Essays in empirical finance

Jankauskas, Tomas

DOI:
10.26116/qh2x-nq45

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

CentER

Essays in Empirical Finance

TOMAS JANKAUSKAS

TILBURG UNIVERSITY

Essays in Empirical Finance

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. W.B.H.J. van de Donk, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op vrijdag 25 augustus 2023 om 16.00 uur

door

Tomas Jankauskas

geboren te Vilnius, Litouwen.

Promotores: Prof. dr. J.J.A.G. Driessen (Tilburg University)
Prof. dr. L.T.M. Baele (Tilburg University)

Promotiecommissie: Prof. dr. B.J.M. Werker (Tilburg University)
Prof. dr. M.F. Boons (Tilburg University)
Prof. dr. V. Ioannidou (City University of London)
Dr. M.M.J.E. Cosemans (Erasmus University Rotterdam)

Acknowledgments

My Ph.D. was a long journey with ups and many valleys of struggle. After completing my undergraduate degree, it took additional seven years to finish my doctoral studies, and I would not be able to endure this challenge without the vast support from others. It is only fair that I start this dissertation by expressing my heartfelt gratitude to all those who cared for me, aided me, and served as a source of inspiration.

Throughout my Ph.D., my supervisors were pivotal in providing guidance and direction. I want to thank my supervisor, Joost Driessen, for his immense support over the last six years. I met him in the summer of 2017 for my first finance research project, and I was exceptionally fortunate to become his student. During my Ph.D. years, Joost explored several different venues in his career. However, he always found time for me and displayed a huge amount of dedication to my supervision. There are no words to describe the significance of his role in my stay here at Tilburg. I attempted to absorb as much knowledge as I could from his supervision, but I never reached the depths of what he had to offer. Thank you, Joost. I will forever be grateful for your help.

I want to thank Lieven Baele, my second supervisor, for his unwavering support, particularly during the final years of my Ph.D. Perhaps few may be aware that it was Lieven who wrote the reference letter facilitating my admission to the Research Master program at Tilburg University. By some coincidence, Joost invited Lieven to become my supervisor too. Despite his role as the head of the department, Lieven always made time to discuss any concerns, whether they pertained to research or personal matters. At the same time, Lieven has always trained me to distill the core message of my papers, a skill that proved particularly valuable during the thesis writing process and my job market endeavors. Thank you, Lieven. I always felt that you played a fatherly role in my supervision.

To both of my supervisors, I extend my gratitude for their patience while I was dealing with my countless health complaints. I am sure that my experiences yielded lasting memories for many years ahead. Hopefully, your future students, including myself, will not repeat or beat my streak of misfortunes in the first three Ph.D. years. I appreciate your understanding during those difficult periods.

This thesis would not be the same without my co-author and unofficial supervisor, Nicola Pavanini. Nicola taught me a great deal about research, and without his involvement, my visit to the European Central Bank and our joint project would not have been possible. Meeting him during my Research Master's years was akin to winning the lottery. The same words of appreciation extend to Stefano Giglio, who hosted my stay at the Yale

School of Management in 2022. With little to no hesitation, Stefano invited me for a visit which fulfilled my long-term dream. I was amazed by the attention and the opportunities that Stefano selflessly provided. Meeting Stefano inspired me to explore career prospects beyond the Atlantic Ocean, which eventually became a reality.

I want to thank my committee members Bas Werker, Martijn Boons, Mathijs Cosemans, and Vasso Ioannidou for thoroughly reading my dissertation and for all your valuable input. I would also like to express my gratitude to the entire finance department of Tilburg University for raising me as a young financial economist. I was lucky enough to interact with almost every faculty member, and all of you deserve my compliments. Special mention goes to Frank de Jong, Fabio Braggion, Julio Crego, Martijn de Vries, Magdalena Rola-Janicka, and Rik Frehen, who actively offered their helping hand during my job market.

I could not have the same memories of Tilburg without the presence of many friends and colleagues. I thank Daniel Kárpáti for his friendship. Daniel offered his shoulder in many challenging moments of my life, and I could fill pages in this dissertation listing the times he answered my calls for help. I want to thank Zhaneta Tancheva for her tremendous support and care during my struggles and for all the invaluable lessons we shared, which helped me grow as a person. I would like to thank Rafael Greminger and Martin Kapons for inviting me to countless walks in the Tilburg University forest. During the coronavirus pandemic, they were my few colleagues that helped me to stay sound and hopeful. And, of course, I would like to express my gratitude for all the laughter and happy moments I shared with my colleagues Giuseppe, Jan, Zihao, Giovanni, and Michael. I also greatly appreciate the time spent with Jorgo, Artur, Shuo, Samia, Ertunc, Gleb, and Joren, as well as all other colleagues I met at Tilburg University, the European Central Bank, the Yale School of Management, and conferences.

Finally, I would like to thank my parents for patiently listening to my complaints about the Ph.D. for so many years. It took me over a decade to close this chapter of my life, and finally, I can inform my relatives that I am no longer a university student. My parents, particularly my mother, had to play the roles of coach and emotional therapist on a near-weekly basis. I hope my dissertation will make them proud of me.

Introduction

This dissertation consists of three chapters. The first chapter is a joint collaboration with my co-authors Ugo Albertazzi and Lorenzo Burlon from the European Central Bank, and Nicola Pavanini from Tilburg University. The second chapter is my single-authored job market paper. The last chapter is a joint project with my Ph.D. supervisors prof. Joost Driessen and prof. Lieven Baele.

In the first chapter **"The Shadow Value of Unconventional Monetary Policy"**, we quantify how central bank unconventional monetary policy, in the form of funding facilities, reduced the banking sector's fragility in the euro area in 2014-2021. We estimate a micro-structural model of imperfect competition in the banking sector that allows for multiple equilibria with bank runs, banks' default and contagion, and central bank funding. Our framework incorporates demand and supply for insured and uninsured deposits, for loans to firms and households, and borrowers' default. We use confidential granular data for the euro area banking sector, including information on banks' borrowing from the European Central Bank (ECB). We document the presence of alternative equilibria with run-type features, but also that central bank interventions exerted a crucial role in containing this risk. Our counterfactuals show that, on average across equilibria, a 1 percentage point reduction in the ECB lending rate leads to a 1.4 percentage points reduction in banks' default probability.

In the second chapter **"The Term Structure of Corporate Bond Risk Premia"**, I construct an implied estimate of the term structure of corporate bond risk premia based on yields and estimated default probabilities. I document an upward-sloping term structure of risk premia both at the aggregate bond market level and across various cross-sectional sorts over 2002-2020. The implied estimates reveal that most of the credit risk, value and size premia in corporate bonds are earned on short-duration bonds, suggesting that the investors are particularly averse to the short- and intermediate-horizon risks. These patterns are not detectable by the average realized returns that in the same sample period exhibit a slightly downward-sloping term structure. I show that a significant fraction of the realized returns is driven by downward-trending risk premia over the last 20 years, which substantially reduced the informativeness of realized returns about risk premia.

In the third chapter **"The Implied Equity Term Structure"**, we propose a new methodology to estimate the term structure of equity risk premia from the cross-section of stock prices. Our implied equity term structure equates market prices to the discounted value of projected firm-level cash flows. Our method allows for cross-sectional and time-series variation in the term structure. We find the market-wide term structure to be upward sloping on average, but flatten out during recessions. Stocks sorted within most portfolios have an upward sloping term structure, but we do find the smallest value firms and small firms with low credit ratings to have a slightly downward sloping term structure.

Contents

Chapter 1

The Shadow Value of Unconventional Monetary Policy*

Co-authored with Ugo Albertazzi, Lorenzo Burlon and Nicola Pavanini

*The authors gratefully acknowledge Carlos Cañon, Fabio Castiglionesi, Alessandro Gavazza, Sergio Mayordomo, Christian Offermanns, Rafael Repullo, Roberto Robatto, Enrico Sette, Fabiano Schivardi, Adi Sunderam, as well as seminar participants at the European Central Bank, Tilburg University, 23^{rd} Central Bank Macro Modelling Workshop (Norges Bank), 4^{th} Annual Workshop of the ESCB Research Cluster on Monetary Economics (Bundesbank), 19^{th} Workshop on Macroeconomic Dynamics (Bank of Italy), Workshop on Monetary policy, Macroprudential Policy and Financial Stability (European Central Bank), IIOC 2021, Dutch National Bank, 2021 Stanford Institute for Theoretical Economics conference on Financial Regulation, Zürich Workshop on Sustainable Banking. The opinions in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank or the Eurosystem.

1.1 Introduction

Since the seminal contribution by Diamond & Dybvig (1983) and Goldstein & Pauzner (2005), it is well understood that banks are intrinsically fragile institutions, as they are subject to a host of strategic complementarities such as expectations over the performance of their exposures, the evolution of their funding costs, or the behaviour of competitors in lending and deposit markets. In this context, it is recognized that one of the main effects and purposes of central bank unconventional monetary policy, in the form of liquidity injections and refinancing operations, is to prevent the materialization of adverse equilibria with runs on retail or wholesale bank funding, eventually resulting in disorderly deleveraging with potentially very large welfare losses. Largely motivated by this purpose, in the last ten years the European Central Bank launched a series of massive short and long term refinancing operations, with peak take up of €2.2 trillions, corresponding to over 18% of the euro area GDP.

While other institutional features exist to temper the risk of bank runs, notably the presence of deposit insurance schemes, monetary policy can be considered to maintain a crucial role in dealing with banks' intrinsic fragility, due to several factors.[1] First, moral hazard considerations explain why deposit insurance schemes universally envisage only a partial coverage I.A.D.I. (2013). Second, deposit insurance could be ineffective at preventing systemic runs because it is often not financed upfront, but instead based on ex post contributions provided on a mutualistic basis by other intermediaries within the same banking sectors. Third, deposit insurance could also fail to work when the solvability of the domestic government, often considered the ultimate explicit or implicit guarantor of bank liabilities, is doubtful to begin with or is put at stake by the bank run itself, through the so-called sovereign bank nexus Dell'Ariccia, Ferreira, Jenkinson, Laeven, Martin, Minoiu & Popov (2018). Finally, as clearly shown by the experience of the global financial crisis, runs can concern not only retail deposits but also, if not primarily, wholesale ones Gorton (2010).

Surprisingly, despite the presence of this source of tail risk and the widely acknowledged role played by central banks in this context, the empirical evidence of the relevance of these prevention mechanisms is, at best, still scant. In principle, in order to be able to quantify the effectiveness of central banks' interventions in this context, one should identify and compare episodes where central banks exogenously did not intervene with comparable episodes where they intervened. This is clearly a daunting task because of the endogenous timing of these measures, which are adopted by central banks whenever the risk of a systemic run emerges. If such endogeneity issue is not adequately dealt with, both the risk of a run and the stabilization impact of monetary policy interventions will be largely underestimated. In other words, it could be argued that it is essentially impossible

[1] In our paper, we use terms fragility and intrinsic fragility interchangeably. We measure fragility using (average) bank's default probabilities at the bank (country-time equilibrium) level. Admittedly, the mean of the distribution may not be able to accurately gauge the level of sector-wide risks. For this reason, to capture aspects of contagion, we also look at the standard deviations of bank default risk in each country-time equilibrium.

to grasp the role of monetary policy in taming the risk of runs when, in equilibrium, runs are actually hardly ever observed. Moreover, run equilibria can be averted even in the absence of an explicit central bank intervention, because the very fact that agents expect the central bank to step in can be sufficient at inducing them to coordinate on a non-run equilibrium. These considerations suggest that, in order to tackle this crucial and thorny identification challenge, the empirical strategy should be based on a framework that allows constructing simulated counterfactual scenarios which can quantify the shadow value of central bank's interventions, that is what would have happened in the absence of these policies.

Our paper addresses this challenge developing and estimating such framework, and simulating those counterfactual scenarios, to quantify the effectiveness of central bank's refinancing operations at preventing bank runs in the form of multiple equilibria. Specifically, we build and estimate a structural equilibrium framework of the euro area banking sector, modeling demand and supply in imperfectly competitive deposit and loan markets, as well as borrowers' and banks' default risk, and the central bank's funding interventions. We generalize the approach of Egan, Hortaçsu & Matvos (2017), which quantifies multiple equilibria with bank run features for the US banking sector, and we analyze the effects of unconventional monetary policy on the multiplicity of equilibria and welfare. In order to quantify the value of the central bank's interventions, we extend their framework along two crucial dimensions.

First, we allow for the presence of a central bank which is willing to inject liquidity in the banking sector at pre-determined conditions. This is in line with the central bank's function of lender of last resort, through which it can alter the competition for deposits in the banking sectors, and potentially eradicate run equilibria. This role allows the central bank to lower the severity of feedback loops between high deposit rates, low profitability, and higher banks' default probabilities, reducing the multiplicity of equilibria and increasing the resilience of the system. Second, we introduce into the model a market for bank loans to the real sector under asymmetric information, following Crawford, Pavanini & Schivardi (2018). Modeling simultaneously loan granting and deposit taking is not only done for the sake of realism, but also because it is a crucial ingredient to be able to assess the implications of banks' intrinsic fragility on banks' lending capacity and ultimately on the real economy, as captured by developments in borrowing firms' default rates.

The structural model provides a characterization of banks' activity with high degree of detail, and allows for various dimensions of heterogeneity across banks. In particular, our framework models banks' behavior at the individual intermediary-level for what concerns both lending and liabilities, following the empirical industrial organization literature on demand for differentiated products Berry (1994), Berry, Levinsohn & Pakes (1995). On the deposit demand side, we distinguish between insured and uninsured depositors and estimate their preferences for bank characteristics, including interest rates and banks' default risk, while on the supply side banks compete on interest rates and have heterogeneous and time-varying marginal costs of providing deposit services. Banks are allowed to raise capital not only through deposits, but also through bonds and via borrowing from the central bank. On the lending side, we distinguish between loan demand for households and non-financial corporations (NFCs) and estimate their preferences for bank characteristics, including interest rates, while on the lending supply side banks compete on interest rates, have heterogeneous marginal costs, but also form expectations over borrowers' default risk that affect their pricing. Last, banks have limited liability and may default if a shortfall in profits exceeds their franchise value next period. The ability of the model to identify and characterize all possible multiple equilibria that would be admissible, with the same fundamentals and monetary policy that determined the observed data, allows to evaluate the resilience of the banking system to run-like episodes during its recent historical experience. The possibility to do the same under counterfactual scenarios for the monetary policy allows gauging the shadow value of such policy interventions.

