint Louis
Interest rate	Moody's Seasoned Baa Corporate Bond Yield	FED Saint Louis
Interest rate	Baa-Aaa Bond Spread	FED Saint Louis
IP Index	Industrial Production: Manufacturing (SIC)	Federal Reserve Board
IP Index	Industrial Production: Mining : crude oil	Federal Reserve Board

Table (2.A.3) List of the variables used to estimate PCA-factors

Group	Variable	Source
IP Index	Industrial Production: durable goods : ow steel	Federal Reserve Board
IP Index	Industrial Production: durable manuf : vehicle	Federal Reserve Board
IP Index	Industrial Production: mining : gold and silver	Federal Reserve Board
IP Index	Industrial Production: mining	Federal Reserve Board
IP Index	Industrial Production: consumer good	Federal Reserve Board
IP Index	Industrial Production: durable consumer good	Federal Reserve Board
IP Index	Industrial Production: non durable manuf : food alcool beverage	Federal Reserve Board
IP Index	Industrial Production: durable manuf : machinery	Federal Reserve Board
IP Index	Industrial Production: business equipement	Federal Reserve Board
IP Index	Industrial Production: non durable manuf : chimestrey	Federal Reserve Board
IP Index	Industrial Production: durable manuf : computer	Federal Reserve Board
IP Index	Industrial Production: Material	Federal Reserve Board
IP Index	Industrial Production: consruction supplies	Federal Reserve Board
IP Index	Industrial Production: Mining :oil & gas extraction	Federal Reserve Board
IP Index	Industrial Production: Non durable consumer good	Federal Reserve Board
IP Index	Industrial Production: Durable manufacturing: Electrical equipment, appliance, and component	Federal Reserve Board
IP Index	Industrial Production: Durable manufacturing: Aerospace	Federal Reserve Board
IP Index	Industrial Production: Durable manufacturing:	Federal Reserve Board
IP Index	Industrial Production: Non Durable manufacturing	Federal Reserve Board
IP Index	Industrial Production: Business supplies	Federal Reserve Board
IP Index	Industrial Production: IPI excl. energy (74%)	Federal Reserve Board
IP Index	Industrial Production: Durable material	Federal Reserve Board
IP Index	Industrial Production: Non Durable material	Federal Reserve Board
IP Index	Industrial Production: Industrial equipment	Federal Reserve Board
IP Index	Industrial Production: manufacturing exluding vehicle	Federal Reserve Board
IP Index	Industrial Production: SA equipment total	Federal Reserve Board
IP Index	Industrial Production: electric & gas utilities	Federal Reserve Board
IP Index	Industrial Production: Total Index	FED Saint Louis
Labor statistics	Unemployed level, thousands	US BLS
Labor statistics	Employment level, thousands	US BLS
Labor statistics	US employment rate: Age 25 to 54	OECD
Labor statistics	Employment population ratio	US BLS
Labor statistics	All Employees: Total Nonfarm	US BLS
Labor statistics	US unemployment rate	US BLS
Labor statistics	Continued Claims (Insured Unemployment)	U.S. ETA
Leading Indicator	Chicago Fed National Activity Index	FED Saint Louis
Leading Indicator	Future New Orders; Diffusion Index for FRB - Philadelphia District	FED Philadelphia
Leading Indicator	Orders: Manufacturing: Total orders: Value for the United States	OECD
Leading Indicator	Manufacturers' New Orders for All Manufacturing Industries	U.S. Census Bureau
Leading Indicator	Manufacturers' New Orders durable goods	U.S. Census Bureau
Leading Indicator	Advance Real Retail and Food Services Sales	FED Saint Louis
Leading Indicator	Advance Retail Sales: Retail (Excluding Food Services)	FED Saint Louis
Leading Indicator	Advance Retail Sales: Retail and Food Services, Total	FED Saint Louis
Leading Indicator	Advance Retail Sales: Building Materials, Garden Equipment and Supplies Dealers	FED Saint Louis
Leading Indicator	Advance Retail Sales: Clothing and Clothing Accessory Stores	FED Saint Louis
Leading Indicator	Advance Retail Sales: Food Services and Drinking Places	FED Saint Louis
Leading Indicator	Advance Retail Sales: Furniture and Home Furnishings Stores	FED Saint Louis
Leading Indicator	Advance Retail Sales: Retail Trade and Food Services, Excluding Motor Vehicle and Gasoline Station	FED Saint Louis
Leading Indicator	Advance Retail Sales: Gasoline Stations	FED Saint Louis
Leading Indicator	Advance Retail Sales: Electronics and Appliance Stores	FED Saint Louis
Leading Indicator	Advance Retail Sales: Auto and Other Motor Vehicle	FED Saint Louis
Leading Indicator	Advance Retail Sales: Nonstore Retailers	FED Saint Louis
Leading Indicator	Advance Retail Sales: Motor Vehicle and Parts Dealers	FED Saint Louis
Leading Indicator	Advance Retail Sales: Food and Beverage Store	FED Saint Louis
Leading Indicator	Advance Retail Sales: Sporting Goods, Hobby, Book, and Music Stores	FED Saint Louis
Leading Indicator	Advance Retail Sales: Health and Personal Care Stores	FED Saint Louis
Leading Indicator	Advance Retail Sales: Retail Trade and Food Services, Excluding Motor Vehicle and Gasoline Station	FED Saint Louis
Leading Indicator	Advance Retail Sales: Retail Trade and Food Services, Excluding Gasoline Stations	FED Saint Louis
Leading Indicator	Leading Indicators OECD: Component series: CS - Confidence indicator	OECD
Surveys	Business Surveys: Order Books: Level	OECD
Surveys	Business Surveys: Export Order Books or Demand	OECD
Surveys	Business Surveys: Confidence Indicators (OECD)	OECD
Surveys	Business Surveys: Capacity Utilization	OECD
Surveys	Business Surveys: Confidence Indicators (European Commission)	OECD
Surveys	Business Surveys: Orders Inflow	OECD
Surveys	Business Surveys: Production	OECD
Surveys	Consumer Opinion Surveys: Confidence Indicators	OECD

2.A.4 Estimated factor

Figure (2.A.3) Estimated Factor, Market DY and Lead IP growth (In-sample estimates, forecasting over a 12-month horizon)



Note: The Figure represents the estimated factor (in red) based on an in-sample forecasting exercise over a 12-month horizon, the Market DY (in purple) as well as the IP growth lead by 12 month. For visual purposes we represent here the opposite of the Market DY and we normalized the three variables.

3. Structural estimation of time-varying spillovers: An application to credit risk transmission

(with Lukas Boeckelmann¹)

Abstract

We propose a novel approach to quantify time-varying financial spillovers based on a structural version of the Diebold-Yilmaz framework. Key to our approach is a SVAR-GARCH that is statistically identified by heteroskedasticity, economically identified by maximum shock contribution and that allows for time-varying FEVDs. We analyze spillovers between Euro Area sovereign and bank CDS. The spillovers estimated are a good fit for known spillover events and give more reactive signals compared to alternative models. We find spillovers to explain 37% of the variation in our sample, amid strong variations of the spillovers over time and in the cross section.

Nous proposons une nouvelle approche pour quantifier les phénomènes de contagion entre marchés financiers sur la base d'une version structurelle du modèle de [Diebold and Yilmaz \(2009\)](#). Nous nous reposons essentiellement sur un modèle SVAR-GARCH identifié statistiquement par hétéroscédasticité, identifié économiquement par la contribution maximale des chocs et qui permet d'obtenir des décompositions des erreurs de prévision non-constantes dans le temps. Nous analysons la propagation des chocs de risque de crédit dans la Zone Euro entre CDS souverains et bancaires. Du point de vue de la méthodologie, nous trouvons que notre modèle permet de mieux identifier les chocs de crédit par rapport aux autres modèles de contagion de la littérature, et qu'il est par ailleurs plus réactif aux événements que les modèles basés sur des estimations par fenêtres roulantes. Du point de vue économique, nous trouvons que les phénomènes de contagion expliquent seulement 37% de la variance de nos variables, avec toutefois de fortes variations dans le temps.

Keywords: *CDS, spillover, sovereign debt, SVAR, identification by heteroskedasticity.*

JEL codes: C58, G01, G18, G21.

¹European Central Bank, email: Lukas.Boeckelmann@ecb.europa.eu

3.1. Introduction

Credit risk spillovers have been among the major challenges for financial stability in the Euro Area. Recent spillover episodes include the collapse of the investment bank Lehman Brothers in 2008, the emergence of a sovereign bank nexus in multiple Euro Area countries during the Euro area debt crisis and the Italian political turmoil in 2018. These examples highlight that spillovers can occur at multiple and interdependent dimensions: between banking systems internationally, between sovereigns internationally and between banking systems and sovereigns in the same country.

Assessing credit risk spillovers on financial markets is challenging. The first challenge concerns shock identification: to evaluate how a specific shock propagated from one market to another requires first to identify this shock. Yet, this task may cause significant difficulties as asset prices contemporaneously affect each other and thus co-move markedly. The second challenge concerns time variation: Spillover episodes tend to be short lived and to vary substantially over time.

In this paper we propose a framework to estimate credit risk spillovers that combines an attractive identification approach for a set of endogenous variables with time variation in the spillover estimates. The approach relies on a Structural Vector Autoregression with a GARCH error structure (SVAR-GARCH) that is identified by heteroskedasticity. On the SVAR estimates we apply the framework of [Diebold and Yilmaz \(2009\)](#) and measure spillovers by the off-diagonal elements of the time-varying Forecast Error Variance Decomposition (FEVD). The approach allows a timely monitoring of spillovers and up-to-date assessment of financial stability risks. We estimate the model on a sample of 16 banking sector and sovereign CDS series in the Eurozone (EZ), ranging between 2008 and 2019. We estimate international spillovers between banking sectors and between sovereigns and national spillovers between sovereigns and banks in one mutually consistent framework.

The seminal work by [Diebold and Yilmaz \(2009\)](#), as well as a large number of subsequent papers (for example [Alter and Beyer, 2014](#); [Claeys and Vařćek, 2014](#); [Demirer et al., 2018](#); [De Santis and Zimic, 2018](#)), propose to base spillover estimates on the off-diagonal entries the FEVDs of rolling window structural vector autoregressions. While the approach allows for the construction of mutual consistent spillovers, the literature faces the econometric challenge of identification ([De Santis and Zimic, 2018](#)). Earlier papers rely on short-run zero restrictions for the coefficients of the SVAR (for example [Diebold and Yilmaz, 2009](#)). However this assumption is unlikely to hold with financial data that reacts almost instantaneously to news (see [Alter and Beyer, 2014](#)). Later papers sidestep any structural identification by using reduced form shocks in the form of Generalized FEVD analysis (GFEVD, see [Pesaran and Shin, 1998](#)). Yet, reduced form shocks have no economic interpretation and cannot be used for quantifying causal relationships of the data ([Kilian and Lütkepohl, 2017](#)). Other standard identification approaches are not appealing either: sign restrictions ([Fry and Pagan, 2011](#)), for example, are not exploitable as we do not want to restrict the impacts of the shocks *a priori*. [De Santis and Zimic \(2018\)](#) and [De Santis and Zimic \(2019\)](#) propose attractive identification schemes using magnitude restrictions. However, as most of the literature, they rely on rolling window estimations in order to generate time variation in their spillover estimates. Such rolling window estimations come with a significant drawback: at each point in time they deliver average spillover effects over large time horizons where new spillover estimates are averaged out with outdated estimates. Spillover estimates therefore do not represent up-to-date information.

We propose a novel approach for estimating time-varying spillovers by exploiting a SVAR-GARCH model that is statistically identified by the heteroskedasticity in the data. Also [Normandin and Phaneuf \(2004\)](#); [Bouakez and Normandin \(2010\)](#); [Lütkepohl and Milunovich \(2016\)](#) and others take advantage of the conditional heteroskedasticity in a SVAR-GARCH to identify structural shocks. We show that beyond this property, the model is attractive as it yields time-varying FEVDs based on the conditional variances of estimated structural errors. To the best of our knowledge, we are the first to exploit the properties of the conditional variances in a SVAR-

GARCH model to construct time-varying spillover estimates in financial networks. Moreover, we show that it is feasible to achieve economic identification between structural shocks and financial market variables in a nontrivial one-to-one relationship, even in a system of 16 variables. We label shocks with a maximum contribution to the forecast error variance of a variable as a shock of precisely that variable (similar to [Grosse Steffen and Podstawski, 2016](#) and [Dungey et al., 2010](#)). Due to the GARCH component in our estimation, spillover estimates are up-to-date instead of being averaged out in a moving window (as in [Diebold and Yilmaz, 2009, 2012, 2014](#)).

We investigate the properties of the SVAR-GARCH model estimated on the sample of Euro Area banking sector and sovereign CDS series. We show that the identification of the SVAR-GARCH model yields shock estimates that fit known economic and market events, supporting the choice of economic identification by maximum shock contribution. We manage to match major shocks to credit risk to 117 news events, either for bank or for sovereign CDS. In a second exercise, we compare the spillovers implied by the SVAR-GARCH with estimates stemming from other identification strategies used in the literature. For a range of established spillover events, either based on the events we identify, the events identified by [Candelon et al. \(2011\)](#) or the events identified by [Alexandre et al. \(2016\)](#), we apply a horse race between the competing models. We find that for either event list, the SVAR-GARCH outperforms identification schemes used in [Fengler and Herwartz \(2018\)](#), [Diebold and Yilmaz \(2009\)](#) or [Diebold and Yilmaz \(2012\)](#).

Overall, we find credit risk in the Euro Area to be less integrated than suggested by estimates based on traditional Diebold-Yilmaz approaches. We estimate that, on average, credit risk spillovers explain about 37% of the total variation in our sample. Yet, we show that the importance of spillover fluctuates distinctively, peaking at 61% during the Great Financial Crisis. Spillovers differ also largely in the cross-section. For example, we find that during the European debt crisis, spillovers from periphery sovereigns increased markedly, affecting the strongest credit risk of other periphery sovereigns and banking sectors. We also find strong credit risk

spillovers from periphery banking sector shocks, for example at the beginning of 2013 when investor worries surfaced about the health of the Italian banking system amid high non performing loan ratios and excessive reliance on debt. In contrast, we find for the period of the Great Financial Crisis elevated spillovers from core Euro Area countries.

We investigate the economic propagation channels underlying our spillover estimates. We find *international* credit risk spillovers between sovereigns to be higher when the countries have stronger ties in trade and portfolio investments, in line with the business cycle network literature (Foerster et al., 2011). We find *international* credit risk spillovers between banking systems to be higher when they exhibit more similar portfolios; yet we find spillovers not to be significantly associated with bank cross-holdings (similar to the findings in Brunetti et al., 2019). Concerning the national sovereign-bank nexus, we find that (i) a lower capital ratio and higher debt to GDP ratio increase *domestic* bank to sovereign spillovers in both low and high debt countries; while (ii) reliance of the non bank sector on domestic bank funding is significantly associated with domestic bank to sovereign spillovers only in low debt countries. In turn, we find *domestic* sovereign to bank spillovers to be higher for countries with a stronger bank exposure to domestic government debt. Moreover, we find that in high debt countries *domestic* sovereign to bank spillover are stronger when the domestic banking sector shows higher non-performing loan ratios and disposes of a lower share of liquid assets to short term liabilities.

The rest of the paper is structured as follows: Section 3.2 discusses the related literature, Section 3.3 details the methodology, Section 3.4 introduces the data, Section 3.5 reports the results of the SVAR-GARCH and Section 3.6 concludes.

3.2. Estimating Spillovers in the Literature

Throughout this paper we define spillovers as the degree to which exogenous shocks to one CDS market drive the variation of CDS spreads in other markets, based on the off-diagonals of

forecast error variance decompositions. However, the definition of spillovers may differ in the literature. [De Santis and Zimic \(2018\)](#) characterize spillovers as the impulse response of one shock to another variable while they label estimates based on FEVDs as “connectedness”. Additionally, [Forbes and Rigobon \(2002\)](#), [Claeys and Vašíček \(2014\)](#) and [Dungey et al. \(2015\)](#) term contagion as significant changes in the propagation mechanism, not the propagation mechanism itself.

[Diebold and Yilmaz \(2009, 2012, 2014\)](#) propose in a set of papers a prominent approach to quantify time-varying spillovers on financial markets. The model is widely reused in the literature (e.g. [Claeys and Vašíček, 2014](#); [Alter and Beyer, 2014](#); [Adams et al., 2014](#); [Fengler and Gisler, 2015](#); [Diebold et al., 2018](#); [Hale and Lopez, 2018](#); [Greenwood-Nimmo et al., 2017, 2019](#)). The key challenge of the approach is the identification of shocks in the underlying SVARs.

Three different strains of the spillover-literature do offer attractive identification strategies. First, [De Santis and Zimic \(2018\)](#) and [De Santis and Zimic \(2019\)](#) apply a methodology close to ours. They gauge the spillovers between sovereign debt markets and between medium-term interest rates with a Diebold-Yilmaz approach based on a SVAR that is identified by “magnitude restrictions”. The approach relies on the assumption that a shock originating from one country impacts the strongest the financial market in that very same country. Second, [Ando et al. \(2018\)](#) add numerous exogenous variables to their vector autoregressions with the aim to purge their variables from common factors. Once this filtering is done, they obtain (quasi) orthogonal shocks. Finally, several papers focusing on financial spillovers ([Ehrmann et al., 2011](#); [Dungey et al., 2015](#); [Ehrmann and Fratzscher, 2017](#); [Fratzscher and Rieth, 2019](#)) apply the idea of [Rigobon \(2003\)](#) and rely on the identification by heteroskedasticity. The authors use the variations in the variance-covariance matrix of the reduced form shocks to identify the structural shocks.