The model is estimated with mostly three proprietary ECB datasets on euro-area banks and allows to analyze the various liquidity operations adopted by the ECB since 2009. In our setting, we focus on the latest rounds of the Targeted Longer-Term Refinancing Operations (TLTROs) during the period 2014-2021, which covered almost all of the ECB funding to banks with a peak take up at €2.2 trillions. Our estimation is based on the Individual Balance Sheet Indicators (IBSI) database, which reports at the unconsolidated level the main asset and liability items of over 300 banks resident in the euro area from August 2007 to July 2021. This dataset provides information on the amount of outstanding deposits, loans, and other relevant bank balance sheet information. We complement IBSI with the Individual Monetary and Financial Institutions Interest Rates (IMIR) database, which contains information on deposits and lending rates. Information on the quality of bank loans' portfolios and the breakdown between insured and uninsured deposits is obtained from confidential supervisory statistical reports. The merge of our rich data yields a representative sample of the euro area banking sector, consisting of an unbalanced panel of 64 banks for 168 months from August 2007 to July 2021, covering 13 euro area countries (Austria, Belgium, France, Germany, Greece, Ireland, Latvia, Lithuania, Italy, Portugal, Slovakia, Spain, and the Netherlands). The banks in our sample represent over 50% of their domestic loan and deposit markets on average across countries.

The demand schedules included in our structural model are estimated with instrumental variables. The main results can be summarized as follows. We find that insured depositors are considerably more price sensitive than uninsured depositors, with demand elasticities of 0.7 and 0.2, respectively. As expected, our estimates show that uninsured deposits' market shares are decreasing with banks' default probabilities, while we find that insured deposits' market shares react very mildly to banks' idiosyncratic risks. The presence of some sensitivity to bank risk also for insured deposits can be explained by the possible perceived solvability issues for some of the euro area domestic governments, especially during the sovereign debt crisis. The stronger relationship between banks' share of uninsured deposits and their default risk is what generates a potential mechanism of financial contagion across banks, which can be summarized as follows. Distressed banks, finding it hard to attract depositors in the uninsured sector, will be forced to offer more attractive rates in the insured deposit market to make up for the loss of capital. This however will push solvent banks to raise their rates too, in order not to lose insured deposits, increasing their cost of capital and negatively affecting their solvency. Crucially, this type of cross-bank contagion mechanism can be grasped only in a micro-structural framework, such as ours, that models individual banks' behavior. As expected, we find a negative demand elasticity for loans, with households being more sensitive than firms, and find that borrowers' expected default rates are increasing in loan interest rates, consistent with evidence of either adverse selection or moral hazard.

In terms of documenting multiple equilibria under the actual policy rate, our main findings can be summarized as follows. We show that on average banks' default probabilities are 9 percentage points higher in the alternative equilibria relative to the realized ones. This implies that in those alternative equilibria banks need to compensate their default risk to depositors with higher deposit rates and increase their reliance on central bank funding. The extra funding that banks can collect also maps into lower loan rates. We also document that the distribution of banks' riskiness in the alternative equilibria is characterized by a significantly thicker right tail, representing a higher risk of bank runs, which predominantly involves banks that already had a high level of riskiness in the observed equilibrium.

We complete our analysis with a series of counterfactual exercises, where we simulate scenarios with higher or lower central bank policy rates, and quantify the effect of these changes on the main outcomes of our model. We show that 1 percentage point increase in the policy rate increases on average banks' default probability by about 1.4 percentage points. We also investigate the impact of such change on the country-year weighted average and weighted standard deviation of banks' riskiness, with weights given by banks' assets, to capture the effect on the average stability if a country's banking system as well as on its volatility. We find that one percentage point increase in the policy rate increase the mean and standard deviation of banks' default risk by around 3 and 2 percentage points, respectively. We also find evidence of an asymmetric effect depending on whether the policy rate increases or decreases relative to the realized one, with policy rate increases having a significantly larger effect on the mean and standard deviation of banks' default probabilities. Last, we document that one percentage point increase in the policy rate reduces total welfare by about €22bn, equivalent to a 15% drop relative to the baseline, mostly caused by a decrease in banks' franchise value and higher expected deposit insurance costs.

Related Literature. A number of empirical studies of central bank liquidity injections, based on granular datasets and on a difference-in-differences approach, look at their impact on credit supply.[2] These studies capture the stimulative effect of the accommodative conditions at which these funds were provided, that is at cheaper conditions compared to what otherwise available in funding markets. However, these papers cannot capture

[2]For the euro area see, for example, Carpinelli & Crosignani (2018), Jasova, Mendicino & Supera (2018), Andrade, Cahn, Fraisse & Mèsonnier (2019), and Garcìa-Posada & Marchetti (2016). Similar analysis for the US and focusing on money-market mutual funds, instead of banks, is presented in Duygan-Bump, Parkinson, Rosengren, Suarez & Willen (2013).

the role of liquidity injections in averting the materialization of runs, which might have huge but yet unobservable consequences if the central bank intervention itself is successful. To better clarify the difference between the two channels, it can be pointed out that the cheap funding channel is by definition not active if central bank funds are provided at market conditions. The channel that we are instead looking at, that is the impact of central bank interventions in avoiding the realization of inefficient run equilibria, could in principle be active even if the rates applied were above prevailing market conditions.

As mentioned before, the closest article in terms of methodology is Egan et al. (2017), placing our work among a growing recent strand of papers applying structural equilibrium models from the empirical industrial organization literature to financial markets. This includes applications to insurance Koijen & Yogo (2016), asset demand Koijen & Yogo (2019), deposits Ho & Ishii (2011), Xiao (2020), commercial loans Crawford et al. (2018), Ioannidou, Pavanini & Peng (2022), Darmouni (2020), and mortgages Benetton (2021), Robles-Garcia (2020). A recent paper by Wang, Whited, Wu & Xiao (2022) also estimates a micro-structural model of the banking sector to explore the transmission of monetary policy. Their objective is to document the high importance of the banking sector's market structure in affecting the monetary policy transmission mechanism. The paper does not envisage multiplicity of equilibria and therefore does not explore the relevance of what we define as non-fundamental risk nor the role played by monetary policy in abating it. In this respect, closer to our paper is also the analysis by Robatto (2019), who develops and calibrates a macro model of the banking sector with multiple equilibria, and shows how large enough liquidity injections may eradicate bad equilibria. The most relevant difference with our approach is that, by constructing, estimating and calibrating a structural micro-level banking model, we can better capture the role of heterogeneity in the banking sector, and assess the possibility of contagion of both fundamental and non-fundamental risk (bank-specific) shocks.

Last, we also contribute to the empirical work on runs both in the banking sector and in other financial markets Iyer & Puri (2012), Iyer, Puri & Ryan (2016), Calomiris & Mason (2003). Pèrignon, Thesmar & Vuillemey (2018), by focusing on wholesale markets, can identify and explore some episodes of funding dry-ups. However, as they point out, they do not observe market freeze, possibly reflecting the presence of stabilizing factors and, in particular, of lender of last resort facilities. Moreover, the episodes they consider largely refer to intermediaries in deep distress, which hardly provides overall evidence of the systemic relevance of banks' intrinsic fragility. More recently, Artavanis, Paravisini, Robles-Garcia, Seru & Tsoutsoura (2019) provide interesting and convincing empirical evidence of run-like deposit withdrawals by examining the variation in the cost of withdrawal induced by the maturity expiration of time-deposits in Greece, but do not assess the stabilizing role of monetary policy. While their identification strategy forces them to focus on the panic-driven withdrawals triggered by a fundamental shock on bank funding, our framework can instead assess the relevance of non-fundamental risk also if totally unrelated to a deterioration of the fundamentals.

The rest of the paper is organized as follows. Section 1.2 introduces the institutional background and the data, Section 1.3 describes the model, Section 1.4 presents the estimation strategy and results, Section 1.5 displays and discusses multiple equilibria under the actual policy, Section 1.6 simulates alternative scenarios with different policies, Section 1.7 discusses limitations and extensions, and Section 1.8 concludes.

1.2 Institutional Background and Data

Since the outbreak of global financial crisis, the euro area banking sector has been exposed to a number of systemic shocks that led to significant impairment in its funding and lending capacity, leading to the adoption of unprecedented monetary policy measures.[3] The freeze in international money markets experienced in 2007 was followed soon after by the so called global financial crisis, ignited by the collapse of Lehman Brothers in September 2008. This immediately reverberated outside the US economy via a dry-up in some funding segments, such as wholesale deposits placed by non-residents, and the euro banking sector was heavily affected. In the following years Greece, Ireland, Italy, Portugal, and Spain (hereafter, the "vulnerable" countries) were involved in sovereign debt crises that strongly impaired wholesale funding conditions of the domestic banking sector. These tensions strained financial conditions due to banks' sovereign exposures, rising non-performing loan levels, and, in particular, the fact that the domestic sovereign was perceived by market participants as the explicit or implicit guarantor of bank liabilities.

In a bank-based economy such as the euro area, the fear that a material impairment in funding conditions could lead to a credit crunch, or at least prevent the transmission to the real sector of the stimulus provided by the accommodative monetary policy, motivated the adoption of a number of operations providing credit intermediaries with short-term liquidity and longer term funding.[4] Since 2008 there have been four types of unconventional monetary policy interventions based on refinancing operations.

First, starting in October 2008 the ECB allowed banks to obtain unlimited short-term liquidity at a fixed rate as long as they pledged sufficient collateral, through the so-called Fixed-Rate Full-Allotment policy. For every amount of eligible collateral, the banks could access an equal amount of liquidity minus a haircut that depended on the characteristics of the pledged collateral (asset class, residual maturity, rating, coupon structure).[5] The rate was the same as that on the Main Refinancing Operations (MROs).

Second, the ECB promoted a series of Longer-Term Refinancing Operations (LTROs). Differently from standard operations with a maturity of up to three months, these new operations extended liquidity with maturities of one year (in July 2009) and three years (in December 2011 -vLTRO I- and February 2012 -vLTRO II-), with the aim of reducing

[3]See Rostagno, Altavilla, Carboni, Lemke, Motto, Saint Guilhem & Yiangou (2021) for a detailed and comprehensive review of the conduct of monetary policy in the euro area.

[4]The sovereign crisis had opposite effects for banks in non-vulnerable countries, which experienced positive revaluations of their domestic government bond holdings and stable macroeconomic conditions.

[5]Eligible assets included government and regional bonds, covered bonds, corporate bonds, asset-backed securities, and other uncovered credit debt instruments. The large majority of the collateral was provided by government bond holdings.

roll-over risks and favoring longer-term investment. Funds available to banks were still constrained by the collateral requirements. While the central bank balance sheet was protected by the adoption of haircuts, which depended on the degree of liquidity of the assets pledged, the subsequent revision of the collateral policy substantially relaxed the collateral constraints existing for banks in accessing those funding facilities.[6] The interest rate applied was equal to the rate applied on regular short-term operations, on average over the time span of each operation, so to reflect the accommodative monetary policy stance. All these factors contributed to a high take up by banks in these operations, especially in stressed countries, and to the massive increase in the liquidity in the system (by more than a trillion euros, approximately 8% of GDP).

Third, even larger amounts of liquidity were injected via the subsequent operations adopted by the ECB. These not only supported the funding conditions and the stability of the banking sector, but also were conceived so as to avoid some of the side effects experienced with the previous operations.[7] Due to these reasons, these Targeted Longer-Term Refinancing Operations (TLTROs) are the main focus of our paper. They were announced in June 2014 (TLTRO I), in March 2016 (TLTRO II), and in March 2019 (TLTRO III). In between the waves of TLTROs, the ECB also updated the rules on borrowing limits, maturities, and early repayment options. Eligibility criteria and haircut schedules for the collateral were the same as the previous operations. Borrowing limits (for TLTRO I) and interest rates (for TLTRO II and III) differed.[8] Even though borrowing limits were in place, they were not perceived as necessarily binding in case of systemic shocks because there was an expectation that in such instances they would have been relaxed.

Finally, the pandemic brought forth, in March and April 2020, a series of re-calibrations of TLTRO III, expanding its borrowing limits, maturity, and early repayment options. Moreover, these re-calibrations consisted in an even lower pricing, which then encompassed a transitory period where the minimum achievable TLTRO III rate, subject to

[6]Pledgeable ABSs started to include securities with a lower rating and with underlying assets comprising residential mortgages and loans to small and medium enterprises (excluding mixed-class ABSs and ABSs with non-performing, structured, syndicated, or leveraged loans). Crucially, the list of pledgeable assets was extended and included an increasing number of assets, also relatively less liquid, such as individual bank loans (so-called Additional Credit Claims -ACCs-). It is also worth noting that the risk of losses on these assets remained with the corresponding national central banks instead of with the entire Eurosystem.

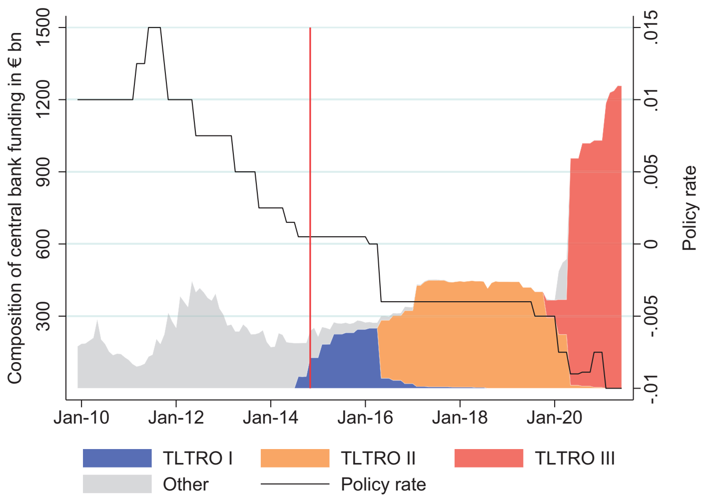
[7]See Albertazzi, Barbiero, Marqués-Ibañez, Popov, Rodriguez D'Acri & Vlassopolous (2020) for a comparative review of the papers assessing the financial stability spillover of these and other unconventional monetary policy measures.