The time variation in the first two strains of the literature comes from a rolling window estimation. These papers use relatively long window length in order to have a sufficient accuracy in their parameter estimates. With this feature, models lack in responsiveness as past observations

mitigate the effect of new ones. The third strain of the literature focuses on specific sub-periods (e.g. [Ehrmann and Fratzscher, 2017](#) or [Dungey et al., 2015](#)) and do not provide a continuous estimation of their spillover indices.

In contrast, a recent literature has exploited MGARCH models that are capable of generating up-to-date spillovers ([Fengler and Herwartz, 2018](#); [Strohsal and Weber, 2015](#)). However, these models lack attractive identification approaches for structural analysis.^{2 3}

3.3. Methodology

3.3.1 Measuring spillovers

We follow the key idea of [Diebold and Yilmaz \(2009, 2012, 2014\)](#) and base a set of mutual consistent spillover measures, from pairwise to system wise, on FEVDs. Table 3.3.1 depicts a FEVD which is amended with an additional bottom row that captures the off-diagonal column sums, an additional column on the right that captures the off-diagonal row sums and a bottom right element that captures the grand average of either off-diagonal column or row sums.

Table (3.3.1) Diebold-Yilmaz Spillover Table

	y_1	y_2	\dots	y_N	To Others
y_1	d_{11}^H	d_{12}^H	\dots	d_{1N}^H	$\sum_{j=1}^N d_{1j}^H, j \neq 1$
y_2	d_{21}^H	d_{22}^H	\dots	d_{2N}^H	$\sum_{j=1}^N d_{2j}^H, j \neq 2$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
y_N	d_{N1}^H	d_{N2}^H	\dots	d_{NN}^H	$\sum_{j=1}^N d_{Nj}^H, j \neq N$
From Others	$\sum_{\substack{i=1 \\ i \neq 1}}^N d_{i1}^H$	$\sum_{\substack{i=2 \\ i \neq 2}}^N d_{i2}^H$	\dots	$\sum_{\substack{i=3 \\ i \neq N}}^N d_{i3}^H$	$\frac{1}{N} \sum_{i,j=1}^N d_{ij}^H$

²For example, in [Fengler and Herwartz \(2018\)](#) the orthogonalisation is based on the square root of the variance-covariance matrix of the reduced form shocks. Thus it does not rely on economic intuition and therefore makes the interpretation of the structural shocks difficult.

³The same drawback applies to [Geraci and Gnabo \(2018\)](#) who estimate a time-varying VAR to evaluate interconnectedness between financial institutions, although on the basis of reduced-form coefficients.

The FEVD is populated by elements d_{ij}^H , which give the proportion of the H step forecast error variance of variable y_j that is driven by an orthogonal shock to y_i . Following [Diebold and Yilmaz \(2009, 2012, 2014\)](#) we define d_{ij}^H as a *pairwise directed spillover* from i to j :

$$S_{i \rightarrow j}^H = d_{ij}^H. \quad (3.1)$$

The pairwise spillovers allow to construct more aggregated spillover indices. For example, the off-diagonal column sums indicate to which degree the H step forecast error variation of variable y_j is driven by other variables in the system. [Diebold and Yilmaz \(2009, 2012, 2014\)](#) define therefore *inward spillovers* as:

$$S_{j \leftarrow \bullet}^H = \sum_{\substack{i=1 \\ i \neq j}}^N d_{ij}^H. \quad (3.2)$$

Vice versa, the off-diagonal row sums indicate to what degree variable y_j drives the variation of all other variables in the system. *Outward spillovers* are therefore defined as:

$$S_{j \rightarrow \bullet}^H = \sum_{\substack{i=1 \\ i \neq j}}^N d_{ji}^H. \quad (3.3)$$

Total spillovers in the system are finally defined as average of *inward* or *outward spillovers*.

$$S^H = \frac{1}{N} \sum_{\substack{i,j=1 \\ i \neq j}}^N d_{ij}^H. \quad (3.4)$$

As underlined above, [Diebold and Yilmaz \(2009, 2012, 2014\)](#) estimate time-varying FEVDs based on moving window estimations of vector autoregressions, and identify the SVARs with orthogonalization strategies that can be challenged. The remainder of the paper outlines an approach that allows for a structural estimation of VAR parameters as well as for time-varying FEVDs that do not rely on rolling window estimations.

3.3.2 Description of the Model

For the development of a structural version of the Diebold-Yilmaz index, we rely on a SVAR model with a GARCH error structure and an identification by heteroskedasticity, similar in spirit to [Normandin and Phaneuf \(2004\)](#). We choose the model for the following reasons: first, a GARCH error structure appears a natural choice given that first differences of CDS, alike many other financial variables, show clustering of volatility over time and are therefore well approximated by GARCH processes. Second, the model has the property of time-varying conditional volatility of the errors, given the GARCH structure of the model. This property is crucial for the identification of structural shocks ([Rigobon, 2003](#)). Third, still relying on this property, we can construct time-varying FEVDs. This last feature allows us to estimate the model over the whole period, thus enabling more responsiveness compared to a time-varying FEVD based on a rolling estimation.

SVAR identification through heteroskedasticity

We base the empirical model on a structural vector autoregression of order p , that allows our variables to be determined simultaneously.

$$\mathbf{B}_0 \mathbf{Y}_t = \boldsymbol{\gamma} + \mathbf{B}_1 \mathbf{Y}_{t-1} + \dots + \mathbf{B}_p \mathbf{Y}_{t-p} + \boldsymbol{\epsilon}_t \quad (3.5)$$

where \mathbf{Y}_t is a vector containing the endogenous variables of interest, typically sovereign and bank sector CDS time series. The matrices \mathbf{B}_i contain the contemporaneous and lagged effects of the endogenous variables. $\boldsymbol{\epsilon}_t$ denote structural errors with zero mean and an unconditional diagonal variance covariance matrix $\boldsymbol{\lambda}_\epsilon$. As the SVAR cannot be estimated directly, we first estimate a reduced form VAR:

$$\mathbf{Y}_t = \boldsymbol{\beta} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \boldsymbol{\mu}_t \quad (3.6)$$

where the reduced form shocks μ_t have zero mean and a non-diagonal variance covariance matrix Σ_μ . The structural errors ϵ_t are then defined through μ_t and the contemporaneous interaction matrix B_0 :

$$\epsilon_t = B_0 \mu_t \Leftrightarrow \mu_t = B_0^{-1} \epsilon_t \quad (3.7)$$

The well known VAR identification problem arises as we try to obtain estimates for the contemporaneous interaction matrix B_0 from the relationship $\Sigma_\mu = B_0^{-1} \lambda_\epsilon B_0^{-1'}$. Yet without further restrictions B_0 is not identified since Σ_μ provides only $\frac{N(N+1)}{2}$ equations for N^2 unknowns if we normalize $\lambda_\epsilon = I$.

The SVAR-GARCH model we are using relies on [Rigobon \(2003\)](#) identification scheme that exploits the general heteroskedasticity in financial data. Suppose that the variances (or conditional variances) of μ_t vary over time - implying that the structural error variance does too - while B_0 is constant.⁴ This feature implies that there is more than one volatility regime in the data, defined by a different reduced form variance-covariance matrix $\Sigma_\mu(m)$. If there are M different volatility regimes, then we have:

$$\Sigma_\mu(1) = B_0^{-1} B_0^{-1'}, \quad \Sigma_\mu(m) = B_0^{-1} \lambda_m B_0^{-1'}, m = 2, \dots, M \quad (3.8)$$

where λ_m are the diagonal matrices of the structural shocks (λ_1 is normalized to I). [Lanne and Saikkonen \(2007\)](#) show that B_0 is locally uniquely determined if $\forall (k, l) \in \{1, \dots, K\}^2, k \neq l$, there is an index $j \in \{2, \dots, M\}$ such that $\lambda_{jk} \neq \lambda_{jl}$, i.e. there is sufficient heterogeneity in the volatility changes.

SVAR-GARCH

Conditional heteroskedasticity can be modeled in different ways (see [Lütkepohl and Netsunajev, 2017a](#)). We rely on the methodology first proposed by [Normandin and Phaneuf \(2004\)](#) and assume that it is driven by GARCH processes. Similar models have been applied in [Bouakez and](#)

⁴In Annex 3.A.5 we relax this assumption.

Normandin (2010), Lütkepohl and Milunovich (2016) and Lütkepohl and Netšunajev (2017b).

We assume that the structural shocks are orthogonal and that their variances follow a univariate GARCH(1,1) process:

$$\epsilon_{k,t} = \sigma_{k,t|t-1} e_{k,t} \quad \text{where} \quad e_t \sim \text{i.i.d. } N(\mathbf{0}, \mathbf{I}_N) \quad \text{and} \quad (3.9)$$

$$\sigma_{k,t|t-1}^2 = (1 - \gamma_k - g_k) + \gamma_k (\epsilon_{k,t-1})^2 + g_k \sigma_{k,t-1|t-2}^2 \quad (3.10)$$

where $\gamma_k > 0$, $g_k \geq 0$, $\gamma_k + g_k < 1$, $1 \leq k \leq N$ so that the GARCH(1,1) processes are non-trivial.

Then, we can express the reduced form shocks as:

$$\boldsymbol{\mu}_t = \mathbf{B}_0^{-1} \boldsymbol{\lambda}_{t|t-1}^{\frac{1}{2}} \mathbf{e}_t \quad (3.11)$$

where:

$$\boldsymbol{\lambda}_{t|t-1} = \begin{bmatrix} \sigma_{1,t|t-1}^2 & & 0 \\ & \dots & \\ 0 & & \sigma_{N,t|t-1}^2 \end{bmatrix} \quad (3.12)$$

is a $(N \times N)$ diagonal matrix with the univariate GARCH processes on the diagonal. Therefore, the distribution of $\boldsymbol{\mu}_t$ conditional on past information has mean zero and a covariance matrix:

$$\boldsymbol{\Sigma}_{\boldsymbol{\mu},t|t-1} = \mathbf{B}_0^{-1} \boldsymbol{\lambda}_{t|t-1} \mathbf{B}_0^{-1'} \quad (3.13)$$

Rigobon (2003) shows that for full (local) statistical identification, 2 different volatility regimes is enough. With a SVAR-GARCH we have T (number of observations) different volatility "regimes". In this study, using daily CDS data between 2008 and 2019, this translates into more than 2800 regimes. We estimate the parameters of the SVAR-GARCH model by Maximum Likelihood as in Lütkepohl and Milunovich (2016).

Forecasts for FEVD

Estimates for time-varying conditional variance-covariance matrices allow us to construct FEVDs for each time period, i.e. for each day. Note that for the computation of FEVDs in each period t , one cannot take the actual estimated structural variances $\hat{\lambda}_{t|t-1}$. Instead, we need to compute, by definition of the FEVD, in-sample forecasts for the structural variances $\lambda_{t+h|t}^*$ conditional on the information set in t , as in [Fengler and Herwartz \(2018\)](#). Contrary to the approach in the latter, our matrix B_0 is constant over time, so that the only change between a classic SVAR-FEVD and our approach is the computation of future structural variances.

We have with Equation (3.10):

$$\sigma_{k,t+h|t+h-1}^2 = (1 - \gamma_k - g_k) + \gamma_k(\epsilon_{k,t+h-1})^2 + g_k\sigma_{k,t+h-1|t+h-2}^2 \quad (3.14)$$

Taking conditional expectation at time t , with $h \geq 2$:

$$E_t\sigma_{k,t+h|t+h-1}^2 = (1 - \gamma_k - g_k) + \gamma_k\sigma_{k,t+h-1|t}^2 + g_kE_t\sigma_{k,t+h-1|t+h-2}^2 \quad (3.15)$$

Using the law of iterated expectations, we get:

$$E_t\sigma_{k,t+h|t}^2 = (1 - \gamma_k - g_k) + \gamma_k\sigma_{k,t+h-1|t}^2 + g_kE_t\sigma_{k,t+h-1|t}^2 \quad (3.16)$$

That is:

$$\sigma_{k,t+h|t}^2 = (1 - \gamma_k - g_k) + (\gamma_k + g_k)\sigma_{k,t+h-1|t}^2 \quad (3.17)$$

We thus obtain $\lambda_{t+h|t}^*$ for each h as this matrix is diagonal and is only composed of the different $\sigma_{k,t+h|t}^2$. To build the FEVDs, we then first compute the MSPE. The Θ_i matrices come from the Moving Average (MA) representation of the SVAR as detailed in [Kilian and Lütkepohl \(2017\)](#):

$$\mathbf{Y}_{t+H} - \mathbf{Y}_{t+H|t} = \sum_{i=0}^{H-1} \Theta_i \boldsymbol{\epsilon}_{t+H-i} \quad (3.18)$$

With the structural variances estimated, we get:

$$\begin{aligned} MSPE_t(H) &= E_t(\mathbf{Y}_{t+H} - \mathbf{Y}_{t+H|t})(\mathbf{Y}_{t+H} - \mathbf{Y}_{t+H|t})' \\ &= \sum_{i=0}^{H-1} \Theta_i \boldsymbol{\lambda}_{t+H-i|t}^* \Theta_i' \end{aligned} \quad (3.19)$$

We can then evaluate the contribution of shock j to MSPE of y_{kt} with the usual MSPE-formula, the only difference with a classic SVAR is that variances of structural shocks are no longer normalized to 1. With $\theta_{kj,h}$ the k,j^{th} element of Θ_h :

$$MSPE_{j,t}^k(H) = \theta_{kj,0}^2 \sigma_{j,t+H|t}^2 + \dots + \theta_{kj,H-1}^2 \sigma_{j,t+1|t}^2 \quad (3.20)$$

With:

$$MSPE_t^k(H) = \sum_{j=1}^K MSPE_{j,t}^k(H) \quad (3.21)$$

We get:

$$FEVD_{j,t}^k(H) = \frac{MSPE_{j,t}^k(H)}{MSPE_t^k(H)} \quad (3.22)$$

Eventually the time-varying FEVDs enable to build the time-varying spillover indices, as explained in Section 3.3.1.

3.4. Data and filtering for common shocks

3.4.1 Data

We focus on credit risk of major EZ sovereigns and banks. We attempt to strike a balance between a sufficiently high coverage of important CDS markets and the limited number of variables our empirical approach allows. As a result, we limit the sample to 9 countries (Greece, Ireland, Italy, Portugal, Spain, Germany, France, Belgium and Netherlands). For each country we include two variables in the sample, sovereign credit risk and credit risk in the banking sector, except for Ireland and Greece where we lack banking credit risk series due to data constraints.⁵ This leaves us with 16 variables all together.

As standard in this literature (see [Greenwood-Nimmo et al., 2019](#)), we measure credit risk using CDS spreads on senior unsecured debt, modified-modified restructuring, mid spread and a maturity of 5 years.⁶ We retrieve CDS spreads for non-US sovereigns and US banks denominated in USD while CDS spreads for the US sovereign and European banks are denominated in EUR. Our sample covers daily data between January 2008 and March 2019, covering the GFC, European debt crisis and several sovereign and banking turbulence such as the Italian political turmoil of May 2018.⁷

In our main specification, we estimate the model on the full period. Because the identification approach relies on a B_0 matrix that is constant over time, this implies that we do not take into account potential regime changes in the transmission mechanism of credit risk spillovers. Annex

⁵See [Acharya et al. \(2014\)](#) and [Fratzcher and Rieth \(2019\)](#).

⁶We combine data from three sources: we use principally Thomson Reuters Datastream and extend the sample backwards using growth rates extracted from CDS series from CMA. In case of missing values in the resulting data set, we retrieve growth rates on CDS spreads from Bloomberg.

⁷One shortcoming of using CDS spreads is that they are influenced not only by the probability of default of the corresponding bonds, but also by liquidity factors ([Fabozzi et al., 2007](#)). Therefore, our spillover estimates between CDS spreads may not reflect only risk transfers, but also illiquidity spillovers (see [Cespa and Foucault, 2014](#)). However, CDS appear better suited for our analysis as they are easily available both for individual banks (contrary to bond spreads) and for sovereigns (contrary to equity returns). Additionally, sovereign CDS and bond spreads appear to send a similar signal, as the average correlation of the two variables for the same country in our sample amounts to 75%.

3.A.5 shows that the spillover estimates estimated on the full sample are robust to alternative model specifications.⁸

We construct country banking variables as an unweighted average of bank CDS from that country as in Greenwood-Nimmo et al. (2017). In the selection of banks, we follow Alter and Beyer (2014).⁹ In line with the rest of the literature we first-difference the CDS series. A detailed list of the considered banks as well as descriptive graphs of the CDS series can be found in Annex 3.A.2.

3.4.2 Filtering for common shocks

The literature agrees that global and regional variables may exert a common influence on credit spreads (Longstaff et al., 2011). Ignoring such common shocks that have a simultaneous effect on different variables in an econometric analysis may result in an overestimation of spillover patterns. We would falsely attribute common shocks to the propagation of idiosyncratic shocks. We therefore follow Alter and Beyer (2014) in including the following set of pan-European credit risk factors, including (i) the *Itraxx Europe index* (which comprises investment grade rated European entities, reflecting the overall credit performance of the European real economy), (ii) the *Itraxx Crossover index* (which comprises below investment grade rated European entities, reflecting the lower-end credit performance of the European real economy), and (iii) the spread between the *3-month EURIBOR* and the *3-month EONIA* swap (a proxy for funding liquidity conditions equivalent to the TED spread). Moreover, we control for the *Eurostoxx 50* (the European stock market index), the *VIX index* (as a proxy for investors' risk aversion) and US and UK sovereign and banking CDS series (to account for foreign shocks).