[8]TLTRO I's borrowing limits were direct functions of the amount of loans that banks extended over the period of the operations, while the interest rates were fixed over the time span of each operation at the MRO level prevailing at the time of take-up (plus an additional fixed spread of 10 basis points for the first two TLTRO I auctions). TLTRO II's borrowing limits and interest rates were instead both functions of the loans extended over the period of the operations, with interest rates decreasing with the volume of loans from the MRO rate (which was in parallel reduced to 0%) down to the Deposit Facility Rate (DFR, the rate at which excess reserves are remunerated, which stood at −0.4%).[9] The original pricing design of TLTRO III, settled in July 2019, was similar to TLTRO II's, with the difference of a 10 basis points spread over MRO rate and DFR, which was later waived in September 2019 right before the first TLTRO III operation (together with a further DFR cut to −0.5%).

a milder lending performance criterion, was as low as -1%. The re-calibrations were accompanied by further relaxation of collateral requirements and a series of additional longer-term operations to bridge the gap between announcements of the measures and the actual operations, as well as a series of Pandemic Emergency Longer-Term Refinancing Operations (PELTROs) which acted as a further backstop for those banks whose business models did not allow for meaningful participation to TLTROs.[10]

Figure I reports the time series evolution of the total amount of ECB funding since 2010, with a breakdown across each of the operations described above, as well as the policy rate that was applied. The level of take up is not necessarily related to the degree of stabilization provided by monetary policy. It is so only conditional on the absence of runs. However, as pointed out above, the simple existence of a lender of last resort, or even just the expectation of it, may avert uncoordinated equilibria. This is why a structural model admitting multiple equilibria is needed to be able to make a comprehensive assessment of the role played by monetary policy in sustaining financial stability.

Figure I: ECB Funding and Policy Rate Within Our Sample



Note: TLTRO I, II, and III correspond to the Targeted Longer-Term Refinancing Operations announced respectively in June 2014, March 2016, and March 2019. Other corresponds to the sum of Marginal Lending Facility (MLF), Main Refinancing Operations (MROs), Longer-Term Refinancing Operations (LTROs, including the bridge operations announced in March 2020), Fine-Tuning Operations (FTOs) and Pandemic Emergency Longer-Term Refinancing Operations (PELTROs). Policy rate is the borrowing rate applied to refinancing operations over time. This figure is based on our sample of banks, which corresponds to roughly 50% of overall loan and deposit volumes, and this proportion is also reflected in the amount of ECB funding that our sample covers.

[10]See Barbiero, Boucinha & Burlon (2021) for a description of TLTRO III and the related collateral easing measures.

1.2.1 Data

Our empirical analysis relies on bank level information from various proprietary databases maintained by the ECB. First, we use the Individual Balance Sheet Indicators (IBSI) database, which reports at the unconsolidated level the main asset and liability items of over 300 banks resident in the euro area from August 2007 to July 2021. This dataset provides information on the amount of outstanding deposits, loans, and other relevant bank balance sheet information. Second, we complement IBSI with the Individual Monetary and Financial Institutions Interest Rates (IMIR) database, which contains information on deposits and lending rates. Third, we gather data on banks' Credit Default Swaps (CDS) from Datastream and on firms' Probabilities of Default (PDs) from Supervisory Reports by the Single Supervisory Mechanism of the ECB. Fourth, we add information on bank profitability and Non-Performing Loans (NPLs) from SNL Financial and Bureau van Dijk's BankScope. Lastly, we have granular information on bank's participation in ECB's lending operations and deposits in ECB's deposit facility and current account from ECB's administrative reports.

The merge of our rich data yields a representative sample of the euro area banking sector, consisting of an unbalanced panel of 64 banks for 168 months from August 2007 to July 2021, covering 13 euro area countries (Austria, Belgium, France, Germany, Greece, Ireland, Latvia, Lithuania, Italy, Portugal, Slovakia, Spain, and the Netherlands). The banks in our sample represent over 50% of their domestic loan and deposit markets on average across countries. We express all shares vis-à-vis domestic markets because in the euro area both deposit and loan markets are segmented along country lines. Although some cross-border lending does exist, it is negligible compared to the aggregate. We report the summary statistics of our sample in Table I.

For the deposit demand model we use each bank's market share of the domestic market of deposits, at the month-country level.[11] For the uninsured deposits, we use deposits of domestic corporate clients (overnight, agreed maturity, redeemable at notice), and the bank's composite interest rate on corporate deposits (weighted average of the interest rates across the segments available in the IMIR dataset). For the insured deposits, we use the deposits of domestic household clients (overnight, agreed maturity, redeemable at notice), and the bank's composite interest rate on households' deposits (weighted average of the interest rates across the segments available in the IMIR dataset). Corporate deposits are typically larger than the €100,000 threshold of the Deposit Guarantee Scheme (DGS), while household deposits are typically smaller, which makes them a good proxy for insured deposits. To validate our assumption, we obtain confidential information about the share of insured and uninsured deposits from the Supervisory Reports of the Single Supervisory Mechanism (SSM). This allows us to confirm that the share of corporate to total (household and corporate) overnight deposits is indeed highly correlated with the share of uninsured over total (insured and uninsured) deposits.

[11] Both deposit and loan market shares are calculated based on the total volume of each month-country banking sector, not only based on our sample of banks.

Table I: Summary Statistics

This table presents the summary statistics of the unbalanced panel of 64 banks for 168 months from August 2007 to July 2021, covering 13 EA countries (AT, BE, DE, ES, FR, GR, IE, IT, LT, LV, NL, PT, SK). Avg Lending Rate refers to the weighted average of lending rates to NFCs and Households that is used to estimate the loan default model.

	Obs.	Mean	St.Dev.	Min	p25	p50	p75	Max
Uninsured Deposits								
Deposit Volume (€bn)	8,295	15.96	19.01	0.00	3.04	8.15	21.86	114.28
Deposit Rate (%)	8,295	0.41	0.69	-0.58	0.02	0.11	0.47	4.70
Market Share (%)	8,295	10.63	9.60	0.00	1.64	8.69	17.87	40.71
Insured Deposits								
Deposit Volume (€bn)	8,295	40.46	51.03	0.01	9.13	22.15	55.71	418.70
Deposit Rate (%)	8,295	0.43	0.64	-0.51	0.03	0.18	0.51	5.00
Market Share (%)	8,295	9.45	9.73	0.00	1.39	6.41	14.19	43.08
Loans to NFCs								
Loan Volume (€bn)	8,295	26.97	28.83	0.30	5.93	15.41	40.70	167.91
Lending Rate (%)	8,295	2.66	1.48	-0.23	1.58	2.25	3.38	10.06
Market Share (%)	8,295	9.59	9.39	0.03	1.32	6.56	15.80	41.58
Loans to Households								
Loan Volume (€bn)	8,295	34.42	46.97	0.17	7.36	17.56	42.45	370.97
Lending Rate (%)	8,295	3.55	1.75	0.03	2.10	3.32	4.59	10.31
Market Share (%)	8,295	9.22	9.18	0.01	1.45	6.09	14.61	43.48
CDS Spread (%)	8,295	1.68	2.50	0.13	0.60	0.93	1.66	42.41
Banks' Default Prob (%)	8,295	2.64	3.64	0.21	0.98	1.52	2.69	51.85
Borrowers' Default Prob (%)	8,295	2.04	1.49	0.07	1.19	1.61	2.30	9.02
Avg Lending Rate (%)	8,295	3.14	1.50	0.20	1.91	2.86	4.09	9.58
EONIA (%)	8,295	0.12	0.93	-0.48	-0.36	-0.14	0.25	4.30
Sovereign Rate Spread (%)	8,295	1.96	2.47	-0.35	0.71	1.32	2.45	45.96
ROA (%)	8,295	0.25	0.89	-6.56	0.14	0.34	0.59	2.29
Excess Liquidity Holdings (%)	8,295	3.52	6.14	-0.09	0.00	0.71	4.60	50.88
Securities Holdings (%)	8,295	7.58	5.75	0.00	3.53	6.47	10.45	31.40
Deposit Ratio (%)	8,295	38.86	20.18	0.00	24.19	38.41	53.42	82.71
NPL Ratio (%)	8,295	6.20	6.67	0.42	2.46	4.09	6.96	42.49
Loss Given Default (%)	8,295	27.39	7.63	0.00	22.66	27.53	31.37	52.92
Net Position with CB (€bn)	8,295	-0.35	14.96	-131.55	-1.67	0.00	3.80	76.91
CB Policy Rate (%)	8,295	0.24	1.05	-1.00	-0.40	0.05	0.75	4.25
Other Net Balance (€bn)	8,295	5.32	25.91	-105.57	-5.41	1.36	12.54	169.83
Other Borrowing Rate (%)	8,295	2.08	2.69	-0.71	0.52	1.42	3.15	46.32

Deposit shares range from almost nil to over 40 percent in some jurisdictions, with an average value of 11 percent in the case of uninsured deposits and 9 percent in the case of insured deposits. Deposit rates are close to zero in most countries, reaching at maximum levels slightly below 5 percent. Some deposit rates are negative, with a few reaching levels below the minimum of the DFR in the sample period, at -0.6 percent. The average interest rate on insured and uninsured deposits is around 0.4 percent.

For the loan demand model we use bank's market share of the domestic market of loans, at the month-country level. For loans to NFCs, we use loans to domestic corporate clients, and the bank's interest rate on new corporate loans excluding overdrafts. For loans to households, we use the loans to domestic households, and the bank's interest rate on new household loans excluding overdrafts. The market of loans to NFCs is roughly as concentrated as the market of loans to households, with average shares around 10 and 9 percent, respectively. Shares in some smaller countries can reach up to over 40 percent, similarly to the deposit markets. Loan rates hover around 3 to 4 percent on average, and can reach 10 percent for some banks.

Similarly to Egan et al. (2017), we measure the financial solvency of each bank with the CDS spreads. We derive five-year CDS spreads from Datastream, and calculate the probability of default of each bank under the same risk neutral model with a constant hazard rate and under the same assumptions as in Egan et al. (2017).[12] The average CDS spread in our dataset is 168 basis points, but can reach peaks of over 4,200 basis points during the sovereign debt crisis. Under our assumptions, these peaks correspond to a sizable risk-neutral probability of bank default of 50 percent.

We measure borrowers' default with the probability of default on performing exposures reported in the Supervisory Reports of the Single Supervisory Mechanism of the ECB. In our sample, this probability is on average 2 percent and can reach over 9 percent in the aftermath of the sovereign debt crisis. Consistently, we proxy the aggregate loan interest rate that affects borrowers' default with the average interest rate on loans to the non-financial private sector.

In our regressions we control for a series of time-varying characteristics at the bank level. We include bank's ROA to proxy for profitability, the ratio of excess liquidity over assets to measure the exposure to the negative interest rate policy and the level of liquidity, the ratio of securities holdings over assets for the exposure to the capital gains from asset purchases by the ECB and the level of collateral, the ratio of deposit over assets to proxy for the business models and the exposure to the frictions emerging from the zero lower bound, and the NPL ratio as a proxy of the quality of the loan portfolio. All controls are considered with a one month lag. We also summarize the EONIA rate and the sovereign rate spread that we use as instruments for our demand models. Last, at the bottom of Table I, we report descriptive statistics for banks' loss given default from the Supervisory Reports, the net position of each bank vis-à-vis the ECB (borrowing minus deposits) as well as the policy rate that they were required to pay when borrowing from

[12] We use a 5% risk free rate and bank-month specific recovery rates.

the ECB. To complete the summary of banks' balance sheets, we include the net balance of the other components of banks' assets and liabilities, and the interest rate on these net balances as proxied by the 10-year domestic sovereign yield.

1.3 The Model

Our framework models the behavior of four agents: depositors, borrowers, banks, and the central bank. We distinguish between insured and uninsured depositors, corresponding respectively to households and non-financial corporations, and let them have preferences for banks' characteristics that determine their demand for deposit services. Depositors will consider in their deposit demand not only the interest rate offered, but also a measure of financial fragility of each financial institution. Similarly, we consider borrowers as either households or non-financial corporations, and let them have preferences for banks' characteristics that determine both their demand for loans and likelihood to default. Borrowers will choose their preferred bank based on the offered loan interest rate, which will also have an effect on their default probability, capturing any potential extent of moral hazard and/or adverse selection.

We model banks' supply of deposits and loans as Bertrand-Nash competition on interest rates, following the standard empirical industrial organization literature on demand for differentiated products Berry (1994), Berry et al. (1995). In the spirit of Hortaçsu, Matvos, Shin, Syverson & Venkataraman (2011) and Egan et al. (2017) we also let banks default if, when running a loss, their expected franchise value next period is expected to be lower in absolute value than such loss. Our framework is static and characterized by stationary pure strategy Bayesian Nash equilibria, where banks compete and decide to default within each period, not across periods. The combination of endogenous banks' default due to banks' limited liability, and of depositors' preferences for banks' stability is what allows the model to produce multiple equilibria, a key ingredient for our policy evaluations.

The degree to which the model will allow for multiple equilibria depends directly on the size of the sensitivity to price and risk conditions of the different schedules, representing the behavior of banks, depositors and firms. For instance, if depositors expect a bank to default, their expectations will be self-fulfilling, causing demand of mostly uninsured deposits for that bank to diminish, which can only be offset by offering higher deposit rates for both insured and uninsured deposits. In equilibrium this may not only validate the expectations of a bank's default, but also contribute to a contagion effect, as solvent banks are now forced to increase their deposit rates as well in order not to lose market shares, which will eventually negatively affect their solvency. Alternatively, the distressed bank may also react by charging higher lending rates, raising the riskiness of the loan book and, in turn, of the bank itself. Again, if the deterioration of the asset quality is large enough, the initial expectations of a bank's default get validated. This multiplicity based on depositors' beliefs is not eliminated by the presence of a central bank offering liquidity, as the funding it provides is usually constrained by borrowers' availability of suitable collateral.