We account for common shocks in a two-step approach. First, we regress each CDS series individually on a vector of common factors and then we run the SVAR-GARCH model specified

⁸We relax the assumption of constant transmission mechanisms and estimate the model on different sub-periods, allowing for a certain degree of time variation in the B_0 matrix.

⁹We exclude those banks that defaulted over the observation period.

in Section 3.3.2 on the obtained residuals, as in [Dungey et al. \(2010\)](#).¹⁰ That is, in a first step, we filter first differences bank and sovereign CDS series by the following OLS regression:

$$\Delta z_{jt} = \alpha_j + \Delta \mathbf{X}'_t \boldsymbol{\beta}_j + y_{jt} \quad (3.23)$$

where Δz_{jt} represents the first difference of a CDS series j in the sample, α_j is a constant and $\Delta \mathbf{X}_t$ is a vector of common factors in first differences. y_{jt} contains the residuals of the regression and serves as input data for the SVAR-GARCH. Annex 3.A.5 reports robustness checks using a smaller set of exogenous variables.

3.5. Results

In this section we present the results for the SVAR-GARCH model outlined above. We estimate the model with 2 lags as indicated by the information criteria from a simple VAR estimated on the same dataset. Moreover, in line with [Diebold and Yilmaz \(2009, 2012, 2014\)](#), we choose a forecast horizon for the FEVD of 10 days. In Section 3.5.1 we present the results of our identification approach, that is the labeling of structural shocks, as well as comparisons of timeliness and of identification performances with traditional spillover models. In Section 3.5.2 we present the economic results of our application.

3.5.1 Econometric results

Statistical and economic identification

Statistical identification is achieved when the number of univariate GARCH components underlying the GARCH structure are larger or equal to $N-1$. That means that for full local identification we may have at most one series that is not well approximated by a GARCH process in

¹⁰This approach is similar to including a vector of exogenous variables directly into the SVAR. [Alter and Beyer \(2014\)](#) find similar results between the two approaches. In our case, a two-step approach is preferable as we found that including a vector of exogenous variables in the SVAR-GARCH significantly increases the time to convergence.

order to have sufficient heteroskedasticity in the structural shocks. We follow the identification test proposed by Lanne and Saikkonen (2007) and reject fewer than N-1 GARCH processes in our sample (see Annex 3.A.6).

However, full local identification implies only statistical identification up to sign changes and ordering. To make the orthogonal shocks economic meaningful we need to label them, ideally in such a way that each orthogonal shock corresponds to a different variable. In line with Grosse Steffen and Podstawski (2016), we label shocks with the maximum contribution to the forecast error variance of a variable as a shock from this particular variable (for example the German banking sector). Exact economic identification is obtained if for each CDS series there is only one structural shock with a maximum contribution to the forecast error variance of that specific CDS series. As we estimate one FEVD for each day we focus on average shock contributions over time. However the labelling would be exactly the same if we focus, at each point in time, on individual FEVDs. Table 3.5.1 reports a FEVD that is averaged over all time periods and for which shocks are labelled accordingly. It is clear from the diagonal of the table that each shock has a maximum contribution to a different CDS series, allowing a clear labelling of the orthogonal shocks.

Table (3.5.1) Forecast error variance decomposition, average over time (%)

	BE bk	FR bk	DE bk	IT bk	NL bk	ES bk	PT bk	DE	BE	FR	GR	NL	ES	IT	PT	IE
BE bk	81	1	3	0	2	0	0	2	1	1	0	1	0	0	0	1
FR bk	0	54	5	2	0	6	0	2	16	11	1	2	4	6	3	5
DE bk	3	9	78	3	2	3	1	4	11	7	2	4	11	8	10	10
IT bk	3	5	1	83	17	16	4	0	1	0	5	0	2	2	3	0
NL bk	1	2	0	5	61	1	1	3	3	3	3	7	19	4	9	4
ES bk	2	23	3	1	1	56	1	10	4	3	1	1	1	1	1	4
PT bk	2	1	0	1	0	1	84	0	0	1	2	0	1	1	2	1
DE	4	1	1	0	2	2	0	70	13	8	2	10	4	5	3	6
BE	0	2	0	0	1	1	1	1	34	0	1	3	1	2	2	5
FR	1	2	0	0	1	2	1	0	5	61	1	1	1	3	1	1
GR	0	0	0	0	1	0	0	1	0	0	76	0	1	0	0	1
NL	1	0	0	1	1	2	0	2	9	3	1	68	1	2	1	7
ES	1	0	6	2	8	8	1	1	0	1	0	1	50	7	15	1
IT	0	0	0	0	1	1	0	1	1	1	2	1	2	58	3	2
PT	0	0	0	0	0	0	1	0	1	0	2	1	2	1	40	0
IE	1	0	0	1	0	0	3	1	0	1	0	0	1	1	8	53

This table represents the average over time of the FEVDs obtained with the SVAR-GARCH. We can see that the originating shocks (in line) impact the most their own variables (in column).

Our economic identification approach receives further confirmation from the fact that the estimated structural time series of shocks ($\hat{\epsilon}_t$ in Equation (3.5)) correspond to a large number of historical events. In the spirit of [Antolín-Díaz and Rubio-Ramírez \(2018\)](#), we compare major shocks with historical economic and market events.¹¹ We define major shocks as those shocks that are higher than 6 times their own standard deviations. Of the 79 shocks that meet this criteria, we are able to match 62 events (covering 78% of major shocks).¹² On Figure 3.5.1 we present the time series of the estimated structural shocks (in black) along with the timing of the matched events (red vertical lines). Figure 3.5.1 shows also isolated events of spillovers that fall short of the threshold for major events. Again, we are able to match a large amount of such shocks to economic and financial events, extending the list of events to 117 items. The identified events are typically rating downgrades or political shocks (for sovereigns) or bank stress episodes (for banking sectors). Annex 3.A.7 reports the exhaustive list of events. This exercise suggests that our identification strategy based on major shock contribution is further supported by the event-analysis on structural shocks of Figure 3.5.1, something which is rarely performed in the SVAR literature.

Total Spillover Comparison

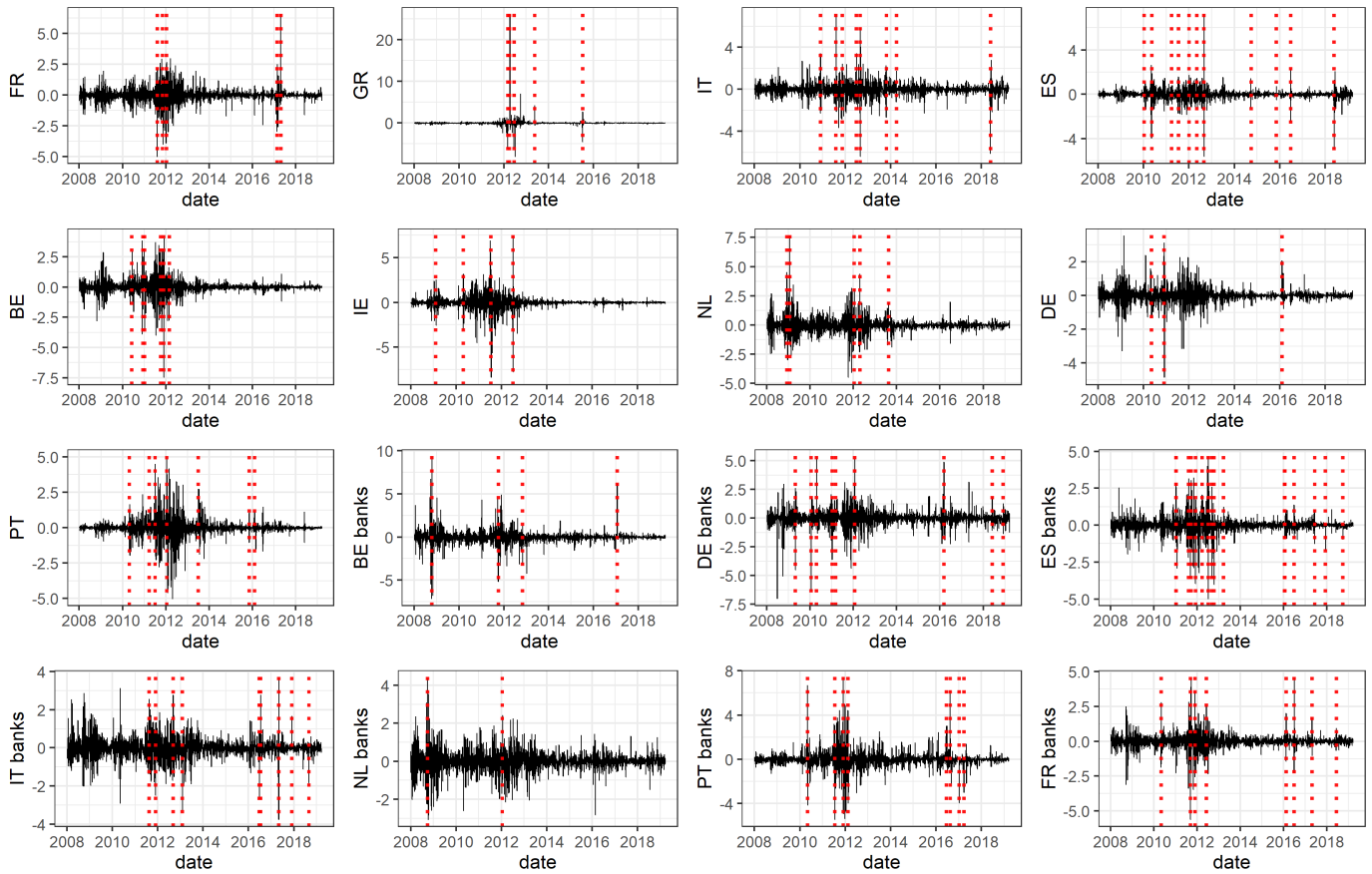
We now compare the total spillovers S^H defined in Equation (3.4) and estimated by the SVAR-GARCH model with total spillover estimates from alternative spillover models. The purpose of the exercise is to compare the magnitude and reactivity of spillover estimates. The models we compare measure spillovers through elements in FEVDs based either on SVARs or on multivariate GARCH models:¹³

¹¹More precisely, for days with a large structural shock surge, we investigate the existence of major market events in the financial press.

¹²We consider that a peak is identified if we can match it with an event 5 days before or after the date of the peak. This ratio is robust to changes in the threshold value as well to changes in the number of days considered.

¹³The [De Santis and Zimic \(2018\)](#) model is computationally intensive to estimate and therefore does not yield daily spillover estimates. For this reason we cannot integrate it in this comparison.

Figure (3.5.1) Structural Shocks and Events



On the different graphs above are represented the estimated structural shocks of the model ($\hat{\epsilon}_t$) as well as identified historical events for each variable represented in vertical red lines. The list of events used is available in Annex 3.A.7.

Model 1 A SVAR estimated on a rolling window and identified by Cholesky a decomposition (as in [Diebold and Yilmaz, 2009](#)) We label the model VAR Cholesky.

Model 2 A SVAR estimated on a rolling window and identified by GIRF/GFEVD (as in [Diebold and Yilmaz, 2012](#), labeled here VAR GIRF);

Model 3 A DCC-GARCH identified by a Cholesky decomposition and estimated over the full sample (in the spirit of [Elder and Serletis, 2010](#)). More precisely, we estimate a DCC GARCH

as a reduced form VAR, that is:

$$\mathbf{Y}_t = \boldsymbol{\beta} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \boldsymbol{\mu}_t$$

with:

$$\boldsymbol{\mu}_t \sim N(\mathbf{0}, \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t) \text{ and } \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t = \mathbf{H}_t$$

We then switch to the structural form with a Cholesky decomposition at each period t :

$$\mathbf{B}_{0t}^{-1} \mathbf{B}_{0t}^{-1'} = \mathbf{H}_t. \text{ We label the model DCC Cholesky.}$$

Model 4 Similarly to [Model 3](#), we estimate a VAR-GARCH based on a DCC-GARCH, but identify:

$$\mathbf{B}_{0t}^{-1} = \mathbf{H}_t^{1/2} \text{ (as in } \text{Fengler and Herwartz, 2018}).^{14} \text{ We label the model DCC Fengler.}$$

Figure [3.5.2](#) presents the model implied S^H . Several results are worth noting. First, the SVAR-GARCH evaluates credit risk in the EZ to be significantly less integrated than suggested by VAR Cholesky and VAR GIRF spillover estimates. This is not surprising, [De Santis and Zimic \(2018\)](#) show that GIRF identification tends to overestimate total spillovers and that Cholesky identification imposes restrictions that are likely to be at odds with the data generating process leading to spillover misspecification.¹⁵ Similarly, the identification by [Fengler and Herwartz \(2018\)](#) puts heavy restrictions on the Data Generating Process by assuming a symmetric structure between the variables. This may explain the lower spillover estimates by the DCC Fengler model compared to the SVAR-GARCH. We come back more formally on these points in [Annex 3.A.8](#).

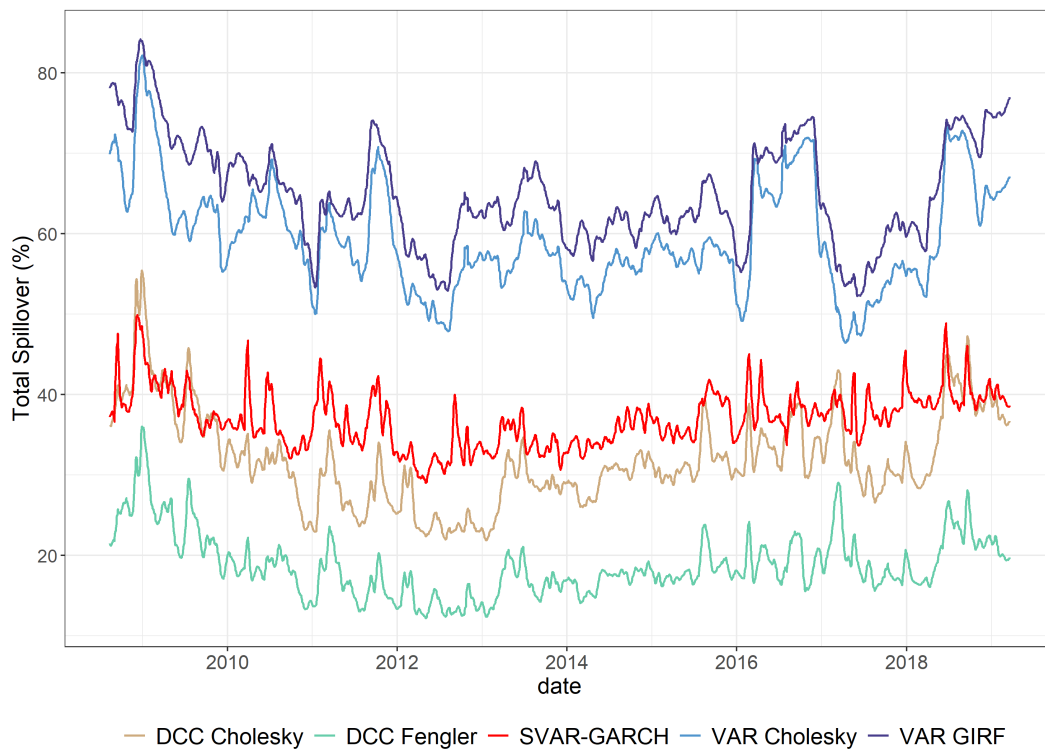
Second, we investigate whether spillover indices relying on rolling window estimations are less responsive to new events compared with spillover estimates that time-vary due to a GARCH

¹⁴Here, as in [Fengler and Herwartz \(2018\)](#), the square root of a symmetric positive definite matrix \mathbf{H} is defined as $\mathbf{H}_t^{1/2} = \boldsymbol{\Gamma} \boldsymbol{\Lambda}^{1/2} \boldsymbol{\Gamma}'$ where the columns of $\boldsymbol{\Gamma}$ contain the eigenvectors of \mathbf{H} and $\boldsymbol{\Lambda}^{1/2}$ is diagonal with the positive square roots of the eigenvalues on its diagonal.

¹⁵When contemporaneous interaction effects between variables are not equal to 0, the estimated standard errors of structural shocks obtained with GIRF are biased upwards, equally biasing upwards spillovers estimates based on FEVDs. The 0 restrictions the Cholesky identification introduces are likely to be at odds with the data generating process.

component. Intuitively, they ought to be while a priori we do not expect a clear distinction in reactivity between the different spillover models with GARCH components. We investigate these hypotheses using a Granger causality analysis between the different S^H . We report results in Table 3.A.2 in annex 3.A.3. The analysis shows that S^H estimated by SVAR-GARCH model Granger causes S^H computed by moving window approaches but not S^H computed by alternative GARCH models.

Figure (3.5.2) Total Spillover Indices from different Models



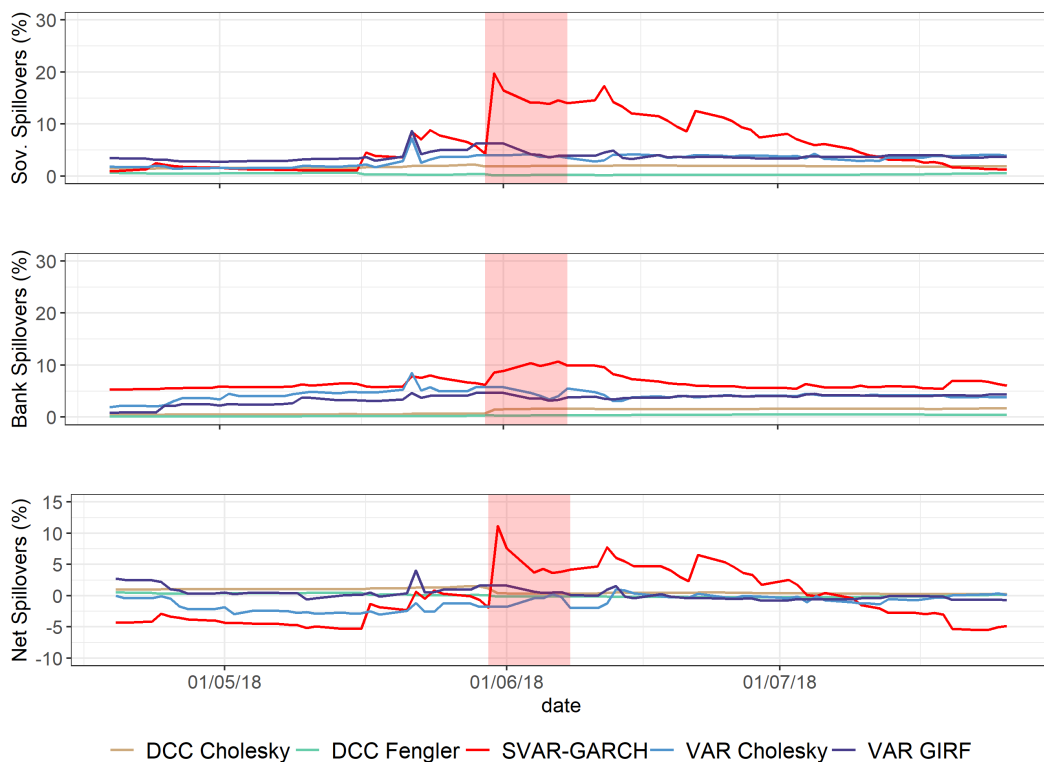
The different lines represent the Total Spillover indices S^H built from the five different models outlined above. The rolling window models are estimated on a 100-day period, as standard in this strand of literature. For readability we show 10 day moving averages of the indices.

Spillover Comparison and Narrative Events

To evaluate the relative performance of our identification strategy, we now compare how spillovers estimated by different approaches evolve along well-known narrative events.

To showcase our approach, we focus on the May 2018 political turmoil in Italy. At that time, the formation of a Eurosceptic coalition brought about a sharp increase in Italian sovereign CDS.¹⁶ We argue that this event should be interpreted as sovereign shock, not as a bank shock. Also, one would expect an increase in outward spillovers ($S_{j \rightarrow \bullet}^H$ of Equation (3.3)) from the Italian sovereign at the time of the event. Yet, the upper part of Figure 3.5.3 shows that only the spillover estimates from the SVAR-GARCH do so during this period (highlighted in red), while other spillover estimates remain subdued.

Figure (3.5.3) Outward Spillovers from Italian sovereign and bank shocks



The two upper parts of the graph represent the Outward Spillover Index ($S_{j \rightarrow \bullet}^H$) from, respectively, the Italian sovereign and the Italian banks, built from the five different models outlined above. The bottom part of the graph represents the “net” spillovers (outward sovereign spillovers minus outward bank spillovers). The periods highlighted in red represent the May 2018 Italian political turmoil.

¹⁶see <https://www.ft.com/content/eed97b90-6306-11e8-90c2-9563a0613e56>, or Reuters <https://www.reuters.com/article/us-eurozone-bonds/italian-bonds-suffer-worst-day-in-more-than-25-years-idUSKCN1IU16G>

As CDS spreads tend to comove substantially (Longstaff et al., 2011), and especially between sovereign and bank CDS series from the same country, there is a high risk that a model confuses bank shocks with their corresponding sovereign shocks. Expressed differently, at the time of a sovereign event, outward spillovers from the country's banking sector should remain flat or decrease. Taken together, for a sovereign event to be correctly identified, not only the sovereign spillovers should increase, they should also increase by more than the corresponding bank spillovers. On the middle and lower parts of Figure 3.5.3 we display the outward spillovers from Italian banks as well as the difference between sovereign and bank outward spillovers ("net" spillovers). While most of the models exhibit flat or negative net spillovers, only the SVAR-GARCH manages well to identify this specific event on this measure.

To evaluate on a more systematic basis the identification strategies of the different models, we replicate the analysis from Figure 3.5.3 over the set of our identified events available in Annex 3.A.7. We estimate that a sovereign (bank) event is well identified if, 5 days around the day of the event, the spillover estimate stemming from the sovereign (banking sector) increases more than the spillover estimate from the banking sector (sovereign) in the same country. We evaluate the identification performance of the models on different sets of events: (i) a subset of the least contestable sovereign events (i.e. only elections, sovereign rating downgrades, or political events) of the list identified in Section 3.5.1 and shown in Annex 3.A.7 covering 18 events, (ii) all sovereign events in Annex 3.A.7 covering 54 events and (iii) all events, bank and sovereign, in Annex 3.A.7 covering 117 events. As the sets of events (i), (ii) and (iii) are generated from our model, we corroborate the analysis with two exogenous lists of sovereign events, from (iv) Candelon et al. (2011) which covers 11 events and (v) Alexandre et al. (2016) which includes 8 events.

Table ?? suggests that the SVAR-GARCH outperforms or ties with, on every set of events, the other models in terms of identification. Note also that the competing models barely exceed the 50% threshold of identification, meaning that they tend to confuse more sovereign events with

Table (3.5.2) Percentage of good event-identification by model

	DCC Fengler	DCC Cholesky	VAR GIRF	VAR Cholesky	SVAR - GARCH
(i) Subset of sovereign events	11.0	22.0	44.0	39.0	78.0
(ii) Total sovereign events	30.8	33.3	33.3	38.5	64.1
(iii) Total sovereign and bank events	34.2	36.8	39.5	47.4	68.4
(iv) Candelon et al. (2011)	36.3	45.4	36.4	44.6	63.6
(v) Alexandre et al. (2016)	12.5	25.0	75.0	0.5	75.0

Note: This table reports the percentage of correct event identifications by each model. E.g. we consider a sovereign event to be correctly identified if, 5 days around the event, the outward spillover stemming from the sovereign increases more than the outward spillover stemming from the corresponding banking sector. The results are reported for (i) uncontroversial sovereign events (sovereign rating downgrades and votes) (ii) all the sovereign events previously identified (iii) all the sovereign and banking events identified (iv) the sovereign event list of [Candelon et al. \(2011\)](#) (v) the sovereign event list of [Alexandre et al. \(2016\)](#).

banking events than a random selection.¹⁷

3.5.2 Economic results

Figure 3.5.2 on total spillover indices shows that we estimate credit risk to be less integrated than other models would suggest. According to our S^H estimates, on average about 37% of the variation in the filtered CDS rates can be explained by spillovers. Yet, we find substantial variation in this magnitude over time. To investigate the sources of heightened spillovers, this section analyses first the time-variation of both bank and sovereign spillovers from the EZ countries, and then the economic channels behind the spillovers we estimate.

¹⁷Results reported in Table ?? are robust to a large number of specifications (by analysing the % change instead of absolute changes, with different window lengths or with pairwise spillovers instead of outward spillovers).

Group pairwise spillovers

In this section, we analyse credit risk spillovers in terms of (i) timing, (ii) magnitude and (iii) origin. Given, that we estimate spillovers between 16 CDS series, presenting the resulting 240 pairwise spillovers is not feasible. We focus therefore on pairwise spillovers from different sets of countries/banking sectors. In the “Peripheric” group are included the high-debt countries at the time of the EZ debt crisis: Italy, Spain, Portugal, Greece, Belgium and Ireland.¹⁸ The “Core” group, on the reverse, is constituted by Germany, France and the Netherlands. The “Peripheric banks” and “Core banks” include the corresponding banking sectors. However, as indicated in Section 3.4, due to data-constraints the group “Peripheric banks” does not include Greek and Irish banking sectors.

Figure 3.5.4 presents estimates of group pairwise spillovers for each variable sets defined above. In line with Section 3.3.1 we define the group pairwise spillover from group G_1 to group G_2 as the average outward spillovers from G_1 restricted to the variables of G_2 . More formally we have:

$$S_{G_1 \rightarrow G_2}^H = \frac{1}{N_{G_1} N_{G_2} - N_{G_1} \mathbb{1}_{\{G_1=G_2\}}} \sum_{i \in G_1} \sum_{\substack{j \in G_2 \\ j \neq i}} d_{ij}^H \quad (3.24)$$

With N_{G_1} and N_{G_2} the number of variables in G_1 and G_2 .¹⁹ Each line represents by how much shocks from a variable set drive the variation of other variable sets on average. The analysis of time-varying spillovers here differs from the presentation of snapshots spillovers around narrative events in Section 3.5.1 (as we focus here on a much broader time period) and also from the presentation of the shocks in Section 3.5.1. This is because spillover estimates in our DY-framework are not only functions of the time-varying variances of structural shocks

¹⁸Note that throughout his paper, we include Belgium in the Periphery-group as the country exhibited high public debt/GDP ratio. However the results are very similar if we define Belgium as a Core country.

¹⁹Contrary to Equations (3.2) and (3.3), we divide here the index by the number of pairwise directed spillovers considered. Likewise the different indices of Figure 3.5.4 are expressed in the same unit, that is: by how much, on average, a single variable of G_1 has an impact on a single different variable of G_2 .

(λ_t) , but also of the interaction matrix (B_0) and of the VAR coefficients (A_i). Therefore a large structural shock does not necessarily translate into a large spillover if it is associated with low coefficients in the corresponding matrices or if the magnitude of the shock is low relatively to other shocks' variances.

Figure 3.5.4 can be read in two ways, either from a *shock to* perspective by rows, or from a *shock from* perspective by columns. In the following, we take the *shock to* perspective. Figure 3.5.4 shows significant variation across time and groups. We find, in line with conventional wisdom, sizeable effects of sovereign periphery shocks to the rest of the EZ clustered around the beginning of the debt crisis in 2010. For example, at the height of the EZ debt crisis around mid-December 2010 when Moody's put Spain's rating on review, single variables from periphery sovereign shocks explained on average 4% and 3.5% of the variation of single variables from periphery sovereign and periphery banking groups respectively. Other major events of spillovers from periphery sovereigns include the Irish request for financial support to the EUs Financial Stability Facility and the IMF, the EU finance minister gathering to decide Greece's fate in 2015 or the 2018 Italian election crisis amid fears of new elections and voter support for Eurosceptics (see the subcaption of Figure 3.5.4 for exact dates). Figure 3.5.4 suggests that periphery sovereign shocks affect strongest CDS rates in other sovereign periphery countries followed by periphery banking sectors. Yet, core sovereigns and banks were also significantly affected by periphery sovereign shocks.

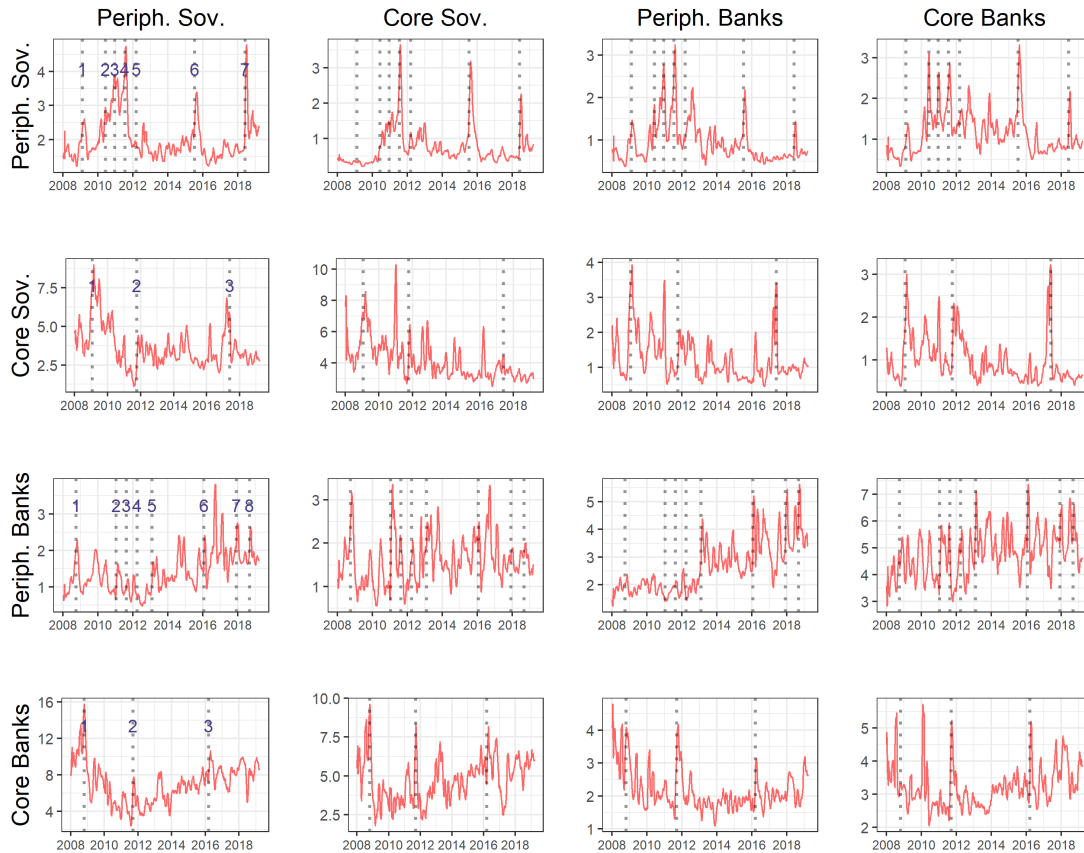
We also find sizable spillovers from the periphery banking sector to other blocks in the EZ. For example, Figure 3.5.4 shows elevated spillovers at the beginning of 2013, when investors worried about the health of the Italian banking sector (due to high NPL ratios amid excessive reliance on debt), as well as at the beginning of 2016, when again concerns about NPLs and the lack of credibility in the Italian banking sector heightened. We also find increased spillovers around dates between 2011 and mid-2012 when the Spanish banking sector signaled problems.

While we find spillovers from periphery EZ countries to increase with the beginning of the Euro

debt crisis, we find spillovers from core EZ countries to be stronger during 2008/09 financial crisis. For example, we estimate strong sovereign core spillovers in January 2009 when the Dutch government announced plans to provide a backup facility to cover the risks of the ING's securitised mortgage portfolio. Moreover, Figure 3.5.4 shows increased sovereign core spillovers around dates that coincide with a downgrade of France by S&P as well as the second round presidential election stand-off between Emmanuel Macron and Marine Le Pen. Finally, we find strong core bank spillovers, for example around the dates when ING received 10bn EUR from the Dutch government or when BNP entered a liquidity crunch when the bank was no longer able to borrow in USD.

Overall, compared to their periphery counterparts, we find sudden increases of spillovers from core countries to be less frequent. However, their size is much larger; spillovers from core eurozone countries appear to be an order of magnitude larger than those from periphery eurozone countries. This ought to be due to the fact that on average economies and banks tend to be larger in core eurozone countries compared to their periphery counterparts. We investigate in more detail the channels underlying the estimated spillovers in the next section.

Figure (3.5.4) Pairwise spillovers from EZ sovereigns and banks (%)



The different lines represent group pairwise spillovers ($S_{G_1 \rightarrow G_2}^H$) for the 4 groups: *Peripheral sovereigns, Core Sovereigns, Peripheral banks, Core banks*. For readability we show 30 day moving averages of the indices. **Peripheral Sovereigns:** 1) Ireland recapitalizes its two main banks 11/02/09, 2) Belgium struggles to raise debt among political uncertainty 07/06/10, 3) Moody's puts Spain's ratings on review 15/12/10, 4) Market pressure on Spanish and Italian sovereigns 17/07/11, 5) ISDA declares Greece in default 09/03/12, 6) EU finance minister gathering to decide Greece's fate 11/07/15, 7) Italy election crisis spreads as CB chief warns about investor trust 30/05/18, **Core Sovereigns:** 1) Dutch government announces plans to rescue banks 26/01/09, 2) SP mistakenly downgrades France 10/11/11, 3) French elections, spike in sovereign CDS 24/05/17, **Peripheral Banks:** 1) Dexia bailed out 30/09/08, 2) Concerns on

Spanish banks 14/01/11, 3) Trading suspension for Italian bank IS 17/08/11, 4) Need for Spanish bailout is underlined by EU officials 28/03/12, 5) MPS asks for 3.9bn bailout 01/02/13, 6) Market sentiment turns against Spain's banking sector 21/01/16, 7) ECB underlines Italian banks' NPL problems 29/11/17, 8) UniCredit and IS fall on news of increased political uncertainty 31/08/18 **Core Banks:** 1) ING receives 10 bn from Dutch government 19/10/08, 2) BNP can no longer borrow USD 13/09/11, 3) Deutsche and UBS defeated in UK tax avoidance case 10/03/16.

What economic channels explain spillovers?

While in the previous section we discussed the sources and time-variation of outward and total spillovers, this section focuses on the economic channels underlying the pairwise spillovers we estimate. More specifically, given a shock to a sovereign or banking sector in our sample, we vet whether the resulting pairwise spillovers match the economic channels proposed by the theoretic and empirical literature. We focus here on four different types of spillovers: (i) *international* sovereign to sovereign spillovers, (ii) *international* bank to bank spillovers, (iii) *national* bank to sovereign spillovers and finally (iv) *national* sovereign to bank spillovers.