Note that ours is not a model with bank runs in a narrow sense, because we do not have maturity transformation. However, our mechanism is not different from standard models of bank runs Diamond & Dybvig (1983). In fact, in our model a similar strategic complementarity emerges, because the withdrawal of some deposits increases the risk for other depositors. This is not because the bank is forced to liquidate (at a loss) long term illiquid projects, but simply because the withdrawals increase banks' funding costs.

Last, but crucially, we introduce a central bank, that is the ECB in our empirical application, that is willing to provide liquidity to banks at predetermined rates. No arbitrage considerations imply that the potentially unlimited availability of central bank funds determines the cost at which banks are marginally willing to borrow from comparable alternative funding sources, such as international wholesale markets, as well as the return of comparable assets. In what follows we outline the specifics of the deposit demand models, the loan demand and default models, lenders' supply through deposit and loan pricing, and banks' default decisions.

1.3.1 Deposit Demand

We model demand for deposits by specifying the indirect utilities that determine uninsured $\mathcal{N}$ (i.e. non-financial corporations) and insured $\mathcal{I}$ (i.e. households) depositors' choice of bank, where banks are allowed to provide differentiated services. More specifically, depositor i of type $d = \{\mathcal{N}, \mathcal{I}\}$ has the following indirect utility from depositing at bank j in country m at month t:

$$U^d_{ijmt} = \alpha^d P^d_{jmt} + \gamma^d F_{jmt} + \delta^d_j + \zeta^d_{mt} + \xi^d_{jmt} + \varepsilon^d_{ijmt}, \tag{1.1}$$

where P^d_{jmt} is the interest rate on deposits, F_{jmt} is a measure of bank's fragility, δ^d_j are bank fixed effects controlling for differences in depositors' mean utilities due to observed and unobserved (by the econometrician) bank characteristics, ζ^d_{mt} are country-month fixed effects absorbing any macroeconomic factor, ξ^d_{jmt} are bank-country-month unobserved characteristics (by the econometrician), and ε^d_{ijmt} are IID shocks that follow a Type 1 Extreme Value distribution. We normalize to zero the utility from choosing the outside option, that is a set of small fringe banks.[13] We allow not only uninsured depositors, but also insured ones to be sensitive to banks' fragility, to capture any costs that insured depositors might face in case of bank's default, as well as potential delays in the implementation of the deposit insurance scheme.

From these indirect utilities we can derive each bank's market share in country m at month t, both for uninsured and insured deposits, as follows:

[13] Our choice of inside vs outside option banks is mostly driven by data availability. We focus on banks for which we can observe the CDS spreads, our measure of banks' fragility, the borrowers' default probability, and the loss given default. Our final sample of (inside) banks corresponds to the largest institutions representing on average 40% of both aggregate deposits' and loans' volumes.

$$S^d_{jmt} = \frac{\exp\left(\alpha^d P^d_{jmt} + \gamma^d F_{jmt} + \delta^d_j + \zeta^d_{mt} + \xi^d_{jmt}\right)}{1 + \sum_k \exp\left(\alpha^d P^d_{kmt} + \gamma^d F_{kmt} + \delta^d_k + \zeta^d_{mt} + \xi^d_{kmt}\right)}. \tag{1.2}$$

As reported in the descriptive statistics in Table I, a small but increasing over time fraction of deposit interest rates are actually below the Zero Lower Bound (ZLB), only for uninsured depositors. Based on a recent strand of literature looking at deposit markets with rates below the ZLB Heider, Saidi & Schepens (2019), Altavilla, Burlon, Giannetti & Holton (2021), we investigated in the context of our deposit demand model whether depositors had non-linear preferences for deposit rates, which would justify a stronger demand response to deposit rates below zero, hence limiting banks' incentives to set negative deposit rates. We experimented with a quadratic term for deposit rates in the indirect utility function for both insured and uninsured depositors, as well as with as interaction of deposit rates with a dummy for negative rates for the case of uninsured depositors only, for which we have observations with negative rates. We found that none of these nonlinearities are statistically significant, possibly reflecting the presence of a negative effective lower bound below the negative values reached by deposit rates in the sample. We therefore rejected any difference in depositors' response to interest rates above or below the ZLB, and maintained the current specification with a linear relationship.

1.3.2 Loan Demand and Borrowers' Default

We model demand for loans in a similar way as demand for deposits. In particular, we define borrowers as either firms $\mathcal{F}$ (i.e. non-financial corporations) or households $\mathcal{H}$, and let each borrower $b = 1, ..., B$ of type $\ell = \{\mathcal{F}, \mathcal{H}\}$ have the following indirect utilities from taking a loan from bank j in country m in month t:

$$U^\ell_{bjmt} = \alpha^\ell P^\ell_{jmt} + \delta^\ell_j + \zeta^\ell_{mt} + \xi^\ell_{jmt} + \varepsilon^\ell_{bjmt}, \tag{1.3}$$

where $P^{\mathcal{F}}_{jmt}, P^{\mathcal{H}}_{jmt}$ are respectively the average loan interest rates for firms and households, δ^ℓ_j are bank fixed effects, ζ^ℓ_{mt} are country-month fixed effects, ξ^ℓ_{jmt} are unobserved bank-country-month attributes, and ε^ℓ_{bjmt} are IID shocks that follow a Type 1 Extreme Value distribution. We let borrowers choose an outside option, that is any small fringe bank, and normalize to zero the utility from that option. Hence, these indirect utilities allow us to derive each bank's market share for firm and household borrowers in country m at month t as:

$$S^\ell_{jmt} = \frac{\exp\left(\alpha^\ell P^\ell_{jmt} + \delta^\ell_j + \zeta^\ell_{mt} + \xi^\ell_{jmt}\right)}{1 + \sum_k \exp\left(\alpha^\ell P^\ell_{kmt} + \delta^\ell_k + \zeta^\ell_{mt} + \xi^\ell_{kmt}\right)}, \tag{1.4}$$

Finally, we let borrowers default on their loans based on the following indirect utility function:

$$U^{\mathcal{D}}_{bjmt} = \beta P^{\mathcal{L}}_{jmt} + \delta^{\mathcal{D}}_j + \zeta^{\mathcal{D}}_{mt} + \xi^{\mathcal{D}}_{jmt} + \varepsilon^{\mathcal{D}}_{bjmt}, \tag{1.5}$$

where $P^{\mathcal{L}}_{jmt} = \frac{S^{\mathcal{F}}_{jmt}}{S^{\mathcal{F}}_{jmt}+S^{\mathcal{H}}_{jmt}} P^{\mathcal{F}}_{jmt} + \frac{S^{\mathcal{H}}_{jmt}}{S^{\mathcal{F}}_{jmt}+S^{\mathcal{H}}_{jmt}} P^{\mathcal{H}}_{jmt} = (1-w^{\mathcal{H}}_{jmt})P^{\mathcal{F}}_{jmt} + w^{\mathcal{H}}_{jmt} P^{\mathcal{H}}_{jmt}$ is the weighted average of the loan interest rates for firms and households,[14] and the other controls and fixed effects follow the same logic as the loan demand models. Hence, the share of defaulting borrowers across firms and households that bank j expects to have is defined as:

$$\mathcal{D}_{jmt} = \frac{\exp\left(\beta P^{\mathcal{L}}_{jmt} + \delta^{\mathcal{D}}_{j} + \zeta^{\mathcal{D}}_{mt} + \xi^{\mathcal{D}}_{jmt}\right)}{1+\exp\left(\beta P^{\mathcal{L}}_{jmt} + \delta^{\mathcal{D}}_{j} + \zeta^{\mathcal{D}}_{mt} + \xi^{\mathcal{D}}_{jmt}\right)}. \tag{1.6}$$

Finally, we assume that once default occurs, only a fraction $\mathcal{X}_{jt}$ of the loan principle and promised interest payment is lost, with $1-\mathcal{X}_{jt}$ measuring bank-month specific recovery rates. This aims at capturing the effect of most loans being collateralized and amortized over time, which means that the default in general does not wipe out the whole principle and accrued interest. As a result, each bank's expected revenue from its loan portfolio can be expressed as:

$$(1-\mathcal{D}_{jmt})(1+P^{\mathcal{L}}_{jmt}) + \mathcal{D}_{jmt}(1-\mathcal{X}_{jt})(1+P^{\mathcal{L}}_{jmt}) = (1-\mathcal{X}_{jt}\mathcal{D}_{jmt})(1+P^{\mathcal{L}}_{jmt}). \tag{1.7}$$

It is important to discuss a restrictive assumption that we are making in this context, which has to do with the total size of the market, both in terms of deposits and loans. We are in fact assuming that banks can attract depositors and borrowers, by increasing their market shares, from a fixed pool of potential deposits' volume $M^{\mathcal{I}}_{mt}, M^{\mathcal{N}}_{mt}$ (insured and uninsured), as well as potential loans' volume $M^{\mathcal{F}}_{mt}, M^{\mathcal{H}}_{mt}$ (for firms and households). These quantities are defined respectively as the total amount of insured and uninsured deposits in country m at time t, and the total amount of loans granted to firms and households in country m at time t. This assumption, in line with Egan et al (2017), means that the model allows for substitution of quantities of deposits and loans across banks, but does not allow the aggregate volume of deposits and loans to change endogenously. Relaxing this assumption is however challenging, as it requires making an assumption over the potential market size for deposits and loans that goes beyond the observed aggregate volumes.

1.3.3 Deposit and Loan Pricing, Bank Default, and ECB Funding

On the supply side, we let banks compete Bertrand-Nash on interest rates in deposit and loan markets, but also decide on their survival depending on whether equity holders, who are subject to limited liability, find it profitable to finance a shortfall of the bank or not. We allow banks to raise capital form three different sources. First, from insured and uninsured depositors, whose interest rates are set by banks to maximize their expected equity value. Second, from the central bank, which sets a borrowing rate, that is also equivalent to a deposit interest rate if banks decide to deposit funds instead or borrowing.

[14] We use this weighted average as we only observe non-performing loans accurately enough at the bank-country-month level, not with breakdown by households and firms.

Last, from any source other than deposits and central bank funding, namely equity, debt security issuances, borrowing from other banks, and financial liabilities. While the costs faced by the banks on the first two elements of their liabilities are endogenously determined within our model, for the third one, which is introduced to match banks' assets, the cost is exogenously given, and we assume to be determined in international capital markets. We set the amount and interest rate on this latter source as fixed across our counterfactuals. On the other hand, banks' assets are represented by two main components. The first are loans granted to households and firms, while the second are any other source of assets. As for the case of liabilities, the last element is exogenously given and included to match banks' assets in the data.

Accordingly, we define the total profits of bank j in country m at month t as:

$$\begin{aligned}\Pi_{jmt} = &\sum_{\ell \in \mathcal{F},\mathcal{H}} M^{\ell}_{mt} S^{\ell}_{jmt}\left[\left(1+P^{\ell}_{jmt}\right)\left[1-\mathcal{X}_{jt}\mathcal{D}_{jmt}\right] - w^{\mathcal{L}}_{jmt} C_{jmt}\right] - M^{\mathcal{F}}_{mt} S^{\mathcal{F}}_{jmt} C^{\mathcal{F}}_{jmt} \\ &- \sum_{d \in \mathcal{I},\mathcal{N}} M^{d}_{mt} S^{d}_{jmt}\left(1+P^{d}_{jmt} + (1-w^{\mathcal{L}}_{jmt}) C_{jmt}\right) - M^{\mathcal{I}}_{mt} S^{\mathcal{I}}_{jmt} C^{\mathcal{I}}_{jmt} \\ &- M^{\mathcal{C}}_{jmt}\left(1+P^{\mathcal{C}}_{t} + C^{\mathcal{C}}_{jmt}\right) - M^{\mathcal{B}}_{jmt}(1+P^{\mathcal{B}}_{jmt}),\end{aligned} \tag{1.8}$$

where $M^{\mathcal{I}}_{mt}, M^{\mathcal{N}}_{mt}$ are respectively the total amount of insured and uninsured deposits in country m in month t, $M^{\mathcal{F}}_{mt}, M^{\mathcal{H}}_{mt}$ are the total amount of loans for firms and households, $C^{\mathcal{F}}_{jmt}$ are extra costs of providing loans to firms relative to households, and $C^{\mathcal{I}}_{jmt}$ are extra costs of providing insured deposits relative to uninsured ones. We let $M^{\mathcal{B}}_{jmt}$ be any source of capital for banks other than deposits and central bank liquidity injections, and $P^{\mathcal{B}}_{jmt}$ be its price. We take this cost of funding as exogenous, and define as $M^{\mathcal{C}}_{jmt}$ the amount that bank j borrows from the central bank, which decides on a common rate $P^{\mathcal{C}}_{t}$.[15] $C^{\mathcal{C}}_{jmt}$ captures any extra cost that bank j faces when borrowing from the central bank, such like hitting the target of maximum amount that can be borrowed as a function of its pleadable assets. Last, C_{jmt} represents any lending or deposit related stochastic costs, including administrative costs, marketing, screening and monitoring costs, and borrowers' default costs not predicted by $\mathcal{D}_{jmt}$ or other cost variables. We assume that $C_{jmt} \sim N(\mu_{jmt}, \sigma^2_{jmt})$ and that these costs are shared across loans and deposits with normalized weights $w^{\mathcal{L}}_{jmt}$ and $1-w^{\mathcal{L}}_{jmt}$.[16] We let banks' returns to be defined as:

$$R_{jmt} = \sum_{\ell \in \mathcal{F},\mathcal{H}} M^{\ell}_{mt} S^{\ell}_{jmt}\left[\left(1+P^{\ell}_{jmt}\right)\left[1-\mathcal{X}_{jt}\mathcal{D}_{jmt}\right] - 1 - P^{\mathcal{C}}_{t} - C^{\mathcal{C}}_{jmt}\right] - M^{*}_{mt} S^{*}_{jmt} C_{jmt}, \tag{1.9}$$

[15] Note that in some cases we can have $M^{\mathcal{B}}_{jmt} = M^{\mathcal{F}}_{mt} S^{\mathcal{F}}_{jmt} + M^{\mathcal{H}}_{mt} S^{\mathcal{H}}_{jmt} - M^{\mathcal{I}}_{mt} S^{\mathcal{I}}_{jmt} - M^{\mathcal{N}}_{mt} S^{\mathcal{N}}_{jmt} - M^{\mathcal{C}}_{jmt} < 0$, which means that the bank borrows more than what it lends through loans. This will then become a source of revenue with the same price/cost structure.