International spillovers

First, we address the following question: given a sovereign shock in country i , what factors are the international spillovers to the sovereign risk in country j associated with? We follow broadly the regression approach by [De Santis and Zimic \(2018\)](#) and regress the credit risk spillover of sovereign i on sovereign j in quarter t on a set of regressors that can be divided into two main groups: distance and exposure. We estimate:

$$\bar{\omega}_{i \rightarrow j, t}(h) = \beta_i + \alpha_t + \beta_2 d_{i, j, t}^{GDP} + \beta_3 d_{i, j, t}^{\frac{D}{GDP}} + \beta_4 \text{exposure}_{j \rightarrow i, t}^k + \epsilon_{i, j, t} \quad (3.25)$$

where $\bar{\omega}_{i \rightarrow j, t}(h)$ is the average spillover from i to j at forecast horizon h over the quarter t , all

variables d are distance measures that include (i) the squared difference between country i and country j 's GDP growth in t and (ii) the squared difference between country i and country j 's government debt to GDP ratio in t .²⁰ Exposure $_{j \rightarrow i, t}^k$ is the exposure of country j to country i in respect of either the share of exports or portfolio assets (equity and bonds). The choice of those exposure variables follows the empirical work by [De Santis and Zimic \(2018\)](#) and the theoretical work by [Foerster et al. \(2011\)](#) (see Annex 3.A.4 for data sources and construction of the explanatory variables). We use time fixed effects and, following the regression design in [De Santis and Zimic \(2018\)](#), fixed effects for the origin of the sovereign shock.

Table (3.5.3) Factors associated with spillovers from sovereigns to sovereigns

	(1)	(2)	(3)
Similar BC	-0.00005 (0.001)	-0.002 (0.001)	-0.001 (0.001)
Similar D/GDP	0.021*** (0.002)	0.007*** (0.002)	0.014*** (0.002)
Trade exposure		0.433*** (0.020)	
Investment exposure			0.236*** (0.028)
Time fixed effects?	Yes	Yes	Yes
i fixed effects?	Yes	Yes	Yes
Observations	3,240	3,240	3,171
R ²	0.448	0.598	0.490
Adjusted R ²	0.438	0.591	0.481

*p<0.05; **p<0.01; ***p<0.001

The results, reported in Table 3.5.3 suggest that similarity in business cycles cannot explain spillovers in sovereign risk. Instead, we find that similar credit risk in terms of similar debt to GDP ratios as well as both stronger trade and portfolio exposure are significantly related to higher sovereign risk spillovers. This finding supports the business cycle network literature (such as [Foerster et al., 2011](#)) which models propagation channels through exactly those two

²⁰We multiply difference variables by -1, such that the indicators increase in similarity.

exposure variables.²¹

We repeat the same exercise to investigate the determinants of a spillover from the banking sector in country i to the banking sector in country j . Similar to Equation (3.25), we regress pairwise banking spillovers on a fixed effect for the shocking banking sector, as well as distance and exposure variables.

$$\bar{\omega}_{i \rightarrow j, t}(h) = \beta_i + \alpha_t + \beta_2 d_{i, t}^{NPL} + \beta_3 d_{i, t}^{Lev.R.} + \beta_4 \text{exposure}_{j \rightarrow i, t}^k + \epsilon_{i, t} \quad (3.26)$$

The distance variables include credit risk distances which we estimate by the squared difference between country i and country j 's banking sector's non-performing loans and capital ratios in period t .²² In terms of exposures we test for two economic channels that are frequently used to model financial institution linkages: cross asset holdings and similarities in portfolios across banking sectors (see Giudici et al., 2020, Brunetti et al., 2019, or Greenwood et al., 2015). We construct bank sector portfolios from BIS Consolidated Banking Statistics data and calculate squared differences of those portfolios for each time period t . Cross asset holdings between banking systems are measured as the share of banks claims of country j visàvis country i .

The results, shown in Table 3.5.4, suggest that cross-asset holdings are not significantly linked to the bank to bank spillovers. We do find however, that portfolio similarities are significantly associated with bank to bank spillovers. Both these findings are in line with the literature (Brunetti et al., 2019). Similarly to the sovereign regressions, risk distances have some explicative power: we find that international bank spillovers are significantly associated with similar capital ratios for pairs of banking systems. However, similar NPL ratios turn out not to be of statistical significance.

Spillovers in the national sovereign - bank nexus

²¹As the use of generated dependent variables in the regression can induce heteroskedasticity (see De Santis and Zimic, 2018), we report White heteroskedasticity-consistent standard errors.

²²Here again, we multiply difference variables by -1, such that the indicators increase in similarity.

Table (3.5.4) Factors associated with spillovers from banks to banks

	(1)	(2)	(3)
Similar NPLs	-0.001 (0.003)	-0.003 (0.003)	-0.001 (0.003)
Sim. Capital ratios	0.037*** (0.009)	0.029** (0.010)	0.040*** (0.010)
Similar portfolio		7.304*** (1.355)	
Bank claims			-0.013 (0.009)
Time fixed effects?	Yes	Yes	Yes
i fixed effects?	Yes	Yes	Yes
Observations	1,812	1,812	1,812
R ²	0.434	0.439	0.435
Adjusted R ²	0.417	0.422	0.417

*p<0.05; **p<0.01; ***p<0.001

While in the previous two regression sets we have focused on international spillovers, we investigate in the next two regressions the economic determinants of the national sovereign bank-nexus. In this section, we differentiate between high debt (Belgium, Italy, Portugal and Spain) and low debt (France, Germany and the Netherlands) countries as in the European debt crisis periphery and core countries experienced substantially different degrees of sovereign-bank nexuses (Podstawski and Velinov, 2018).

We focus first on the economic transmission channels of domestic spillovers from banks to sovereigns. One reason for higher spillovers may simply be a more vulnerable economy. We include in the regression measures of debt to GDP ratios, current account and GDP growth as predictor variables. Dell’Ariccia et al. (2018) identifies three additional propagation channels: First, bank risk may affect domestic sovereign risk through the “safety net channel”; explicit or implicit public guarantees that take effect in case a banking sector is in distress (Alter and Schüler, 2012). To proxy this effect, we add as a predictor the capital ratio of the banking sector. Intuitively, the “safety net channel” should be significant if domestic banks are under-

capitalized and in potential need for public support. The second channel is the “sovereign exposure channel”: when a banking sector is in distress it can trigger fire sales of the government bonds it holds, increasing in turn the credit risk of the sovereign issuer. The third channel is the “macroeconomic channel”: bank distress risks to decrease (domestic) lending activity and increase sovereign risk as economic growth slows (Podstawski and Velinov, 2018). We proxy the second and third channels, with two exposure variables in the regression: the share of domestic government bonds and the share of domestic non-bank assets that the banking sector holds. Denoting v_s^k the vulnerability variable k for sector s , Equation (3.27) restates the OLS regressions we estimate:

$$\begin{aligned} \bar{\omega}_{bank_i \rightarrow sov_i, t}(h) = & \beta_0 + \alpha_t + \beta_1 v_{bank_i, t}^{Lev.R.} + \beta_2 v_{sov_i, t}^{D/GDP} + \beta_3 v_{sov_i, t}^{CA} + \beta_4 v_{sov_i, t}^{gGDP} \\ & + \beta_5 \text{exposure}_{sov_i \rightarrow bank_i, t}^k + \epsilon_{bank_i, sov_i, t} \end{aligned} \quad (3.27)$$

For the vulnerability variables, we use dummies instead of continuous variables contrary to Equations 3.25 and 3.26.²³ We define high and low realisations of the variables with regards to their overall sample mean.²⁴ Since the sample is split according to debt levels, using the mean debt/GDP as threshold for the construction of a debt dummy does not yield much variation in the high debt subsamples. We therefore use the subsample mean for the high debt country group, and the overall sample mean for the low debt country group.²⁵

²³The underlying reason for using continuous variables for Equations 3.25 and 3.26 is that investors on the CDS markets may pass the shock of one sovereign (bank) to the price of another sovereign (bank) CDS if they judge them as similar. However for bank-to-sovereign or sovereign-to-bank regressions we cannot rely on such similarity metrics as the giving and the receiving variables are of different types. Therefore for Equations 3.27 and 3.28 we consider that investors pass the shock of a bank (sovereign) to a sovereign (bank) CDS if they judge the receiving variable as not resilient enough. This kind of reasoning is discrete, therefore we turn to dummy variable so as to illustrate the threshold that investors may consider.

²⁴Defined by 15.2% for the capital ratio, 0.5% for the capital account and 0.3% for real GDP growth.

²⁵We use the overall sample mean (86.5%) for the low debt group, and not the subsample mean (70%) as crossing the latter is unlikely to appear as warning signal for investors. Indeed, Germany has crossed this threshold between 2009 and 2016 while keeping its status as safe heaven. Using the subsample mean for the low debt groups renders the coefficients for the debt ratio insignificant while all other regression estimates are robust to the threshold change. The subsample mean is 101.7% for the high debt group.

We find in Table 3.5.5 that low capital ratios and high debt to GDP ratios are significantly associated with stronger domestic spillovers from banks to sovereigns. This suggests that the “safety net channel” may indeed be important in explaining the sovereign-bank nexus. We find that neither the capital account nor GDP growth is significantly associated with the spillovers we estimate. While the vulnerability variables yield similar results concerning the significance of the indicators across country groups, the results for the exposure variables differ. For high debt countries, both the dependence of the domestic non-bank corporate sector and government on domestic bank lending are not significant. In contrast, we find for low debt countries that higher non-bank exposure to domestic lending is significantly associated with higher domestic bank to sovereign spillover, suggesting that reduced lending activity in the case of a banking shock may indeed feed through the corporate sector into sovereign risk (see Pagano, 2018). As for high debt countries, we also find non-significant effects of sovereign debt exposure to domestic bank for low debt countries.

Finally, we investigate the determinants of domestic credit risk spillovers from a country’s sovereign to its banking sector (see Equation (3.28)). We test the following hypotheses: First, are domestic spillovers to banks stronger if the banking sector is more vulnerable? We proxy bank vulnerability with capital ratio, liquidity (measured by liquid assets to short term liabilities) and NPL ratios. Second, are spillovers stronger if the domestic banking sector holds more domestic government debt, expressed in % of total assets (the “sovereign exposure channel”, see Angeloni and Wolff, 2012 and Buch et al., 2016)? Third, are spillovers stronger if the domestic banking sector holds more assets of domestic non-financial firms, expressed in % of total assets (the “macroeconomic channel”)?²⁶ Here again, we express vulnerability variables in terms of dummies, where the thresholds between high and low realisations are set to sample averages.²⁷

We estimate:

²⁶The underlying rationale for this hypothesis is that a sovereign shock can feed into the real sector and then affect domestic banks, e.g. through increased taxes and lower consumer spending, or through a downgrade of non-financial companies. This last channel occurs because of the “rating channel”: companies cannot have a better rating than their own sovereigns, so when the sovereign is downgraded this also affects private companies, Arezki et al. (2010).

²⁷NPL ratios of 3.6%, 15.2% for capital ratios, 80.0% for liquid assets to short term liabilities.

Table (3.5.5) Factors associated with spillovers from banks to sovereigns in same country

	<i>High debt countries</i>			<i>Low debt countries</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Capital	-0.91*** (0.21)	-0.91*** (0.21)	-1.01** (0.31)	-5.01*** (1.04)	-5.07*** (1.05)	-3.88*** (1.07)
Debt to GDP	1.00*** (0.17)	1.01*** (0.17)	0.88*** (0.21)	3.74** (1.29)	3.56** (1.28)	3.50** (1.22)
Current Account	0.13 (0.20)	0.13 (0.26)	0.23 (0.28)	-0.63 (1.26)	-0.62 (1.28)	0.16 (1.31)
GDP growth	-0.06 (0.22)	-0.05 (0.23)	-0.05 (0.22)	-0.29 (0.66)	-0.31 (0.66)	-0.28 (0.68)
Sov. exposure		0.06 (1.92)			2.94 (9.63)	
Non-bank exposure			-0.88 (1.50)			20.52*** (6.02)
Time fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	174	174	174	123	123	123
R ²	0.74	0.74	0.74	0.93	0.93	0.94
Adjusted R ²	0.64	0.64	0.64	0.89	0.89	0.90

* p<0.05; ** p<0.01; *** p<0.001

$$\begin{aligned} \bar{\omega}_{sov_i \rightarrow bank_i, t}(h) = & \beta_0 + \alpha_t + \beta_1 v_{bank_i, t}^{NPL} + \beta_2 v_{bank_i, t}^{Lev.R.} + \beta_3 v_{bank_i, t}^{Liq.R.} + \\ & + \beta_5 \text{exposure}_{bank_i \rightarrow sov_i, t}^k + \epsilon_{sov_i, bank_i, t} \end{aligned} \quad (3.28)$$

Table (3.5.6) Factors associated with spillovers from sovereigns to banks in the same country

	<i>High debt countries</i>			<i>Low debt countries</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
NPLs	2.61** (0.81)	2.24** (0.81)	2.74** (0.84)	-0.10 (0.18)	-0.12 (0.17)	-0.19 (0.17)
Capital	-1.38 (0.77)	-0.80 (0.76)	-2.49 (1.31)	-0.06 (0.23)	0.34 (0.23)	0.41 (0.31)
Liquid assets	-4.03*** (0.48)	-6.17*** (1.01)	-4.60*** (0.69)	-0.92*** (0.18)	-0.11 (0.29)	-0.43 (0.29)
Exposure domestic gov. debt		0.18** (0.07)			0.18*** (0.05)	
Exposure domestic NFCs			-0.04 (0.03)			0.05** (0.02)
Time fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	174	174	174	121	121	121
R ²	0.54	0.55	0.54	0.88	0.89	0.88
Adjusted R ²	0.37	0.38	0.37	0.80	0.82	0.81

*p<0.05; **p<0.01; ***p<0.001

The results in Table 3.5.6 suggest that, in line with the literature, the “sovereign exposure channel” plays a major role for both high and low debt countries, as underlined by the positive and significant coefficients associated with government bond exposures. The “macroeconomic channel” seems to matter only for low debt countries (positive and significant coefficient for NFC exposures). Concerning the role of bank vulnerability, we find mixed results across country groups: For high debt countries, both higher NPL and lower liquidity ratios are significantly associated with higher domestic sovereign to bank spillovers in contrast to the capital ratio, which we find not to be significantly linked to the latter. For low debt countries, we find both NPL and capital ratios not to be significantly linked to spillovers, while we find lower liquidity ratios to

be significantly associated with higher spillover only in one out of three regressions.

3.6. Conclusion

We propose a novel approach of the popular Diebold-Yilmaz framework by exploiting a SVAR-GARCH model that is statistically identified by the heteroskedasticity in the data. We show that this identification approach is attractive as it yields time-varying FEVDs based on the conditional variances of estimated structural errors. Moreover, we show that it is feasible to achieve economic identification between structural shocks and financial market variables in a nontrivial bijective relationship, even in a system of 16 variables. We show the advantages of this methodological contribution by comparing the results with other common identification approaches used in the time-varying spillover literature. Overall, the identification scheme is supported by the fact that the results outperform other models in terms of timeliness and narrative fit. Additionally, we show that the obtained pairwise spillovers match propagation channels suggested by the literature.

This study has some limitations that could be addressed in future search. First, our identification approach relies on a constant B_0 matrix over the full sample period.²⁸ In principal, this constraint can be relaxed by estimating the model on shorter subsamples, for example defined on dates for which the researcher expects a structural break in interdependencies. While in Annex 3.A.5 we allow for a single change in the B_0 matrix, we leave a more profound analysis of this avenue for future research. Second, by imposing fewer constraints than previous models, the SVAR-GARCH could be applied to investigate spillover patterns for time series that have been less considered in the literature, notably market liquidity data.

²⁸That implies that while varying shock sizes may generate time variation in spillovers, elasticities between the variables stay constant.

Bibliography

- Acharya, V., Drechsler, I., and Schnabl, P. (2014). A Pyrrhic Victory? Bank Bailouts and Sovereign Credit Risk. *Journal of Finance*, 69(6):2689–2739.
- Adams, Z., Füss, R., and Gropp, R. (2014). Spillover effects among financial institutions: A state-dependent sensitivity value-at-risk approach. *Journal of Financial and Quantitative Analysis*, 49(3):575–598.
- Alexandre, H., Guillemin, F., and Refait-Alexandre, C. (2016). Crise de la dette souveraine dans l'Union Européenne: Transparence des banques et spreads de CDS. *Revue Economique*, 67(5):1007–1035.
- Alter, A. and Beyer, A. (2014). The dynamics of spillover effects during the European sovereign debt turmoil. *Journal of Banking and Finance*, 42(1):134–153.
- Alter, A. and Schüler, Y. S. (2012). Credit spread interdependencies of European states and banks during the financial crisis. *Journal of Banking and Finance*, 36(12):3444–3468.
- Ando, T., Greenwood-Nimmo, M., and Shin, Y. (2018). Quantile Connectedness: Modelling Tail Behaviour in the Topology of Financial Networks. *SSRN Electronic Journal*.
- Angeloni, C. and Wolff, G. B. (2012). Are banks affected by their holdings of government debt? *Bruegel Working Paper*, 07(March).
- Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018). Narrative sign restrictions for SVARs. *American Economic Review*, 108(10):2802–2829.
- Arezki, R., Candelon, B., and Sy, A. N. R. (2010). Bad news spreads. *Finance and Development*, 47(4):36–37.
- Bouakez, H. and Normandin, M. (2010). Fluctuations in the foreign exchange market: How important are monetary policy shocks? *Journal of International Economics*, 81(1):139–153.