[16] In our current estimation and counterfactual exercises we are setting $w^{\mathcal{L}}_{jmt} = 0.45$. This degree of cost sharing is calculated based on an exercise whereby operating costs are regressed on deposit and lending volumes, to capture the relative importance of each element in driving the dependent variable.

where $M^*_{mt}S^*_{jmt} = w^{\mathcal{L}}_{jmt}(M^{\mathcal{F}}_{mt}S^{\mathcal{F}}_{jmt} + M^{\mathcal{H}}_{mt}S^{\mathcal{H}}_{jmt}) + (1 - w^{\mathcal{L}}_{jmt})(M^{\mathcal{I}}_{mt}S^{\mathcal{I}}_{jmt} + M^{\mathcal{N}}_{mt}S^{\mathcal{N}}_{jmt})$. Banks' risk neutral equity holders will decide to finance a shortfall if the equity value of the bank next period E_{jmt} exceeds the shortfall, based on the following condition:

$$\Pi_{jmt} + \frac{1}{1+r}E_{jmt} > 0, \tag{1.10}$$

where the equity value next period is determined by the expected value of banks' returns R_{jmt} conditional on survival, times their survival probability. This means that we are not explicitly assuming that banks have any equity. There will be a threshold level of C_{jmt} such that equity holders are indifferent between financing the bank in country m and month t and letting it default, defined as $\overline{C}_{jmt}$. We can then solve for the optimal cutoff rule as follows:

$$\begin{aligned}
-\Pi_{jmt}\left(\overline{C}_{jmt}\right) &= \frac{1}{1+r} \overbrace{\overbrace{\Phi\left(\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)}^{\text{Survival Prob}} \left[\overbrace{\mathbb{E}\left(R_{jmt}\left(C_{jmt}\right) - R_{jmt}\left(\overline{C}_{jmt}\right) \mid R_{jmt}\left(C_{jmt}\right) - R_{jmt}\left(\overline{C}_{jmt}\right) \geq 0\right)}^{\text{Expected Return conditional on survival}}\right]}^{E_{jmt}} \\
&= \frac{1}{1+r}\Phi\left(\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)\left[M^*_{mt}S^*_{jmt}\left(\overline{C}_{jmt} - \mu_{jmt} + \sigma_{jmt}\lambda\left(-\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)\right)\right],
\end{aligned} \tag{1.11}$$

where we let $M^{\mathcal{L}}_{mt}S^{\mathcal{L}}_{jmt} = M^{\mathcal{F}}_{mt}S^{\mathcal{F}}_{jmt} + M^{\mathcal{H}}_{mt}S^{\mathcal{H}}_{jmt}$ and $\lambda(.)$ is the inverse Mills ratio.[17] Similarly to Egan et al. (2017), a crucial feature of the first order condition in equation (1.11) is that it can be satisfied by multiple values of bank's default probability, which gives rise to multiplicity of equilibria for the same model primitives (preferences and costs). The feedback loop between depositors' demand depending on bank's risk, and bank's risk depending on depositors' demand, implies that banks' default probabilities perceived by depositors can be self fulfilling, generating panic-based runs as in Diamond & Dybvig (1983) and Goldstein & Pauzner (2005).

Before observing the realization of the costs C_{jmt}, banks set deposit and loan interest rates $P^{\mathcal{I}}_{jmt}$, $P^{\mathcal{N}}_{jmt}$, $P^{\mathcal{F}}_{jmt}$, $P^{\mathcal{H}}_{jmt}$ maximizing their equity value, solving the following optimization problem under limited liability and risk neutrality:

$$\begin{aligned}
E_{jmt} &= \max_{P^{\mathcal{I}}_{jmt}, P^{\mathcal{N}}_{jmt}, P^{\mathcal{F}}_{jmt}, P^{\mathcal{H}}_{jmt}} \int_{-\infty}^{\overline{C}_{jmt}} \left[\Pi_{jmt} + \frac{1}{1+r}E_{jmt}\right] dF(C_{jmt}) \\
&\equiv \max_{P^{\mathcal{I}}_{jmt}, P^{\mathcal{N}}_{jmt}, P^{\mathcal{F}}_{jmt}, P^{\mathcal{H}}_{jmt}} \Bigg[R_{jmt} - M^{\mathcal{F}}_{mt}S^{\mathcal{F}}_{jmt}C^{\mathcal{F}}_{jmt} - M^{\mathcal{I}}_{mt}S^{\mathcal{I}}_{jmt}\left(P^{\mathcal{I}}_{jmt} + C^{\mathcal{I}}_{jmt} - P^{\mathcal{C}}_{t} - C^{\mathcal{C}}_{jmt}\right) \\
&\quad - M^{\mathcal{N}}_{mt}S^{\mathcal{N}}_{jmt}\left(P^{\mathcal{N}}_{jmt} - P^{\mathcal{C}}_{t} - C^{\mathcal{C}}_{jmt}\right) - M^{\mathcal{B}}_{jmt}\left(P^{\mathcal{B}}_{jmt} - P^{\mathcal{C}}_{jmt} - C^{\mathcal{C}}_{jmt}\right)\Bigg]\Phi\left(\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right),
\end{aligned} \tag{1.12}$$

[17] The formula for the inverse Mills ratio is $\lambda\left(-\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right) = \frac{\phi\left(\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)}{\Phi\left(\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)}$.

where:

$$C_{jmt} = \mu_{jmt} - \sigma_{jmt}\lambda\left(-\frac{\overline{C}_{jmt} - \mu_{jmt}}{\sigma_{jmt}}\right). \tag{1.13}$$

We use the four first order conditions of this optimization problem to back out the unobserved cost components of the bank's objective function, as described in detail in the Appendix B. Those equilibrium conditions, together with the optimal cutoff rule of equation (1.11), allows us to derive $C^{\mathcal{I}}_{jmt}, C^{\mathcal{F}}_{jmt}, C^{\mathcal{C}}_{jmt}, \mu_{jmt}, \sigma_{jmt}$.

1.4 Estimation

We estimate four separate but rather similar demand systems, respectively demand for uninsured and insured deposits, as well as households' and firms' demand for loans. Moreover, we estimate a similar model to determine borrowers' default probabilities. We follow an instrumental variables approach in the spirit of Berry (1994), based on aggregate market shares at the bank-country-month level for each type of depositors and borrowers.

The estimation for deposit demand is based on the following regression equation:

$$\ln S^d_{jmt} - \ln S^d_{0mt} = \alpha^d P^d_{jmt} + \gamma^d F_{jmt} + \delta^d_j + \zeta^d_{mt} + \xi^d_{jmt}, \tag{1.14}$$

where S^d_{0mt} is the market share of the outside option, that is the fringe of small banks. Note that the country-month fixed effects ζ^d_{mt} absorb the variation of the outside good, therefore we do not need to normalize the explanatory variables as difference between the value corresponding to bank j and the value corresponding to the outside good.

We address the identification concerns for both α^d and γ^d using instrumental variables. Our instruments for deposit rates is the bank-specific pass-through of the Euro Overnight Index Average (EONIA), constructed in the spirit of Villas-Boas (2007) as interactions of the EONIA with bank dummies. Our instrument for banks' CDS spreads instead is a measure of bank-specific pass-through of sovereign risk, constructed again as interactions of bank dummies with the spread between each country's sovereign yield and the EONIA. The basic idea is to identify the slope of households' demand for deposits by exploiting the variation in deposit rates which reflects shifts in banks' willingness to rely on this source of funding. Changes in the monetary policy rate are transmitted to deposit rates differently across banks, largely reflecting banks' specific characteristics, such as in particular their pricing power in the deposit market. For example, after a monetary policy tightening, some bankers will be less eager or less quick to increase deposit rates because they can rely on higher market power. Analogous considerations hold for the slope of household' demand with respect to the level of bank risk. We find these instruments to be strongly relevant in the first stage across all five models. The economic interpretation of the instruments adopted in the regressions below mimics that of the deposit demand equation.

Similarly, the estimation for the loan demand will result in the following regression equation:

$$\ln S^{\ell}_{jmt} - \ln S^{\ell}_{0mt} = \alpha^{\ell} P^{\ell}_{jmt} + \delta^{\ell}_{j} + \zeta^{\ell}_{mt} + \xi^{\ell}_{jmt}. \tag{1.15}$$

Last, the estimation for borrowers' default is based on the following regression equation:

$$\ln \mathcal{D}_{jmt} - \ln (1 - \mathcal{D}_{jmt}) = \beta P^{\mathcal{L}}_{jmt} + \delta^{\mathcal{D}}_{j} + \zeta^{\mathcal{D}}_{mt} + \xi^{\mathcal{D}}_{jmt}. \tag{1.16}$$

We use the set of instruments of equation (1.14) also in equation (1.15) and (1.16).

1.4.1 Results

We report the main estimates of the five models in Table II, while a more detailed summary of the results can be found in the Appendix in Tables A.I, A.II, and A.III. We first look at the demand for uninsured deposits. The results in column 1 of Table II highlight a positive effect of the remuneration of deposits on the demand for such contracts. However, they also highlight the sensitivity of deposited funds to the risk profile of the bank. A higher default probability prompts a lower demand for uninsured deposits in that bank, and this emerges even after controlling for unobserved heterogeneity related to bank-specific characteristics (i.e. bank fixed effects) or aggregate developments in the country of residence (i.e. country-month fixed effects). We then turn to the demand for insured deposits. In principle, this demand should be price-elastic, just as in the case of the demand for uninsured deposits, but should not react to banks' default probability, as the government guarantee should separate deposit safety from banks' creditworthiness for these types of contracts. We do in fact find that banks' default probabilities have no significant effect on demand for insured deposits, as reported in column 2 of Table II.

Price elasticities between the two deposit types are significantly different. In terms of magnitudes, the price elasticity is around 23 percent for uninsured deposits and 67 percent for insured deposits at the level of a 5 percent market share of the domestic market and a 1 percent interest rate. Moreover, with a 5 percent share in the domestic market, uninsured deposits' demand declines by 2 percentage points for a 1 percentage point increase in the default probability. The price elasticity of uninsured depositors is much lower than that of insured depositors. This empirical pattern is in line with the explanation that insured depositors only care about interest rates and, thus, insured deposits are much more homogeneous products. Similarly, one can argue that uninsured depositors, which are firms in our sample, depend on other bank services, which makes them less sensitive to interest rates.

We report the estimates for the loan demand in columns 3 and 4 of Table II. Similarly to deposit markets, we find that households are more sensitive to loan interest rates than firms, with a price elasticity of 14 for households and of 4 for firms. The firms display smaller elasticities for the same reasons as uninsured depositors. Most likely, these firms depend on a bundle of services from the bank, which introduces higher switching costs and makes them less responsive to changes in interest rates alone.

The last piece of the model is the equation describing borrowers' default. We report the estimates of its parameters in column 5 of Table II. We find that indeed increases in aggregate interest rate lead to a riskier borrowers' pool. A 1.5 percent increase in the aggregate lending rate, which roughly corresponds to 1 standard deviation in our sample, leads to a 4 basis points increase in borrowers' default. Considering that the standard deviation of the latter is 1.5 percent, the default equation describes a mechanism that explains over 3 percent of the unconditional variation in borrowers' default (gross of bank observables and bank and country-time fixed effects).

Table II: Deposit and Loan Demand, Borrowers' Default

The table presents the deposit and loan demand, and borrower's default estimation results. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Bank controls include ROA, excess liquidity holdings, securities holdings, deposit ratio, and NPL ratio.

	DEPOSITS		**LOANS**		
	Uninsured	**Insured**	**Firms**	**Households**	**Default**
Interest Rate	24.49***	70.48***	-4.63***	-14.57***	2.41***
	(3.32)	(4.00)	(1.39)	(1.61)	(0.76)
Bank Default Probability	-2.35***	-0.25			
	(0.93)	(0.62)			
Bank Controls	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes
Country-Month FE	Yes	Yes	Yes	Yes	Yes
Observations	8,295	8,295	8,295	8,295	8,295
R-squared	0.031	0.076	0.061	-0.017	0.015

1.4.2 Model Fit

In this Section we display model-based quantifications of a set of parameters not directly observable, with the aim to check if the time series and cross sectional patterns obtained are consistent with the financial instability episodes, as well as with the monetary policy and regulatory initiatives observed over the same period. By doing so we conduct a additional qualitative check about the overall plausibility of our modeling framework, providing an overall analysis of the fit of the model. In particular, we focus on the mean μ_{jmt} and variance σ_{jmt} of banks' unobserved costs C_{jmt}, as well as on the incremental cost $C^{\mathcal{I}}_{jmt}$ of providing insured deposits relative to uninsured, the incremental cost $C^{\mathcal{F}}_{jmt}$ of granting loans to NFCs relative to households, and the extra cost $C^{\mathcal{C}}_{jmt}$ of borrowing from the central bank.

The model-implied accounting of assets (loans) and liabilities (deposits and central bank funding) leaves a net balance sheet position for each bank in our sample. The evolution of this residual variable reflects three main developments (Figure A.I). First, in the aftermath of the crisis, euro area intermediaries have started a deleveraging process that is still ongoing. Second, the deposit base expanded across euro area countries, but especially among non-vulnerable countries. Third, deleveraging in vulnerable countries was made possible only after the adoption of the Unconventional Monetary Policies (UMPs), presumably reflecting a re-capitalization process that has been going on in parallel with the adoption of unconventional monetary policy measures.

The expected cost of lending which, based on the model, is implicit in the pricing of loans, seems to be strongly countercyclical (Figure A.II). Banks perceive borrowers' defaults as more expensive in crisis times, and the distribution normalizes again only after the adoption of UMPs. A possible and interesting interpretation of this is that in the context of a systemic crisis banks anticipate the possibility of fire sales depressing asset values, including loan collateral, thereby increasing the losses incurred form defaulted loans. The average but also the dispersion in this measure is particularly pronounced in vulnerable countries, with higher tails on both ends of the distribution (Figure A.III).

The variance of costs of borrowers' default σ_{jmt} follows a long-term downward trend (Figure A.IV). This implies, together with Figure A.II, that μ_{jmt} and σ_{jmt} were negatively correlated, at least until the adoption of UMPs, which is coherent with the notion that fire sales, by depressing collateral values in the entire economy, increase default costs across the board, diminishing cross sectional heterogeneity in default costs. The adoption of UMPs is instead associated with a decline in both parameters. In the comparison across countries, σ_{jmt} is evenly spread across intermediaries between vulnerable and non-vulnerable countries, with a lower average variance in vulnerable countries.

The opportunity cost of issuing insured deposits as opposed to uninsured ones $C^{\mathcal{I}}_{jmt}$ became permanently lower after the crisis (Figure A.VI), possibly reflecting stronger appetite for this source of funding, but also gradual perceived improvements in the institutional framework, ultimately leading to the banking union. The distribution became also more asymmetric, with a thicker left tail. The opportunity cost of lending to NFCs as opposed to households $C^{\mathcal{F}}_{jmt}$ did not change significantly over time (Figure A.VII). Instead, costs of central bank funding $C^{\mathcal{C}}_{jmt}$ gradually increased relative to other funding sources, despite the decrease in policy rates (Figure A.VIII), possibly reflecting market stigma associated with these funding sources.

1.5 Multiple Equilibria Under the Actual Policy

In this Section we review the findings obtained by simulating the model under the actual policy, with the main objective to assess whether equilibria other than the realized one are admissible and, if so, how these are characterized. Equilibria are defined as an alternative set of prices (deposit and lending rates for each bank) and banks' default probabilities that satisfy all first-order conditions in a country-year combination, given the estimated demand elasticities and the policy rate. The counterfactual scenarios for the policy rate in Section 1.6 will instead consider exogenously defined higher or lower policy rates in the different years. Any equilibrium will also be characterized by different levels of welfare and default probabilities for borrowers, although the focus will be predominantly on banks' default probabilities, the most direct measure of financial stability. For any given bank in any given year, the dispersion of its default probability across alternative equilibria is defined as non-fundamental risk. This captures the possibility that, even for given “fundamentals”, the default probability is high or low only depending on which equilibrium occurs. We define instead fundamental risk as the average default probability of a bank, in a given year, across alternative equilibria.

We consider a subsample of the data, relative to the estimation sample, for the analysis of multiple equilibria and of alternative policy scenarios, mostly for computational reasons. We focus on eight yearly snapshots of the eight largest countries, covering the main 30 banks. Two preliminary remarks can be done before discussing the properties of the set of equilibria. First, one important consistency check performed was to verify that the realized outcome of the economy (the equilibrium observed in the data) is included in the set of equilibria identified, which turns out to be the case. Second, the analysis below will de-emphasize the number of equilibria identified. One reason for this is that such number is to some extent arbitrary, as it depends on various numerical thresholds used for convergence.[18] What is instead not arbitrary and relevant is the location of such equilibria, defined in terms of the outcome variable analyzed (banks’ default probabilities). A large number of equilibria will be at hand, as it will allow representing smoother distribution functions.

To begin, Table III summarizes some descriptive statistics on the distribution of alternative equilibria, which includes the realized ones, across vulnerable and non-vulnerable countries. Those summary statistics capture the variation across equilibria in the main outcomes of the model, including banks’ default probabilities, deposit and loan volumes and rates, share of non performing loans, total ECB funding net of deposits with the ECB, and changes in depositors’ and borrowers’ surplus, banks’ profits, and total welfare between the realized equilibrium and each alternative one. At the bottom, Table III reports the average number of equilibria per country-year combinations. As expected, bank default probabilities are higher in vulnerable countries than in non-vulnerable ones. This is not reflected into a difference in deposit rates across the two groups of countries, thanks to a higher reliance on central bank funding in vulnerable countries. Moreover, higher loan rates in vulnerable economies reflect both higher levels of observable and unobservable credit risk. Surpluses and profits of the model agents, expressed as difference relative to the observed equilibrium, are consistent with the level of the relevant interest rates.

[18] Relatedly, in line with Egan et al. (2017), equilibria that are considered very similar are grouped and treated as one, which introduces another factor of arbitrariness in the number of equilibria.

Table III: Descriptive Statistics Across Countries, Years, Alternative Equilibria

The table reports descriptive statistics across all model equilibria. The descriptive statistics are calculated across 8 countries, 8 years, and all equilibria, with breakdown by 5 vulnerable countries (IT, ES, GR, IE, PT) and 3 non-vulnerable countries (DE, FR, NL). An equilibrium is counted as a country-year combination. This means that uniqueness (N of equilibria per country-year equal to 1) would imply 40 equilibria in vulnerable countries and 24 equilibria in non-vulnerable countries.

	Vulnerable Countries			**Non-Vulnerable Countries**		
	Mean	**Median**	**Std Dev**	**Mean**	**Median**	**Std Dev**
Bank Default Probability (%)	14.47	2.07	18.84	12.78	1.07	15.41
Total Deposit Volume (€bn)	561.61	410.75	445.42	1089.95	648.77	861.46
Deposit Rates (%)	6.36	0.02	9.33	6.27	0.25	8.64
Total Loan Volume (€bn)	500.25	572.92	251.78	613.77	641.81	245.12
Loan Rates (%)	0.63	2.33	3.81	0.17	1.48	2.83
Borrowers Default Probability (%)	1.95	1.75	1.15	1.30	1.27	0.37
Δ Depositors' Surplus (€bn)	43.03	-0.40	71.21	61.86	-0.07	93.47
Δ Borrowers' Surplus (€bn)	9.55	0.00	21.11	9.16	-0.20	19.11
Δ Banks' Profits (€bn)	-99.90	0.41	158.32	-191.30	0.10	264.89
Δ Total Welfare (€bn)	-47.33	-3.45	73.27	-120.29	-5.28	171.07
Net position vis-à-vis ECB (%)	2.12	9.36	30.51	-16.28	-8.79	36.00
N of Equilibria per Country-Year		12.73			32.67	

We next document in Table IV the differences in several outcomes between the realized and alternative equilibria. We do so regressing each outcome on bank-year fixed effects and a dummy for alternative equilibria, which captures the average difference between realized and alternative equilibria holding fixed any time varying bank level factors. We find that on average bank default probabilities are 9 percentage points higher in the alternative equilibria relative to the realized ones. To compensate depositors for their higher default risk, banks offer deposit rates between 5 and 4 percentage points higher, which leads to larger deposit market shares, and increase their share of funding from the ECB on average by 21 percentage points. This higher availability of funds implies that banks, in order to invest their increased funding, offer loan rates between 2 and 3 percentage points lower, which translates in larger market shares in lending markets as well. The mechanism explaining the reduction in loan rates is the following. As banks experience a costs increase on their liability side, due to their higher default risk, they try to compensate this profit loss with lower loan rates. While this decreases profits per borrower, it allows them to lend to more firms and households, and reduce borrowers' default risk, which in equilibrium leads to larger profits on the asset side.

Table IV: Comparison between Realized and Alternative Equilibria

The table compares the realized and alternative equilibria using the panel regressions. Standard errors in parentheses are clustered at the bank-year level, *** p<0.01, ** p<0.05, * p<0.1.

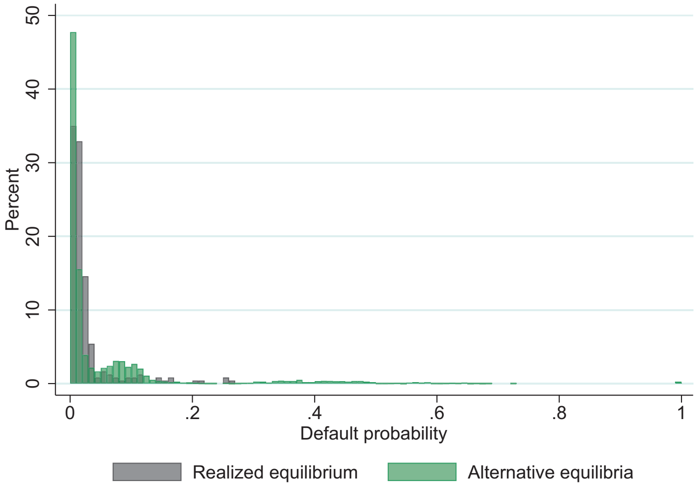
	Banks'	Deposit Rates		Loan Rates		Share CB
	Default Prob	Uninsured	Insured	Firms	Households	Funding
Alternative	0.09***	0.05***	0.04***	-0.02***	-0.03***	0.21***
Equilibrium	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Bank-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,222	6,222	6,222	6,222	6,222	6,222
R-squared	0.195	0.124	0.142	0.172	0.200	0.229

The distribution of banks' default probabilities across bank-year-equilibrium is shown in Figure II, together with the distribution of realized values across bank-year.[19] An important finding of this paper is that the former is characterized by a visibly thicker right tail. In the years considered, and at actual policy rates, the market structure of the euro area banking sector has been consistent with the existence of equilibria other than the realized ones, and characterized by significantly higher default rates. The quantitative relevance of the non-fundamental risk at around 1%, expressed by the median default probability in the bank-year distribution of the realized equilibrium, is the same as the corresponding quantity in the bank-year-equilibrium distribution. In a Diamond-Dybvig framework, the realization of run-type alternative equilibria represents a source of tail risk, which is the risk of low probability but high loss events. The relevance of this non-linearity can be expressed comparing the deterioration of the median values with that of more extreme percentiles. For example, the 75^{th} percentile of the two distributions rises from 2% to 7%, and the 95^{th} from 11% up to 40%. At the same time, it is interesting that the set of alternative equilibria also includes some lower default rates, which can be labelled as high-confidence equilibria. This is visible, for example, by comparing the left tail of the two distributions: the 25^{th} percentile in the bank-year-equilibria distribution is 0.02%, smaller than the corresponding quantity for the distribution in the realized equilibrium at 0.8%.

The interpretation of the findings above is that over the sample period scrutinized, the banking sector has endured some non-negligible levels of non-fundamental risk. Even if the accommodative monetary policy stance adopted has played a stabilizing role for the euro area banking sector, which we will discuss below when comparing alternative policy scenarios, it was not able to completely eradicate the presence of non-fundamental risk, in

[19] We focus on banks' default probabilities because they are intimately connected with the welfare costs or gains of alternative equilibria: equilibria characterized by higher average default probabilities compared to the actual ones are also equilibria where total welfare is generally lower (Table VII).

Figure II: Distribution of Realized and Alternative Default Probabilities



Note: Pooled bank-year observations. "Realized equilibrium" is the data (December observations from 2014 to 2020 plus July 2021 for the balanced panel of 30 banks). "Alternative equilibria" are the sequences of values compatible with FOCs and estimated parameters.

the form of alternative equilibria with different levels of default rates. Among the possible equilibria, moreover, the realized one was close but not coinciding with the most efficient ones, where a high level a self-fulfilling confidence would have reduced further the risk in the banking sector. Our model does not determine which equilibria will realize in the economy, so we are agnostic about the reasons why the bank-run or most efficient scenarios did not materialize. Factors like the monetary policy forward guidance and commitment to saving the euro can be the main drivers that pushed depositors' expectations towards the low-risk equilibrium. Yet, these are all explanations that lie outside of our model.

We complement the graphical evidence on tail risk from Figure II with some regression analysis in Table V, where we show how banks' default probabilities for high risk banks varies in the alternative equilibria relative to the realized one. We define high risk banks in three ways, as those that have a default probability respectively above the 50^{th}, 75^{th}, and 95^{th} percentile of the distribution in the realized equilibrium. More specifically, in Table V we display the results of a regression of bank's default probabilities on dummies for high risk bank interacted with dummies for alternative equilibria. We find that banks with default risk above the 75^{th} and 95^{th} percentile have on average a higher default probability in the alternative equilibria respectively of 8.3 and 29.2 percentage points, consistent with evidence of significant tail risk in the alternative equilibria and of positive correlation between fundamental and non-fundamental risk. This implies that weaker banks are more exposed to the occurrence of adverse equilibria.

Table V: Tail Risk in the Alternative vs Realized Equilibria

The table compares the realized and alternative equilibria using the panel regressions. Standard errors in parentheses are clustered at the bank-year level, *** p<0.01, ** p<0.05, * p<0.1.

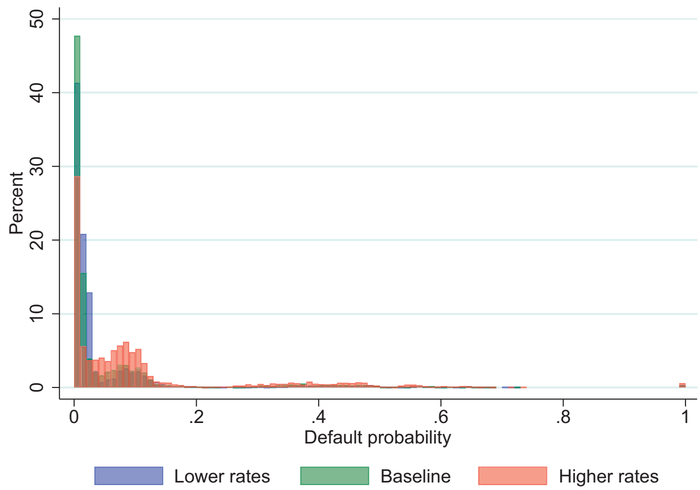
	Banks' Default Prob
High Risk Bank (above 50^{th} pctile) × Alternative Equilibrium	-0.002
	(0.006)
High Risk Bank (above 75^{th} pctile) × Alternative Equilibrium	0.083***
	(0.020)
High Risk Bank (above 95^{th} pctile) × Alternative Equilibrium	0.292***
	(0.070)
Bank-Year FE	Yes
Alternative Equilibrium FE	Yes
Observations	6,222
R-squared	0.209

1.6 Counterfactuals

In what follows we look at the response of key bank-level and country-level variables across equilibria in counterfactual scenarios, in which the policy rates at which intermediaries can borrow from the central bank are either higher (up to two percentage points higher every year from 2014 to 2019) or lower (up to two percentage points lower every year from 2014 to 2019).[20] First, as shown in Figure III, we find that banks' default rates are higher with higher rates and lower with lower rates, compared to the possible equilibria resulting from the actual level of policy rates. As we increase the policy rates, the whole distribution of banks' default probabilities shifts to the right, with an increase in the fatness of the right tail. The opposite occurs as we decrease the policy rates.