- Brunetti, C., Harris, J. H., Mankad, S., and Michailidis, G. (2019). Interconnectedness in the interbank market. *Journal of Financial Economics*, 133(2):520–538.
- Buch, C. M., Koetter, M., and Ohls, J. (2016). Banks and sovereign risk: A granular view. *Journal of Financial Stability*, 25:1–15.
- Candelon, B., Sy, A. N. R., and Arezki, R. (2011). Sovereign Rating News and Financial Markets Spillovers: Evidence From the European Debt Crisis. *IMF Working Papers*, 11(68):1.
- Caporin, M., Pelizzon, L., Ravazzolo, F., and Rigobon, R. (2018). Measuring sovereign contagion in Europe. *Journal of Financial Stability*, 34:150–181.
- Cespa, G. and Foucault, T. (2014). Illiquidity contagion and liquidity crashes. *Review of Financial Studies*, 27(6):1615–1660.
- Claeys, P. and Vašíček, B. (2014). Measuring bilateral spillover and testing contagion on sovereign bond markets in Europe. *Journal of Banking and Finance*, 46(1):151–165.
- De Santis, R. A. and Zimic, S. (2018). Spillovers among sovereign debt markets: Identification through absolute magnitude restrictions. *Journal of Applied Econometrics*, 33(5):727–747.
- De Santis, R. A. and Zimic, S. (2019). Interest rates and foreign spillovers. *ECB Working Paper Series*, No 2221.
- Dell’Ariccia, G., Ferreira, C., Jenkinson, N., Laeven, L., Martin, A., Minoiu, C., and Popov, A. (2018). Managing the Sovereign-Bank Nexus. *Departmental Papers / Policy Papers*, 18(16):1.
- Demirer, M., Diebold, F. X., Liu, L., and Yilmaz, K. (2018). Estimating global bank network connectedness. *Journal of Applied Econometrics*, 33(1):1–15.
- Diebold, F. X., Liu, L., and Yilmaz, K. (2018). Commodity Connectedness. In Mendoza, E., D., S., and E, P., editors, *Monetary Policy and Global Spillovers: Mechanism, Effects and Policy Measures - An Overview*. Bank of Chile Central Banking Series, Santiago.

- Diebold, F. X. and Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, 119(534):158–171.
- Diebold, F. X. and Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1):57–66.
- Diebold, F. X. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.
- Dungey, M., Milunovich, G., and Thorp, S. (2010). Unobservable shocks as carriers of contagion. *Journal of Banking and Finance*, 34(5):1008–1021.
- Dungey, M., Milunovich, G., Thorp, S., and Yang, M. (2015). Endogenous crisis dating and contagion using smooth transition structural GARCH. *Journal of Banking and Finance*, 58:71–79.
- Ehrmann, M. and Fratzscher, M. (2017). Euro Area Government Bonds: Integration and Fragmentation During the Sovereign Debt Crisis. *Journal of International Money and Finance*, 70:26–44.
- Ehrmann, M., Fratzscher, M., and Rigobon, R. (2011). Stocks, bonds, money markets and exchange rates: Measuring international financial transmission. *Journal of Applied Econometrics*, 26(6):948–974.
- Elder, J. and Serletis, A. (2010). Oil price uncertainty. *Journal of Money, Credit and Banking*, 42(6):1137–1159.
- Fabozzi, F. J., Cheng, X., and Chen, R. R. (2007). Exploring the components of credit risk in credit default swaps. *Finance Research Letters*, 4(1):10–18.
- Fengler, M. R. and Gisler, K. I. (2015). A variance spillover analysis without covariances: What do we miss? *Journal of International Money and Finance*, 51(1):174–195.

- Fengler, M. R. and Herwartz, H. (2018). Measuring Spot Variance Spillovers when (Co)variances are Time-varying The Case of Multivariate GARCH Models. *Oxford Bulletin of Economics and Statistics*, 80(1):135–159.
- Foerster, A. T., Sarte, P. D. G., and Watson, M. W. (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119(1):1–38.
- Forbes, K. J. and Rigobon, R. (2002). No Contagion, Only Interdependence: Measuring Stock Market Comovements. *The Journal of Finance*, 57(5):2223–2261.
- Fratzscher, M. and Rieth, M. (2019). Monetary Policy, Bank Bailouts and the Sovereign-Bank Risk Nexus in the Euro Area*. *Review of Finance*, 23(4):745–775.
- Fry, R. and Pagan, A. (2011). Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature*, 49(4):938–960.
- Geraci, M. V. and Gnabo, J. Y. (2018). Measuring Interconnectedness between Financial Institutions with Bayesian Time-Varying Vector Autoregressions. *Journal of Financial and Quantitative Analysis*, 53(3):1371–1390.
- Giudici, P., Sarlin, P., and Spelta, A. (2020). The Multivariate Nature of Systemic Risk: Direct and Common Exposures. *Journal of Banking and Finance*.
- Greenwood, R., Landier, A., and Thesmar, D. (2015). Vulnerable banks. *Journal of Financial Economics*, 115(3):471–485.
- Greenwood-Nimmo, M., Huang, J., and Nguyen, V. H. (2019). Financial sector bailouts, sovereign bailouts, and the transfer of credit risk. *Journal of Financial Markets*, 42:121–142.
- Greenwood-Nimmo, M., Nguyen, V. H., and Shin, Y. (2017). What’s Mine Is Yours: Sovereign Risk Transmission during the European Debt Crisis. *SSRN Electronic Journal*.
- Grosse Steffen, C. and Podstawski, M. (2016). Ambiguity and Time-Varying Risk Aversion in Sovereign Debt Markets. *DIW Berlin Discussion Paper No. 1602*.

- Hale, G. and Lopez, J. A. (2018). Monitoring Banking System Fragility with Big Data. *Forthcoming: Journal of International Economics*, pages 01–41.
- Kilian, L. and Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.
- Lanne, M. and Saikkonen, P. (2007). A Multivariate Generalized Orthogonal Factor GARCH Model. *Journal of Business & Economic Statistics*, 25(1):61–75.
- Longstaff, F. A., Pan, J., Pedersen, L. H., and Singleton, K. J. (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*, 3(2):75–103.
- Lütkepohl, H. and Milunovich, G. (2016). Testing for identification in SVAR-GARCH models. *Journal of Economic Dynamics and Control*, 73:241–258.
- Lütkepohl, H. and Netšunajev, A. (2017a). Structural vector autoregressions with heteroskedasticity: A review of different volatility models. *Econometrics and Statistics*, 1:2–18.
- Lütkepohl, H. and Netšunajev, A. (2017b). Structural vector autoregressions with heteroskedasticity: A review of different volatility models. *Econometrics and Statistics*, 1:2–18.
- Normandin, M. and Phaneuf, L. (2004). Monetary policy shocks:: Testing identification conditions under time-varying conditional volatility. *Journal of Monetary Economics*, 51(6):1217–1243.
- Pagano, M. (2018). The sovereign-bank nexus and the case for European safe bonds. In *Finance and Investment: The European Case*, pages 149–156.
- Pesaran, H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1):17–29.
- Podstawski, M. and Velinov, A. (2018). The state dependent impact of bank exposure on sovereign risk. *Journal of Banking and Finance*, 88:63–75.

Rigobon, R. (2003). Identification Through Heteroskedasticity. *Review of Economics and Statistics*, 85(4):777–792.

Strohsal, T. and Weber, E. (2015). Time-varying international stock market interaction and the identification of volatility signals. *Journal of Banking and Finance*, 56:28–36.

Appendix

3.A. Annex

3.A.1 Derivation of Forecast Error Variance Decomposition

For each univariate structural shock, its variance follows the GARCH process:

$$\sigma_{k,t|t-1}^2 = (1 - \gamma_k - g_k) + \gamma_k(\epsilon_{k,t-1})^2 + g_k\sigma_{k,t-1|t-2}^2 \quad (3.29)$$

Unlike Fengler and Herwartz (2018), our matrix B_0 is constant over time, so that the only change between a classic SVAR-FEVD and our approach is the computation of future structural variances. We have:

$$\sigma_{k,t+h|t+h-1}^2 = (1 - \gamma_k - g_k) + \gamma_k(\epsilon_{k,t+h-1})^2 + g_k\sigma_{k,t+h-1|t+h-2}^2 \quad (3.30)$$

Taking expectations conditional on t , with $h \geq 2$:

$$E_t\sigma_{k,t+h|t+h-1}^2 = (1 - \gamma_k - g_k) + \gamma_k\sigma_{k,t+h-1|t}^2 + g_kE_t\sigma_{k,t+h-1|t+h-2}^2 \quad (3.31)$$

With the law of iterative expectations:

$$E_t \sigma_{k,t+h|t}^2 = (1 - \gamma_k - g_k) + \gamma_k \sigma_{k,t+h-1|t}^2 + g_k E_t \sigma_{k,t+h-1|t}^2 \quad (3.32)$$

That is:

$$\sigma_{k,t+h|t}^2 = (1 - \gamma_k - g_k) + (\gamma_k + g_k) \sigma_{k,t+h-1|t}^2 \quad (3.33)$$

In vector-form:

$$\begin{bmatrix} \sigma_{1,t+h|t}^2 \\ \dots \\ \sigma_{N,t+h|t}^2 \end{bmatrix} = \begin{bmatrix} 1 - \gamma_1 - g_1 \\ \dots \\ 1 - \gamma_N - g_N \end{bmatrix} + \begin{bmatrix} \gamma_1 + g_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \gamma_N + g_N \end{bmatrix} \begin{bmatrix} \sigma_{1,t+h-1|t}^2 \\ \dots \\ \sigma_{N,t+h-1|t}^2 \end{bmatrix} \quad (3.34)$$

To build the FEVDs, we first compute the MSPE:

$$\mathbf{Y}_{t+H} - \mathbf{Y}_{t+H|t} = \sum_{i=0}^{H-1} \Theta_i \boldsymbol{\epsilon}_{t+H-i} \quad (3.35)$$

And with the use of the structural variances estimated in 3.3.2:

$$\begin{aligned} MSPE_t(H) &= E_t(\mathbf{Y}_{t+H} - \mathbf{Y}_{t+H|t})(\mathbf{Y}_{t+H} - \mathbf{Y}_{t+H|t})' \\ &= \sum_{i=0}^{H-1} \Theta_i \boldsymbol{\lambda}_{t+H-i|t}^* \Theta_i' \end{aligned} \quad (3.36)$$

Then we can evaluate the contribution of shock j to MSPE of y_{kt} with the usual MSPE-formula, the only difference with a classic SVAR is that variances of structural shocks are no longer normalized to 1. $\theta_{kj,h}$ the kj^{th} element of Θ_h :

$$MSPE_{j,t}^k(H) = \theta_{kj,0}^2 \sigma_{j,t+H|t}^2 + \dots + \theta_{kj,H-1}^2 \sigma_{j,t+1|t}^2 \quad (3.37)$$

With:

$$MSPE_t^k(H) = \sum_{j=1}^K MSPE_{j,t}^k(H) \quad (3.38)$$

We get:

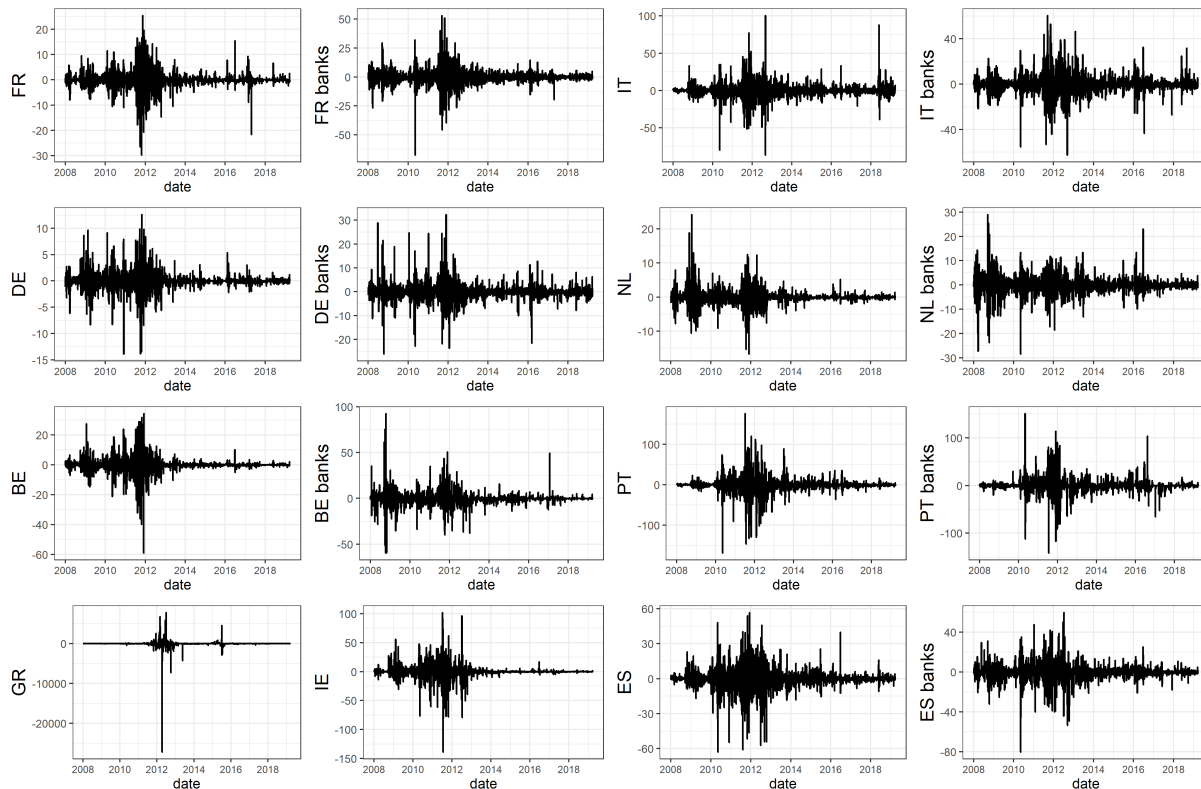
$$FEVD_{j,t}^k(H) = \frac{MSPE_{j,t}^k(H)}{MSPE_t^k(H)} \quad (3.39)$$

3.A.2 CDS Data

Table (3.A.1) List of banks used in bank sector CDS time series

Country	Banks
BE	Dexia, KBC Bank
FR	BNP, Société Générale, Crédit Agricole
DE	Deutsche Bank, Commerzbank, DZ Bank, Landesbank Baden, Landesbank Hessen, HSH Nordbank, WESTLB
ES	BBVA, Banco pastor, Santander, Sabadell, Banco Popolar Espagnol
NL	Rabobank, ING Bank, SNS Bank
IT	Intesa, Unicredit Spa, Banca Montepaschi, Banco PPO Italiana, Unione di Banche
PT	Banco Comercial Portugues, Banco BPI, Caixa Geral

Figure (3.A.1) CDS time series (basis points, first difference)



These graphs represent the raw time series of CDS used in the SVAR-GARCH (before the filtering of common shocks)

3.A.3 Granger causality test

S^H from the SVAR-GARCH does Granger cause S^H indices from the rolling window estimated models (VAR GIRF and VAR Cholesky), but not the indices stemming from a DCC-GARCH (DCC Cholesky and DCC Fengler). When we reverse the perspective, SVAR-GARCH is only Granger caused by DCC Fengler and not by VAR Cholesky or VAR GIRF. In that sense, S^H index estimated by DCC Fengler appears to be the most responsive to new events. However, as we show in the next section, the underlying pairwise spillovers estimated by DCC Fengler and DCC Cholesky are at odds with economic narratives. So that, contrary to the SVAR GARCH, these models fulfill the second condition of a good spillover model (responsiveness) but not the first one (good identification of the events).