[20] We consider four counterfactual levels of the policy rate, namely plus and minus 1 and 2 percentage points compared to the baseline level. We have conducted several robustness checks with different alternative policy rates, such as plus or minus 1 basis point, 50 basis points, 1.5 percentage points, 2.5 percentage points and 3 percentage points. All qualitative results hold across calibrations.

Figure III: Distribution of Default Probability Across Policy Rate Scenarios



Note: Pooled bank-year observations. “Baseline rates” correspond to equilibria with actual policy rates. “Lower rates” (“Higher rates”) correspond to equilibria with 1 and 2 percentage point lower (higher) policy rates than actual ones.

Table VI reports how the main model outcomes presented in Table IV change across different levels of policy rates at the bank level. We regress each model outcome on the policy rate, which includes the baseline level and all the counterfactual values, and control for bank-year fixed effects. We find that 1 percentage point increase in the policy rate increases banks' default probability by about 1.4 percentage points, drives up deposit rates by around 2 percentage points (more for uninsured deposits than for insured ones) and loan rates by 11 basis points, and reduces borrowing from the central bank by over 2.3 percentage points. These results quantify the important role of unconventional monetary policy via the TLTRO refinancing operations, not only to reduce banks' default risk, but also to provide them with cheap liquidity that eases pressure on deposits and can reduce cost of credit for households and firms.

Table VI: Comparison across Equilibria with Different Policy Rates

The table compares the baseline and counterfactual (with different policy rates) equilibria using the panel regressions. Standard errors in parentheses are clustered at the bank-year level, *** p<0.01, ** p<0.05, * p<0.1.

	Banks'	**Deposit Rates**		**Loan Rates**		**Share CB**
	Default Prob	**Uninsured**	**Insured**	**Firms**	**Households**	**Funding**
Policy Rate	1.40***	2.07***	1.86***	0.11**	0.11**	-2.32***
	(0.11)	(0.09)	(0.08)	(0.05)	(0.05)	(0.35)
Bank-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28,031	28,031	28,031	28,031	28,031	28,031
R-squared	0.167	0.153	0.176	0.116	0.136	0.187

Table VII shows the effect of a change in policy rate on the weighted average and weighted standard deviation of banks' default probabilities within a country-year combination, with weights given by banks' assets. These two outcome measures are meant to capture the average stability of a country's banking system as well as its volatility. We regress these two variables on country-year fixed effects, on the policy rate across all our scenarios, but also on the policy rate interacted with dummies for an increase or a decrease in the policy rate relative to the baseline to test for asymmetric responses. We find that one percentage point increase in the policy rate increase the mean and standard deviation of banks' default risk by around 3 and 2 percentage points, respectively. We also find evidence of an asymmetric effect depending on whether the policy rate increases or decreases relative to the realized one, with policy rate increases having a significantly larger effect on the mean and standard deviation of banks' default probabilities.

The asymmetric effects on stability comes from the distributional assumptions about costs and the nature of multiplicity. The realized market equilibrium is characterized by relatively low (2-2.5%) default probabilities. Since the costs follow a normal distribution, it becomes increasingly difficult to reduce default probabilities once they approach 0, as that would mean that even in the most extreme scenarios, which are very unlikely, banks survive. On the other hand, for commensurate increases in the monetary policy rate, the default probability moves up substantially as small decreases in default threshold, $\overline{C}_{ijm}$, translate to non-negligible increases in expected shortfall. In addition, most of this asymmetric effect comes from the impact on the most adverse equilibria, which tend to be sensitive to monetary policy changes. In Appendix E, we present a parsimonious version of our model where we show that this highest-risk equilibrium is very responsive to parameter changes and, therefore, creates most of the variation across counterfactuals.

The last column of Table VII shows that 1 percentage point increase in the policy rate reduces total welfare by almost €22bn on average across countries and years, which corresponds to approximately a 15% drop relative to the baseline. We construct this measure of welfare as the sum of depositors' surplus, borrowers' surplus, the discounted sum of banks' franchise value, the aggregate deposit insurance costs and banks' bankruptcy costs, and report in Appendix C the detailed formulae. The breakdown of the effect of the policy rate on each welfare component is presented in Table A.IV, which shows that a higher policy rate favors depositors, as they obtain higher deposit rates, but harms banks' franchise value and has almost no effect on borrowers.

These results show that the ECB intervention played a significant role at reducing the average instability of euro area countries' banking systems, but also its volatility and the overall welfare level. The asymmetric effect on banks' default risk of an increase or decrease in the policy rate can be interpreted as follows. Banks under the actual policy rate, as reported in Table III, have a low default probability, which implies that they are close to the lower bound of zero in their default risk. As a consequence, while a decrease in the policy rate, which will contribute to a reduction in banks' default risk, is bound in its effect, an increase in the policy rate can almost unboundedly rise banks' default risk.

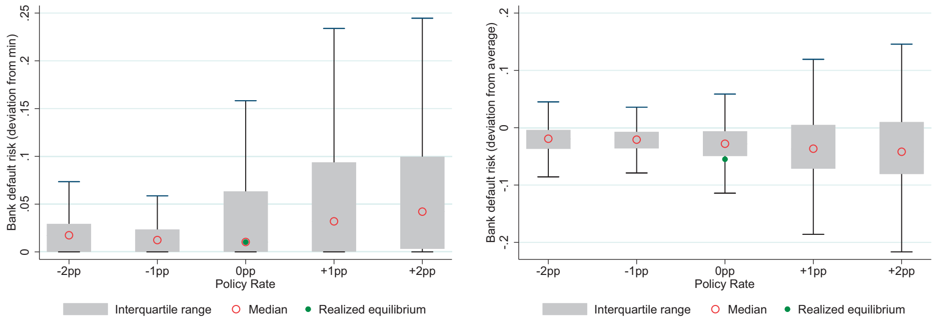
Table VII: Impact of Policy Rate on Non-Fundamental Risk and Welfare

The table compares the baseline and counterfactual (with different policy rates) equilibria using the panel regressions. The unit of observation is a country-year-equilibrium combination. Total Welfare is measured in billions of euros. Standard errors in parentheses are clustered at the country-year level, *** p<0.01, ** p<0.05, * p<0.1.

	Banks' Default Probability				**Total**
	Weighted Average		**Weighted Std Dev**		**Welfare**
Policy Rate	3.01***		1.96***		-21.94***
	(0.48)		(0.29)		(4.89)
Policy Rate × Increase		5.45***		3.21***	
		(0.94)		(0.56)	
Policy Rate × Decrease		1.01**		0.93***	
		(0.55)		(0.36)	
Country-Year FE	Yes	Yes	Yes	Yes	Yes
Observations	5,755	5,755	5,755	5,755	5,755
R-squared	0.261	0.270	0.121	0.129	0.374

Last, we present our counterfactual results also in Figure IV, where we display the distributions of banks' default risk across all the policy rate values that we simulated. The red empty circles represent the median of each distribution, while the green solid dots for the baseline policy rate ("0pp" in the graph) represent the observed equilibrium in the data. The left figure, which displays the distribution of banks' default probabilities across bank-year-equilibrium combinations in deviation from the bank-year minimum, indicates that a higher policy rate leads to a higher median bank's default probability and a higher dispersion. These results reflect the direct impact on banks' solidity of changes in the cost of central bank funding, which we know is disproportionate for the upper tail institutions (the weaker ones). However, these findings also reflect the ability of monetary policy to eradicate equilibria with runs. This is visible in the right figure, which shows the distribution of banks' default probabilities across bank-year-equilibrium combinations in deviation from the bank-year average, and as such is cleaned form the first direct funding cost channel. Lower (higher) rates are associated a with visibly smaller (larger) levels of dispersion across equilibria.

Figure IV: Distributions of Banks' Default Risk across Policy Rates



Notes: The box plots refer to the distribution, for each level of the policy rates and across bank-year-equilibrium combinations, of banks' default probabilities in deviation from the bank-year minimum (left figure) and in deviation from the bank-year average (right figure). Horizontal bars report the whiskers of the distribution.

1.7 Limitations and Extensions

The monetary policies may have broader effects that our model cannot fully incorporate. First of all, in our sample years, the ECB simultaneously pursued the quantitative easing (QE) policies that have boosted financial markets. Indirectly through the bond holdings and lower cost of capital, that could have alleviated the balance sheet constraints and reduced the overall fragility of the banking system. Our model can be generalized in that direction; however, that is not a trivial task. It requires modeling of demand for and supply of risky assets. In recent years, there have been advancements made by Koijen & Yogo (2019) in the demand-based asset pricing literature. Yet, the current tools are still limited to fully gauge the impact of QE. Nevertheless, we believe that the effect of such policies on the banking system's fragility is secondary to the TLROs, which were designed to directly target the bank balance sheets and promote lending in the economy. Therefore, it is likely that our model captures the majority of stability gains.

Unconventional monetary policies can also lead to higher inflation. Our model does not explicitly model the money supply and the real growth rates of the economy, so it does not give predictions for the changes in price levels. However, one could model the impact on inflation in a reduced form. Altavilla, Canova & Ciccarelli (2020) and Rostagno, Altavilla, Carboni, Lemke, Motto & Saint Guilhem (2021) document the TLTRO-related elasticity of inflation to changes in central bank lending rates. In particular, each 10 bps of easing in lending rates due to TLTROs leads to 0.1 p.p. of higher inflation (median impact across studies, the interquartile range goes from 0.02 p.p. to 0.12 p.p.). In our counterfactuals (Table VI), we find that a 1 p.p. increase (decrease) in the policy rate translates to, on average, 0.11 p.p. increase (decrease) in loan rates. That would imply similar magnitude effects on inflation.

1.8 Conclusion

We provide quantitative evidence of the impact that central bank unconventional monetary policy, in the form of funding facilities, have exerted on the reduction of the banking sector's intrinsic fragility. We define fragility as the presence of run-type equilibria, where lack of coordination among bank financiers leads to equilibria with higher default rates, irrespectively of the level of fundamental risk. We do so by constructing, estimating and calibrating a micro-structural model of competition in the banking sector for the euro area, that allows for both runs in the form of multiple equilibria, in the spirit of Diamond & Dybvig (1983), and for central bank liquidity injections. Crucially, our model allows for imperfect competition among banks in both deposit and loan markets. The estimation and the calibration are based on confidential granular data for the euro area banking sector, including information on the amount of deposits covered by the deposit guarantee scheme and the borrowing from the European Central Bank, over the period 2014-2021.

Our main findings can be summarized as follows. First, we document that the presence of non-fundamental risk is highly relevant in the euro area banking sector, as witnessed by the pervasiveness of the multiplicity of equilibria. Second, even under the observed and accommodative monetary policy the economy admitted multiple equilibria, on top of the observed equilibrium. Compared to the latter, the alternative equilibria tend to be characterized by worst aggregate outcomes. Some intrinsic fragility, defined as the possibility that the economy shifts to an inefficient equilibrium, has therefore been present and was not fully eradicated by the accommodative policies actually implemented. Interestingly, in isolated but meaningful cases, the economy also admitted some equilibria that were more efficient than the realized ones. This can be interpreted as suggestive that more confidence could have moved the economy into a more efficient region. We find that on average non-fundamental risk is positively related to fundamental risk, meaning that banks with higher default probability tend to be more exposed to the risk of run-type of equilibria. The simulations of counterfactual scenarios where central bank funds are artificially provided at more or less accommodative conditions indicate that monetary policy has a strong, non-linear impact in mitigating both fundamental and non-fundamental risk.

References

Albertazzi, U., Barbiero, F., Marqués-Ibañez, D., Popov, A., Rodriguez D'Acri, C. & Vlassopolous, T. (2020), Monetary policy and bank stability: The analytical toolbox reviewed. ECB Occasional Papers (forthcoming).

Altavilla, C., Burlon, L., Giannetti, M. & Holton, S. (2021), Is there a zero lower bound? the effects of negative policy rates on banks and firms. forthcoming at *Journal of Financial Economics*.

Altavilla, C., Canova, F. & Ciccarelli, M. (2020), 'Mending the broken link: Heterogeneous bank lending rates and monetary policy pass-through', *Journal of Monetary Economics* **110**, 81–98.

Andrade, P., Cahn, C., Fraisse, H. & Mèsonnier, J.-S. (2019), 'Can the provision of long-term liquidity help to avoid a credit crunch? evidence from the eurosystem's ltro', *Journal of the European Economic Association* **17**(4), 1070–1106.

Artavanis, N., Paravisini, D., Robles-Garcia, C., Seru, A. & Tsoutsoura, M. (2019), Deposit withdrawals. Working Paper.

Barbiero, F., Boucinha, M. & Burlon, L. (2021), 'Tltro iii and bank lending conditions', *Economic Bulletin Articles* **6**.

Benetton, M. (2021), 'Leverage regulation and market structure: An empirical model of the U.K. mortgage market', *Journal of Finance* **LXXVI**(6), 2997–3053.

Berry, S. (1994), 'Estimating discrete-choice models of product differentiation', *RAND Journal of Economics* **25**(2), 242–262.

Berry, S., Levinsohn, J. & Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica* **63**(4), 841–890.

Calomiris, C. W. & Mason, J. R. (2003), 'Fundamental, panics, and bank distress during the depression', *American Economic Review* **93**(5), 1615–1647.

Carpinelli, L. & Crosignani, M. (2018), The design and transmission of central bank liquidity provisions. Working Paper.