Table (3.A.2) Granger causality between the models (First difference, lags=2)

<i>H</i> ₀ : SVAR-GARCH does not Granger cause...	F - test	p-value
<i>Rolling window estimated models</i>		
VAR Cholesky	14.91	3.627e-07***
VAR GIRF	6.0492	0.002391 **
<i>GARCH-related models</i>		
DCC Cholesky	1.0527	0.3491
DCC Fengler	1.3641	0.2558
<i>H</i> ₀ : SVAR-GARCH is not Granger caused by...	F - test	p-value
<i>Rolling window estimated models</i>		
VAR Cholesky	0.5159	0.597
VAR GIRF	0.9483	0.3875
<i>GARCH-related models</i>		
DCC Cholesky	0.4206	0.6567
DCC Fengler	8.8071	0.0001539***

This table indicates the results from the Granger causality tests between the S^H of the different models. Only three Granger causality relationships appear significant: SVAR-GARCH on VAR Cholesky and VAR GIRF, and DCC Fengler on SVAR-GARCH. The marks *, **, *** indicate, respectively, the following significance levels: 0.1, 0.05 and 0.01

3.A.4 Data sources OLS regressions

- Similar Business Cycle : the quarterly squared difference between country i and country j's GDP growth (multiplied by (-1) so that a higher number indicates more similar tendencies) [this is similar to [De Santis and Zimic \(2018\)](#), albeit [De Santis and Zimic \(2018\)](#) sum over over time as they focus on cross sectional effects]; Source: Eurostat
- Similar D/GDP : Same approach for quarterly D/GDP ratios (multiplied by (-1) so that a higher number indicates more similar tendencies); Source: IMF
- Trade exposure: $\frac{\text{Exports}_{j \rightarrow i}}{\text{Total exports}_j}$; Source: IMF
- Investment exposure: $\frac{\text{International investment}_{j \rightarrow i}}{\text{Total International investment}_j}$; Source: Eurostat

- Similar NPLs : the quarterly squared difference between banking sector i and banking sector j 's NPLs (multiplied by (-1) so that a higher number indicates more similar tendencies); Source: IMF Financial Soundness indicators and SNL
- Same approach for capital ratios; Source: IMF Financial Soundness indicators and SNL
- Similar portfolio: In a first step we construct from CBS data portfolio vectors per quarter for each banking sector as in Greenwood et al. (2015): a vector with the holdings of sovereign debt, non-bank financial institutions, households and non-financial corporations; for a large range of counterparty countries. We then express all portfolio items in % to total assets. And finally, we calculate the sum of the squared difference of those portfolio; Source: BIS CBS
- Bank claim exposure: $\frac{\text{Bank claims of country } j \text{ vis-a-vis country } i}{\sum_i \text{Bank claims of country } j \text{ vis-a-vis country } i}$, Source: BIS CBS
- NPLs, capital ratios (regulatory Capital to Risk-Weighted Assets) and liquidity ratios (Liquid Assets to Short Term Liabilities) of banking systems in percent; Source: IMF Financial Soundness indicators and SNL
- Exposure domestic government debt: $\frac{\text{Sovereign debt of country } i \text{ held by banking system } j}{\text{Total assets of banking system } j}$, Source: IMF IFS
- Exposure domestic government NFCs: $\frac{\text{Non-bank assets of country } i \text{ held by banking system } j}{\text{Total assets of banking system } j}$, Source: IMF IFS
- Sovereign exposure : $\frac{\text{Sovereign debt of country } j \text{ held by } i}{\text{Total sovereign debt of country } j}$, Source: BIS CBS
- Nonbank exposure : $\frac{\text{Liabilities of country } j \text{ non-banks held by } i}{\text{Total Liabilities of country } j \text{ non-banks}}$, Source: BIS CBS
- Debt to GDP, Current account, GDP growth in %, Source: Eurostat, OECD and IMF

3.A.5 Robustness checks

We perform several checks to assess the robustness of our results. First, with regards to the exogenous regressors: in our main specification we follow [Alter and Beyer \(2014\)](#) and include a significant number of exogenous variables to account for the strong comovement between CDS spreads. However, one might argue that some of them are endogenous to the sovereign and bank CDS. Therefore, we also estimate the model with a more parsimonious set of exogenous variables. In line with [De Santis and Zimic \(2018\)](#) we consider as alternative exogenous variables: oil prices, global macro news provided by Citibank, as well as US and UK CDS spreads. The upper part of Figure 3.A.2 represents total spillover indices S^H from different specifications, including our main one ("SVAR-GARCH"). As can be seen on the graph, some

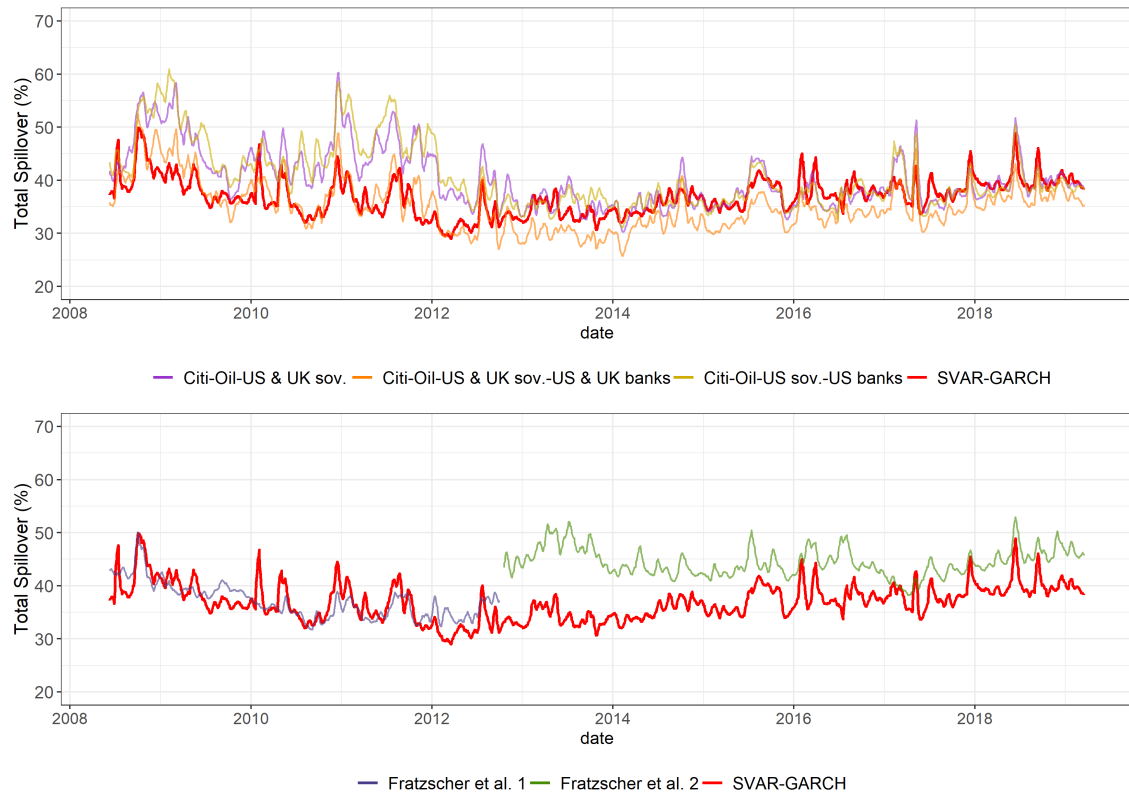
level changes are observable at the beginning of the estimation period, but overall the different indices evolve in a parallel manner.

Second, as exposed in Section 3.3.2, we assume a constant B_0 in our study. Some authors argued that the increase in CDS-correlation during the EZ debt crisis came mainly from changes in volatility and not in propagation mechanisms (Caporin et al., 2018), but this point is disputed (De Santis and Zimic, 2018). To evaluate the robustness of our results to changes in the estimation period, we estimate our model on different subsamples. In line with Ehrmann and Fratzscher (2017), we define two subperiods: a crisis period starting at the beginning of our sample until 01/10/2012, and a post-crisis period running from 01/10/2012 to the end of our sample.²⁹³⁰ The bottom part of Figure 3.A.2 exhibits total spillover indices S^H from the models estimated over the all sample, or over the two subsamples. Here again, apart from level differences on years 2013-2014, our results appear robust to changes in estimation period.

²⁹Note that Ehrmann and Fratzscher (2017) use 3 smaller subperiods, with a "crisis period" for the Greek turmoil that starts in March 2010 and ends in March 2012, and a "post-crisis period" that starts in October 2012. However, to build the total spillover indices S^H , we need the exact identification underlined in Section 3.5.1, i.e. being able to assign each shock to a single variable. This identification is not granted for any subsample, and is hard to achieve on short time intervals. Thus, on Figure 3.A.2 we rely on a large "crisis period" so as to obtain this identification.

³⁰Corresponding to the announcements of the implementation details of the ECB OMT-program.

Figure (3.A.2) Robustness graphs



The different lines represent the Total Spillover indices S^H built from our main specification (“SVAR-GARCH”) as well as from the other specifications outlined above. For the upper part of the graph, the different indices are named according to the exogenous variables included (Oil price, Macro news from Citibank, US and UK bank or sovereign CDS). For the bottom part of the graph, the different indices represent our main specification estimated on subsamples (before and after 01/10/2012 as outlined in [Ehrmann and Fratzscher \(2017\)](#)). For readability we show 10 day moving averages of the indices.

3.A.6 Test for identification and estimated coefficients

We rely on the original test proposed by [Lanne and Saikkonen \(2007\)](#) to test for the identification of B_0 . The recursive test applied here gives strong evidence for full identification of B_0 , see Table 3.A.3. For a more thorough description of the test, see [Lütkepohl and Milunovich \(2016\)](#). Note that the result of the test can be explained despite the reported low power of this latter, because (i) of the size of our dataset (ii) our 16 GARCH processes have a high persistence ($\gamma_k + g_k$ close to $0.9 \forall k$) which tends to increase the power of the test.

Table (3.A.3) Test for identification in SVAR-GARCH

h under H_0	$Q_1(1)$	df	p-value
1	124.3405	1	$< 10^{-5}$
2	113.4685	1	$< 10^{-5}$
3	85.0733	1	$< 10^{-5}$
4	66.6269	1	$< 10^{-5}$
5	60.7231	1	$< 10^{-5}$
6	46.2298	1	$< 10^{-5}$
7	38.0658	1	$< 10^{-5}$
8	35.8007	1	$< 10^{-5}$
9	25.3033	1	$< 10^{-5}$
10	16.2284	1	5.615e-05
11	13.3168	1	0.000263
12	12.6034	1	0.000385
13	517.7083	1	$< 10^{-5}$
14	185.0355	1	$< 10^{-5}$
15	154.8558	1	$< 10^{-5}$

3.A.7 List of Events

Table (3.A.4) Historical Events

Variable	Date	Events	Source
ES	12/01/2010	Spain rows back on measures to enforce economic co-operation	FT
ES	13/05/2010	Tough new austerity measures for Spain	FT
ES	15/12/2010	Moodys puts Spains Aa1 ratings on review for possible downgrade	FT
ES	29/03/2011	Catalan leader Arturo Mas refuses to enforce austerity measures	FT
ES	17/07/2011	Spain and Italy brace for bond market pressure	FT
ES	03/01/2012	Warning over size of Spanish deficit	FT
ES	08/05/2012	Spain set to spend billions on bank rescue	FT
ES	29/08/2012	Catalonia set to call for 5bn bailout	FT
ES	09/11/2015	Catalunya vote for independence	BBC
ES	27/10/2017	Catalan sparks Madrid showdown	FT
ES	27/05/2018	Spain upheaval deepens Italy market jitters	FT
BE	14/07/2008	Belgium government resigns	le Monde
BE	07/06/2010	Uncertainty on Belgium debt (elections coming)	FT
BE	14/12/2010	SP downgrades Belgium perspective	CNBC
BE	22/11/2011	Belgian spreads enlarge	FT
BE	25/11/2011	SP downgrades Belgium	FT
BE	01/03/2012	Belgian State buys back Dexia	Les Echos
IT	01/12/2010	"Premiums that Italy pay hit fresh highs"	FT
IT	30/06/2012	Markets rebound following EU deal "the agreement allow()to buy Italian sovereign bonds"	FT
IT	28/10/2013	Bond yields fall to four-month low as Italy sells 2-year debt	Nasdaq

Table (3.A.4) Historical Events

Variable	Date	Events	Source
IT	15/04/2014	"On April 15th yields on ten-year Italian-government bonds fell to 3.11%, the lowest on record"	The Economist
IT	30/05/2018	Italy election crisis spreads as central bank chief warns investor trust is fading	FT
IT	19/12/2018	Italian bonds and stocks rally as government comes closer to EU pact	FT
GR	09/03/2012	ISDA declares Greece in default (impact on CDS, restructuring)	Reuters
GR	09/04/2012	CDS decrease after Greece restructuring	The Economist
GR	19/06/2012	EZ's Greek poll honeymoon short lived	FT
GR	21/05/2013	Significant decrease in Greek sovereign CDS series	
GR	11/07/2015	EZ finance ministers prepare to decide Greece's fate	FT
PT	27/04/2010	Portugal rating downgraded	CNN
PT	29/03/2011	Portugal rating downgraded	FT
PT	06/07/2011	Portugal rating blow	FT
PT	18/01/2012	Moodys warns of second rescue for Portugal	FT
PT	07/02/2012	Speculation on Portugal debt restructuring	Les Echos
PT	09/03/2012	Renew speculation on Portugal debt	Reuters
PT	02/07/2013	Portuguese government at risk of collapse as foreign minister resigns	Telegraph
PT	10/11/2015	Confidence vote against government, potential left-wing coalition	Business Insider
PT	08/02/2016	Portugal-Germany Yield Spread Widens to Most Since 2014	Bloomberg
IE	11/02/2009	Recapitalisation was carried out at Ireland's two largest banks, Allied Irish Bank (AIB) and Bank of Ireland (BoI)	FT
IE	28/04/2010	Marked increase in Irish 2-year bond yields	The Irish Times
IE	18/07/2011	Record high of Irish CDS in our time series	
IE	06/07/2012	Ireland comes back on sovereign debt markets	FT
FR	11/08/2011	Focus of EZ crisis turns to France	FT
FR	10/11/2011	Standard & Poors mistakenly announced the downgrade of Frances top credit rating on Thursday	Reuters
FR	14/01/2012	SP downgrades France and Austria	FT
FR	22/02/2017	Highest DE-FR spread since 2012	CNBC
FR	28/04/2017	France CDS bounce back after election	IHS Markit
NL	09/10/2008	Government capital injections into banks	BIS
NL	26/01/2009	Bank comprehensive rescue plans (asset insurance)	BIS
NL	14/01/2012	SP puts Netherlands sovereign on negative outlook	Reuters
NL	23/04/2012	PM Rutte resigns after austerity talks	The Guardian
NL	20/08/2013	Netherlands' top rating is affirmed at Fitch amid debt warning	Bloomberg
DE	07/05/2010	German Parliament approves Greek rescue	NYT
DE	29/11/2010	German credit risk jumps to highest since may, debt swaps show	Bloomberg
DE	09/02/2016	Five-year sovereign German CDS rose to almost 22 bps due to hedging activity	Reuters
DE banks	28/04/2009	Profit-taking undermines Deutsche Bank	FT
DE banks	19/04/2010	Bank dividend payments reach record low: Deutsche Bank, plans to pay a dividend of 0.75 for 2009, up from 0.50 in 2008, but still small compared with earnings per share of 7.59	FT
DE banks	11/01/2011	Concerns rise over German bank levy	FT
DE banks	10/03/2011	Sale of stake in Deutsche AM puzzles analysts	FT
DE banks	28/07/2011	Deutsche Bank net revenues in its corporate and investment banking arm fell 27 per cent in the second quarter	FT
DE banks	10/09/2011	Commerzbank hit by 760m Greek writedown	FT
DE banks	24/11/2011	Deutsche bank needs 2 bln to meet EBA's conditions	EBA
DE banks	07/12/2011	SP placed the credit Deutsche Bank and Commerzban under review	Deutsche Welle
DE banks	24/01/2012	Commerzbank buoyant as investors back capital plan	FT
DE banks	10/03/2016	Deutsche and UBS defeated in UK tax avoidance case over bankers bonuses	BBC
DE banks	03/05/2017	HNA raises stake in Deutsche Bank to nearly 10%	FT
DE banks	01/06/2018	SP downgrades Deutsche Bank	FT
DE banks	04/12/2018	Investor fear raids will hit DB turnaround	FT

Table (3.A.4) Historical Events

Variable	Date	Events	Source
FR banks	06/05/2010	BNP Paribas and Société Générale in suffering as the costs of insuring themselves against default rises	FT
FR banks	12/08/2011	French Short-selling ban brings relief for banks	FT
FR banks	13/09/2011	BNP Bank executive says they can no longer borrow USD	WSJ
FR banks	14/10/2011	Big European CDS such as Frances BNP Paribas spiked to 291bp	FT
FR banks	07/11/2011	BNP stock price plunges compared to CAC 40p	Les Echos
FR banks	30/11/2011	SocGen, UniCredit and BNP lose some of Mondays gains	FT
FR banks	12/02/2016	SocGen battles to hit targets amid low rates and volatility	FT
FR banks	11/06/2018	France tells its banks to set aside more capital	FT
IT banks	17/08/2011	Shares in Italys biggest retail bank Intesa Sanpaolo were at one point suspended for excessive losses "European banks at centre of sell-off "	FT
IT banks	30/11/2011	European banks junior debt under review, including a number of Italian banks "European banks junior debt under review"	FT
IT banks	01/02/2013	Monte dei Paschi di Siena asks for 3.9bn bailout amid scandal over loss-making derivatives contracts and alleged fraud	the Guardian
IT banks	11/06/2016	Italian banking crisis, heightened by European financial stress tests	FT
IT banks	27/06/2016	Italian banks struggling "Italy resurrects plans to rescue struggling banks"	FT
IT banks	04/05/2017	Monte Paschi CDS time series spike	
IT banks	29/11/2017	European Central Bank has redoubled warnings that the state of EZ banks is a threat to the regions economic recovery, IT banks with biggest problem of sour loans	FT
IT banks	31/08/2018	UniCredit and Intesa Sanpaolo fall on news of increased political uncertainty	FT
PT banks	06/05/2010	Spectre of counterparty risk, focused attention on to smaller banks in Portugal and Spain	FT
PT banks	19/07/2011	BCP fails Espírito Santo Financia almost fails EBA stress test	EBA
PT banks	30/11/2011	BCP's CDS arrive at record level after Fitch downgrade of covered bonds	Bloomberg
PT banks	16/02/2012	Moodys downgrades state guaranteed debt issued by BCP from Ba2 to Ba3 with negative outlook	FT
PT banks	18/08/2016	Portuguese bonds under pressure after rating agencies warning	FT
PT banks	10/01/2017	Fosun to increase its stake in Millennium BCP to 30%	FT
PT banks	28/03/2017	CDS spread of BCP drops sharply	
BE banks	18/09/2008	Speculative rumors against Fortis	La Libre Belgique
BE banks	30/09/2008	Dexia bailed out	The Guardian
BE banks	15/10/2008	Bank rally, Dexia, the Franco-Belgian bank whose borrowings have already had to be guaranteed, dismissed rumours that it faced impending nationalisation by the Belgian government	FT
BE banks	29/12/2008	KBC loses 1 billion on CDOs	La Libre Belgique
BE banks	30/09/2011	Belgium market authority ends short selling ban on Belgian financial institutions	Fed NY
BE banks	05/10/2011	Dexia shares suspended as break-up takes shape	FT
BE banks	02/11/2012	Three banks Lloyds Banking Group, Commerzbank and Dexia were dropped from the GSifi list	FT
BE banks	29/12/2012	European commission validates Dexia rescue plan	Le Monde
BE banks	25/01/2017	Repricing of Dexia CDS	FT
ES banks	14/01/2011	Spain seeks to show that it is not another Ireland	FT
ES banks	13/07/2011	Spanish bank IPOs under threat	FT
ES banks	09/08/2011	Investors turn agains spanish financials as they bet against the value of what they see as fragile institutions	FT
ES banks	26/09/2011	News about bank rescue plan: ECB expected to boost bank liquidity. Spanish banks affected in particular	FT
ES banks	28/03/2012	EU underlines that Spanish banks need to bailout	FT
ES banks	09/07/2012	Spanish bank bailout talks. Spain to Accept Rescue From Europe for Its Ailing Banks	NYT
ES banks	16/10/2012	Spanish banks rally on hope Madrid ready to request aid	FT
ES banks	25/03/2013	Bankia leads falls across big lenders after EZ comment to toughen bank regime	FT
ES banks	21/01/2016	Market sentiment turns sharply against Spains banking sector	FT
ES banks	08/06/2017	Emergency funds failed to save Banco Popular from death spiral	FT
ES banks	05/12/2017	Strong drop of CDS of Banco Popular	

Table (3.A.4) Historical Events

Variable	Date	Events	Source
ES banks	26/09/2018	Banco Santander changes its chief executive	FT
NL banks	30/09/2008	The Belgian-Dutch Fortis faces state rescue	Reuters
NL banks	19/10/2008	ING receives 10 billion from Dutch government	NYT
NL banks	26/03/2009	Fortis Bank Nederland posts 25.11 billion loss	Reuters
NL banks	16/01/2012	ING benefits from ING gains from Netherlands credit rating	FT
NL banks	13/06/2012	ING to pay USD 619m to settle sanctions case	FT
NL banks	09/07/2012	Former Rabobank traders fired in LIBOR-scandale	Observer

3.A.8 IRF assumptions

An additional advantage of our econometric framework is that it imposes less restrictions on the impulse response functions (IRFs) compared to other spillover models. More specifically, Cholesky-identified SVARs, as used in VAR Cholesky and DCC Cholesky, postulate a recursive structure of the IRFs. Generalized impulse response functions, as used in VAR GIRF, impose that the IRF of a one-unit shock i on variable j has the same initial impact as an IRF from shock j to variable i . Eventually the same criticism applies also to the orthogonalization in [Fengler and Herwartz \(2018\)](#) as used for the model DCC Fengler -see demonstrations below. On the reverse, the SVAR-GARCH framework does not impose such a structure ([Lütkepohl and Netšunajev, 2017a](#)). Therefore the level-differences observed on [Figure 3.5.2](#) may come from the overly strong assumptions of the competing models, over- or underestimating the spillovers.

Why do the identifications used in VAR GIRF and in DCC Fengler assume symmetrical IRFs? A usual VAR-analysis begins with the reduced form VAR (Equation (3.6)), with the objective to go to the structural form of Equation (3.5). Under covariance stationarity, Equation (3.6) is equivalent to its moving average representation:

$$Y_t = \sum_{i=0}^{\infty} \Phi_i \mu_{t-i} \quad (3.40)$$

Which can be rewritten in its structural form:

$$\mathbf{Y}_t = \sum_{i=0}^{\infty} (\Phi_i \mathbf{B}_0) (\mathbf{B}_0^{-1} \boldsymbol{\mu}_{t-i}) = \sum_{i=0}^{\infty} \Theta_i \boldsymbol{\epsilon}_{t-i} \quad (3.41)$$

Generally speaking, the IRF of a vector shock $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$ on \mathbf{Y}_t is defined, at horizon h and with $\boldsymbol{\Omega}_{t-1}$ the information set at t , as:

$$IRF(h, \boldsymbol{\delta}, \boldsymbol{\Omega}_{t-1}) = E(\mathbf{Y}_{t+h} | \boldsymbol{\epsilon}_t = \boldsymbol{\delta}, \boldsymbol{\Omega}_{t-1}) - E(\mathbf{Y}_{t+h} | \boldsymbol{\Omega}_{t-1}) \quad (3.42)$$

Due to the orthogonality of the structural shocks, one uses $\boldsymbol{\delta} = (0, \dots, 0, \delta_j, 0, \dots, 0)$ in order to consider the impact of a single shock. In that case we get, with \mathbf{e}_j a vertical vector full of zeros apart for its j^{th} element that is equal to 1: $IRF(h, \boldsymbol{\delta}, \boldsymbol{\Omega}_{t-1}) = \Phi_h \mathbf{B}_0 \boldsymbol{\delta} = \Phi_h \mathbf{B}_0 \mathbf{e}_j \delta_j$

In our SVAR-GARCH setting, \mathbf{B}_0 is identified by heteroskedasticity and by economic identification (with $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_\mu$ evolving over time and being equal to, respectively, $\boldsymbol{\lambda}_{t|t-1}$ and $\boldsymbol{\Sigma}_{\mu,t|t-1}$, Equation (3.13)). This identification strategy does not impose any structure on the IRFs. Conversely, in [Diebold and Yilmaz \(2009\)](#), \mathbf{B}_0 is identified by using the Cholesky decomposition of $\boldsymbol{\Sigma}_\mu$ with $\mathbf{B}_0^{-1} \mathbf{B}_0^{-1'} = \boldsymbol{\Sigma}_\mu$ as in the first equation of Equation (3.8). Although convenient, this orthogonalization imposes a recursive structure in the Data Generating Process as \mathbf{B}_0 is then lower-triangular.

Identification by GIRF works differently since, instead of considering structural shocks, the GIRF looks at reduced form shocks. Using the notation $\boldsymbol{\Sigma}_\mu = (\sigma_{ij})_{i,j \in \llbracket 1, n \rrbracket^2}$, a one standard deviation shock j and the same remaining notations, the GIRF is defined as:

$$GIRF(h, \sigma_{jj} \mathbf{e}_j, \boldsymbol{\Omega}_{t-1}) = E(\mathbf{Y}_{t+h} | \boldsymbol{\mu}_t = \sigma_{jj} \mathbf{e}_j, \boldsymbol{\Omega}_{t-1}) - E(\mathbf{Y}_{t+h} | \boldsymbol{\Omega}_{t-1}) \quad (3.43)$$

If one assumes that $\boldsymbol{\mu}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\mu)$, then we can write (see [Pesaran and Shin, 1998](#)):

$$E(\boldsymbol{\mu}_t | \mu_{jt} = \sigma_{jj}) = [(\sigma_{1j}, \dots, \sigma_{mj})' \sigma_{jj}^{-1}] \sigma_{jj} = \boldsymbol{\Sigma}_\mu \mathbf{e}_j \quad (3.44)$$

So that the impact of a one standard deviation j shock on variable i at horizon 0 is (with Equation (3.40)):

$$GIRF_i(0, \sigma_{jj} \mathbf{e}_j, \boldsymbol{\Omega}_{t-1}) = \mathbf{e}_i' \boldsymbol{\Phi}_0 \boldsymbol{\Sigma}_\mu \mathbf{e}_j \quad (3.45)$$

As $\boldsymbol{\Phi}_0 = \mathbf{I}$ we get:

$$GIRF_i(0, \sigma_{jj} \mathbf{e}_j, \boldsymbol{\Omega}_{t-1}) = \mathbf{e}_i' \boldsymbol{\Sigma}_\mu \mathbf{e}_j = \sigma_{ij} = \sigma_{ji} = GIRF_j(0, \sigma_{ii} \mathbf{e}_i, \boldsymbol{\Omega}_{t-1}) \quad (3.46)$$

Similarly, the identification strategy of [Fengler and Herwartz \(2018\)](#) used in DCC Fengler yields also symmetric IRFs on impact. This is because the time-varying matrices $\mathbf{B}_{0,t}^{-1}$ (and hence $\mathbf{B}_{0,t}$) are symmetric as, $\forall t$ and knowing that $\boldsymbol{\Lambda}_t$ is diagonal and therefore symmetric:

$$(\mathbf{B}_{0,t}^{-1})' = (\boldsymbol{\Gamma}_t \boldsymbol{\Lambda}_t^{1/2} \boldsymbol{\Gamma}_t')' = \boldsymbol{\Gamma}_t (\boldsymbol{\Gamma}_t \boldsymbol{\Lambda}_t^{1/2})' = \boldsymbol{\Gamma}_t \boldsymbol{\Lambda}_t^{1/2} \boldsymbol{\Gamma}_t' = \mathbf{B}_{0,t}^{-1} \quad (3.47)$$

To conclude, identification with GIRFs or the identification of [Fengler and Herwartz \(2018\)](#) impose a symmetric structure of impulse responses upon impact while identification by Cholesky assumes a recursive one. These assumptions may be controversial when it comes to financial data which tend to respond rapidly to shocks and where variables react asymmetrically to each other.

4. Conclusion

Concluding remarks

We tried, throughout these three chapters, to highlight the importance of micro/sectoral financial data to answer macro-financial questions. These beneficial aspects are ventilated along three dimensions in our case: the contrasting behaviours between sector-level and index-level return predictability, the extra-accuracy that can be gained with sectoral equity variables in macro-forecasting and, eventually, the new insights that can emerge with bank CDS data to better understand the dynamics in sovereign CDS series.

These research works have several policy implications from a central bank point of view. First, identifying financial bubbles, even *ex post*, is a difficult exercise and requires a diversified range of indicators, such as the Cyclically adjusted price-to-earnings ratio (CAPE, [Campbell and Shiller, 1988](#)) or the number of IPO announcements ([Huang et al., 2015](#)). As a result, relying on the alpha-predictability index outlined in the first chapter can be useful in the financial stability toolkit to spot periods of elevated market inefficiencies.

Second, the use of sectoral equity series can definitely help central banks in their macroeconomic diagnosis, given their predictive content that we illustrated in the second chapter of the thesis. In our case, the use of sectoral stock variables within our factor model improves the out-of-sample RMSE by close to 20 % compared to the usual forecasting predictors.

Eventually, the contagion framework based on the SVAR-GARCH model that we developed in the third chapter can be duplicated, beyond CDS-contagion analysis, to evaluate spillover patterns between virtually any financial time series. As a matter of fact, the model has been re-used in Banque de France analysis to analyze spillovers between equity returns or sovereign bond yields (see [Banque de France, 2018](#)).

Bibliography

Banque de France. 2018. Assessment of risks to the French financial system, December 2018.

Campbell, J. Y., and R. J. Shiller. 1988. Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance* 43:661–676.

Huang, D., F. Jiang, J. Tu, and G. Zhou. 2015. Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *The Review of Financial Studies* 28:791–837.

List of Figures

1.1	Hypothetical Micro- and Macro-Predictability according to the different views	11
1.2	Micro- and Macro-Raw Predictability series, over time	18
1.3	Macro-Raw Predictability series and averaged Micro-Raw Predictability, over time	19
1.4	$\overline{R^2}_{i,\alpha,t}$ and US ECY, over time	22
1.5	$\overline{R^2}_{i,\beta,t}$ and US Unemployment rate, over time	23
1.A.1	Individual Micro- and Macro-Raw Predictability series, over time	39
1.A.2	Distribution of Micro-Raw Predictability Statistics	40
1.A.3	Mean and standard deviations of raw predictability series: confidence intervals	42
1.A.4	Macro-Raw Predictability series and averaged Individual Raw Predictability	43
1.A.5	Raw Predictability levels vs. Returns standard deviations	44
1.A.6	Raw Predictability standard deviations vs. Returns standard deviations	45
1.A.7	$\overline{R^2}_{i,\alpha,t}$ with different Factor Specifications	46
1.A.8	$\overline{R^2}_{i,\beta,t}$ with different Factor Specifications	47
2.1.1	S&P 500 and US Industrial Production (100 = Dec.2019)	51
2.4.1	In-Sample RMSE from the different estimated models	61
2.5.1	Out-of-Sample RMSE from the different estimated models	64
2.6.1	Factor weights and DYs correlations with future IP growth and future aggregate returns (In-sample estimates, forecasting over a 12-month horizon)	72
2.6.2	Factor correlations along with Sectoral and Aggregate DY correlations (In-sample estimates, forecasting over a 12-month horizon)	74
2.6.3	Absolute factor weights (In-sample estimates, forecasting over a 12-month horizon)	75
2.A.1	Robustness check, In-Sample RMSE from the different estimated models	85
2.A.2	Robustness check, Out-of-Sample RMSE from the different estimated models	86
2.A.3	Estimated Factor, Market DY and Lead IP growth (In-sample estimates, forecasting over a 12-month horizon)	90
3.5.1	Structural Shocks and Events	110
3.5.2	Total Spillover Indices from different Models	112
3.5.3	Outward Spillovers from Italian sovereign and bank shocks	113
3.5.4	Pairwise spillovers from EZ sovereigns and banks (%)	119

3.A.1 CDS time series (basis points, first difference)	138
3.A.2 Robustness graphs	142

List of Tables

1.1	Regression results for the Alpha-Predictability	26
1.2	Regression results for the Beta-Predictability	27
1.A.1	Estimated Models	35
1.A.1	Estimated Models	36
1.A.2	External regressors used in Model 9 to 20 in Table 1.A.1	37
1.A.3	Additional external regressors used in Sections 1.5.2 and 1.6	38
1.A.4	Additional Regression Results for the Alpha- and Beta-Predictability	48
2.4.1	Predictive coefficients of the estimated factor (In-sample estimates)	62
2.5.1	Difference in RMSE with the main benchmark models (Factor model Corresponding benchmark, Out-of-sample estimates)	65
2.5.2	Difference in RMSE with alternative factor models (Factor model Corresponding PCA-factor benchmark, Out-of-sample estimates)	67
2.5.3	Difference in RMSE by Period (Factor model Corresponding benchmark, Out-of-sample estimates)	69
2.6.1	Absolute factor weights regressions (In-sample estimates, forecasting over a 12-month horizon)	77
2.A.1	Difference in RMSE with volatility models (Factor model Corresponding benchmark, Out-of-sample estimates)	87
2.A.2	Difference in RMSE by country (Factor model Corresponding benchmark, Out-of-sample estimates, 12-month horizon)	87
2.A.3	List of the variables used to estimate PCA-factors	88
2.A.3	List of the variables used to estimate PCA-factors	89
3.3.1	Diebold-Yilmaz Spillover Table	98
3.5.1	Forecast error variance decomposition, average over time (%)	108
3.5.2	Percentage of good event-identification by model	115
3.5.3	Factors associated with spillovers from sovereigns to sovereigns	121
3.5.4	Factors associated with spillovers from banks to banks	123
3.5.5	Factors associated with spillovers from banks to sovereigns in same country	126
3.5.6	Factors associated with spillovers from sovereigns to banks in the same country	127
3.A.1	List of banks used in bank sector CDS time series	137
3.A.2	Granger causality between the models (First difference, lags=2)	139

3.A.3 Test for identification in SVAR-GARCH	143
3.A.4 Historical Events	143
3.A.4 Historical Events	144
3.A.4 Historical Events	145
3.A.4 Historical Events	146

RÉSUMÉ

Marchés financiers et banques centrales entretiennent des liens étroits. D'une part les banques centrales peuvent tirer profit des signaux émis par les marchés financiers, par exemple pour évaluer l'opinion des investisseurs sur le futur niveau de l'activité économique. D'autre part, du point de vue de la stabilité financière, les banques centrales doivent également suivre les marchés financiers afin de jauger des risques qu'ils pourraient faire peser sur l'économie réelle. Ces derniers peuvent notamment émerger de processus de bulles financières ou de la propagation de chocs négatifs de marché à marché. Toutefois, pour l'étude de ces questions, la littérature macro-financière s'est historiquement reposée sur l'utilisation de données financières agrégées. Par opposition, cette thèse de doctorat est centrée sur l'idée que, sur ces sujets, les données désagrégées/sectorielles peuvent ouvrir de nouvelles perspectives de compréhension par rapport aux données macros. Le premier chapitre de cette thèse évalue ainsi dans quelle mesure la prédictibilité des rendements action peut différer entre le niveau indiciel et le niveau individuel/sectoriel. Nous montrons de plus comment cet exercice permet de délimiter les phases durant lesquelles la prédictibilité des rendements est liée aux inefficiences de marché, et ainsi d'identifier les périodes « d'exubérance irrationnelle » (Shiller, 2015). Le deuxième chapitre souligne comment l'utilisation de données sectorielles, et non pas agrégées, tirées du marché action permet d'améliorer notre capacité à prévoir l'activité économique dans le cadre d'un modèle à facteurs. Dans le troisième chapitre, à l'aide d'un modèle économétrique innovant, nous montrons comment l'incorporation de séries temporelles à la fois au niveau des pays et au niveau des secteurs permet de mieux identifier les chocs affectant les secteurs souverains et bancaires et de mieux estimer la propagation de ces derniers.

MOTS CLÉS

Marchés financiers, Inefficiences de Marché, Économétrie financière

ABSTRACT

The relationship between financial markets and central banks is deeply intertwined. On one hand central banks can for example rely on the signals stemming from the markets to assess investors' opinions regarding future economic activity. On the other hand, from a financial stability perspective, central banks need also to monitor the activity on financial markets to assess the build-up of risks for the real economy that can emerge from bubble processes or from the propagation of negative shocks. However, historically, the macro-financial literature dedicated to these issues mostly rely on aggregate data. This PhD thesis is centered on the fact that focusing on a macro-perspective on these subjects can obscure a lot of information compared to using micro/sectoral data. The first chapter of the thesis thus investigates how stock return predictability can differ between the index-level and the micro-level. Through this exercise, we manage also to spot times where stock return predictability stem from market inefficiencies, enabling us to identify periods of "irrational exuberance" (Shiller, 2015). The second chapter highlights how relying on sectoral, rather than on aggregate, equity data within a factor model can improve our ability to forecast future economic activity. In the third chapter, with the use of an innovative econometric model, we show how incorporating both country-level and sector-level time series in our framework helps to better identify sovereign and bank credit shocks as well as the propagation of the latter.

KEYWORDS

Financial Markets, Market Inefficiencies, Financial Econometrics