Crawford, G. S., Pavanini, N. & Schivardi, F. (2018), 'Asymmetric information and imperfect competition in lending markets', *American Economic Review* **108**(7), 1659–1701.

Darmouni, O. (2020), Informational frictions and the credit crunch. forthcoming at *Journal of Finance*.

Dell'Ariccia, G., Ferreira, C., Jenkinson, N., Laeven, L., Martin, A., Minoiu, C. & Popov, A. (2018), Managing the sovereign-bank nexus. IMF Departmental Papers, Paper No. 18/16.

Diamond, D. W. & Dybvig, P. H. (1983), 'Bank runs, deposit insurance, and liquidity', *Journal of Political Economy* **91**(3), 401–419.

Duygan-Bump, B., Parkinson, P., Rosengren, E., Suarez, G. & Willen, P. (2013), 'How effective were the federal reserve emergency liquidity facilities? evidence from the asset-backed commercial paper money market mutual fund liquidity facility', *Journal of Finance* **LXVIII**(2), 715–737.

Egan, M., Hortaçsu, A. & Matvos, G. (2017), 'Deposit competition and financial fragility: Evidence from the U.S. banking sector', *American Economic Review* **107**(1), 169–216.

Garcìa-Posada, M. & Marchetti, M. (2016), 'The bank lending channel of unconventional monetary policy: the impact of the vltros on credit supply in spain', *Economic Modelling* **58**, 427–441.

Goldstein, I. & Pauzner, A. (2005), 'Demand-deposit contracts and the probability of bank runs', *Journal of Finance* **LX**(3), 1293–1327.

Gorton, G. B. (2010), *Slapped by the invisible hand: The panic of 2007*, Oxford University Press.

Heider, F., Saidi, F. & Schepens, G. (2019), 'Life below zero: Bank lending under negative policy rates', *Review of Financial Studies* **32**(10), 3727–3761.

Ho, K. & Ishii, J. (2011), 'Location and competition in retail banking', *International Journal of Industrial Organization* **29**(5), 537–546.

Hortaçsu, A., Matvos, G., Shin, C., Syverson, C. & Venkataraman, S. (2011), 'Is an automaker's road to bankruptcy paved with customers' beliefs?', *American Economic Review Papers & Proceedings* **101**(3), 93–97.

I.A.D.I. (2013), Enhanced guidance for effective deposit insurance systems: Deposit insurance coverage. Guidance Paper.

Ioannidou, V., Pavanini, N. & Peng, Y. (2022), 'Collateral and asymmetric information in lending markets', *Journal of Financial Economics* **144**(1), 93–121.

Iyer, R. & Puri, M. (2012), 'Understanding bank runs: The importance of depositor-bank relationships', *American Economic Review* **102**(4), 1414–1445.

Iyer, R., Puri, M. & Ryan, N. (2016), 'A tale of two runs: Depositor responses to bank solvency risk', *Journal of Finance* **LXXI**(6), 2687–2726.

Jasova, M., Mendicino, C. & Supera, D. (2018), Rollover risk and bank lending behavior: Evidence from unconventional central bank liquidity. Working Paper.

Koijen, R. S. J. & Yogo, M. (2016), 'Shadow insurance', *Econometrica* **84**(3), 1265–1287.

Koijen, R. S. J. & Yogo, M. (2019), 'A demand system approach to asset pricing', *Journal of Political Economy* **127**(4), 1475–1515.

Pèrignon, C., Thesmar, D. & Vuillemey, G. (2018), 'Wholesale funding dry-ups', *Journal of Finance* **LXXIII**(2), 575–617.

Robatto, R. (2019), 'Systemic banking panics, liquidity risk, and monetary policy', *Review of Economic Dynamics* **34**, 20–42.

Robles-Garcia, C. (2020), Competition and incentives in mortgage markets: The role of brokers. Working Paper.

Rostagno, M., Altavilla, C., Carboni, G., Lemke, W., Motto, R. & Saint Guilhem, A. (2021), 'Combining negative rates, forward guidance and asset purchases: identification and impacts of the ecb's unconventional policies', *ECB Working Paper* .

Rostagno, M., Altavilla, C., Carboni, G., Lemke, W., Motto, R., Saint Guilhem, A. & Yiangou, J. (2021), *Monetary Policy in Times of Crisis: A Tale of Two Decades of the European Central Bank*, Oxford University Press.

Spence, M. (1973), 'Job market signaling', *Quarterly Journal of Economics* **87**(3), 355–374.

Villas-Boas, S. B. (2007), 'Vertical relationships between manufacturers and retailers: Inference with limited data', *The Review of Economic Studies* **74**(2), 625–652.

Wang, Y., Whited, T. M., Wu, Y. & Xiao, K. (2022), Bank market power and monetary policy transmission: Evidence from a structural estimation. forthcoming at *Journal of Finance.*

Xiao, K. (2020), 'Monetary transmission through shadow banks', *Review of Financial Studies* **33**(6), 2379–2420.

Appendices

1.A Additional Tables and Figures

Table A.I: Deposit Demand

	Uninsured		Insured	
	OLS	**IV**	**OLS**	**IV**
Interest Rate	13.91***	24.49***	42.06***	70.48***
	(2.73)	(3.32)	(3.00)	(4.00)
Bank Default Prob	-2.24***	-2.35***	-0.66**	-0.25
	(0.26)	(0.39)	(0.30)	(0.62)
ROA	-5.86***	-5.70***	4.16***	4.62***
	(0.62)	(0.65)	(1.35)	(1.24)
Excess Liquidity Holdings	0.86***	0.85***	0.17	0.12
	(0.14)	(0.14)	(0.18)	(0.18)
Securities Holdings	-0.00	0.10	1.05***	1.34***
	(0.13)	(0.13)	(0.13)	(0.15)
Deposit Ratio	0.11	0.12	0.54***	0.54***
	(0.15)	(0.15)	(0.14)	(0.14)
NPL Ratio	0.14	0.22	0.87***	1.04***
	(0.17)	(0.17)	(0.21)	(0.23)
Bank FE	Yes	Yes	Yes	Yes
Country-Month FE	Yes	Yes	Yes	Yes
IV - Monetary Policy	No	Yes	No	Yes
IV - Sovereign Risk	No	Yes	No	Yes
Observations	8,295	8,295	8,295	8,295
R-squared	0.985	0.031	0.984	0.076

Note: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table A.II: Loan Demand

	NFCs		**Households**	
	OLS	**IV**	**OLS**	**IV**
Interest Rate	-2.13***	-4.63***	-2.51**	-14.57***
	(0.75)	(1.38)	(1.19)	(1.61)
ROA	-5.82***	-5.86***	-6.58***	-5.22***
	(0.62)	(0.62)	(0.63)	(0.71)
Excess Liquidity Holdings	-0.38***	-0.38***	-0.06	0.01
	(0.12)	(0.12)	(0.15)	(0.15)
Securities Holdings	-0.51***	-0.55***	-0.32*	-0.17
	(0.12)	(0.12)	(0.17)	(0.17)
Deposit Ratio	-1.08***	-1.07***	-0.48**	-0.52**
	(0.15)	(0.15)	(0.22)	(0.22)
NPL Ratio	1.00***	1.03***	1.29***	1.49***
	(0.15)	(0.15)	(0.25)	(0.25)
Bank FE	Yes	Yes	Yes	Yes
Country-Month FE	Yes	Yes	Yes	Yes
IV - Monetary Policy	No	Yes	No	Yes
IV - Sovereign Risk	No	Yes	No	Yes
Observations	8,295	8,295	8,295	8,295
R-squared	0.981	0.061	0.979	-0.017

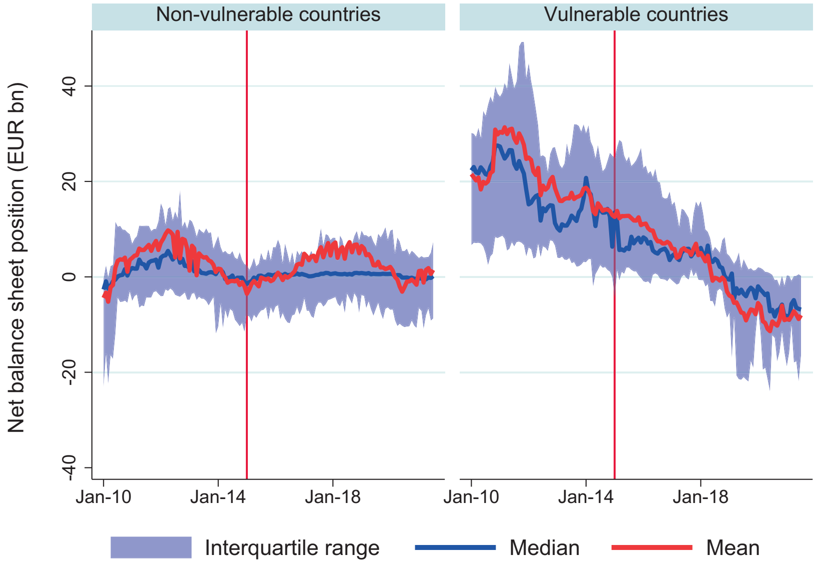
Note: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table A.III: Default Equation

	Dep Var in Level		**Transform**
	OLS	**IV**	**IV**
Avg Lending Rate	0.03***	0.03*	2.41***
	(0.01)	(0.01)	(0.76)
ROA	-0.01*	-0.01	1.09**
	(0.01)	(0.01)	(0.45)
Excess Liquidity Holdings	0.00***	0.00***	0.56***
	(0.00)	(0.00)	(0.08)
Securities Holdings	-0.01***	-0.01***	-0.80***
	(0.00)	(0.00)	(0.14)
Deposit ratio	0.00***	0.00***	-0.39***
	(0.00)	(0.00)	(0.08)
NPL ratio	0.00	0.00	0.36***
	(0.00)	(0.00)	(0.11)
Bank FE	Yes	Yes	Yes
Country-Month FE	Yes	Yes	Yes
IV - Monetary Policy	No	Yes	Yes
IV - Sovereign Risk	No	Yes	Yes
Observations	8,295	8,295	8,295
R-squared	0.971	0.015	0.025

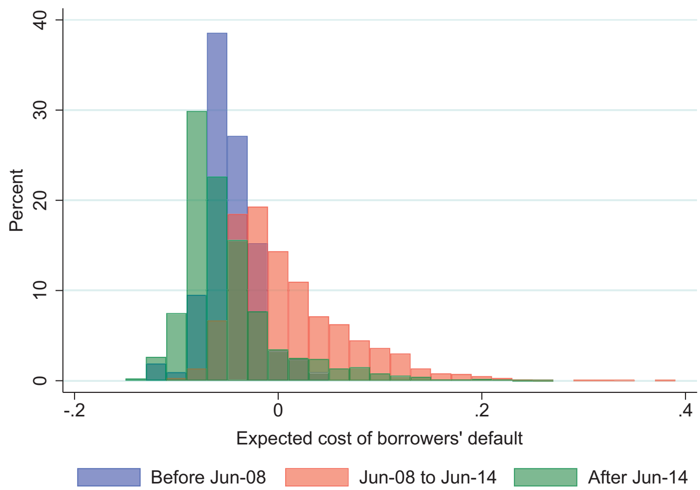
Note: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. The transform of borrowers' default probability is $log([\cdot]) - log(1 - [\cdot])$.

Figure A.I: Model-Implied Variables: Evolution of Net Balance Sheet Position



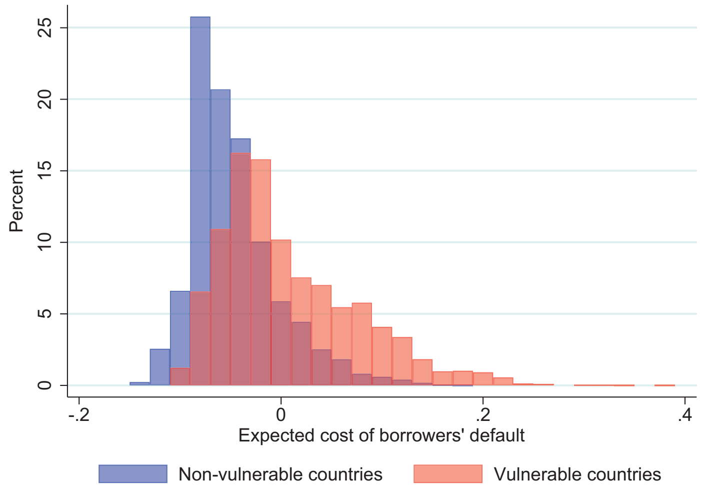
Note: Pooled bank-month observations. "Non-vulnerable countries" include AT, BE, DE, FR, LT, LV, NL, SK. "Vulnerable countries" include IT, ES, GR, IE, PT.

Figure A.II: Model-Implied Variables: Expected Cost of Borrowers' Default



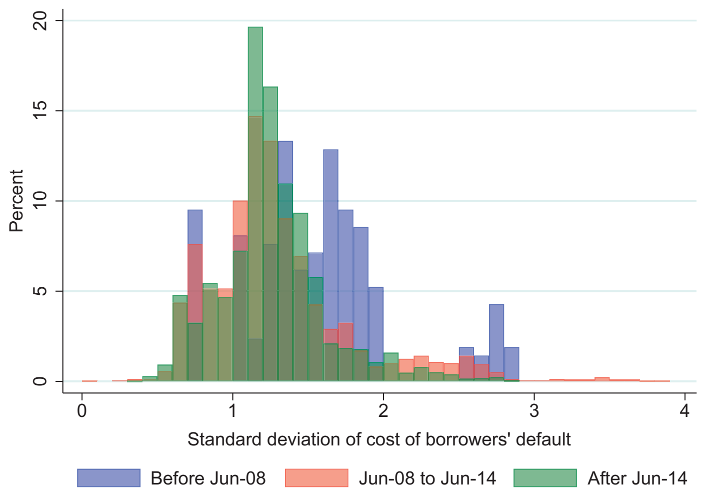
Note: Pooled bank-month observations.

Figure A.III: Model-Implied Variables: Expected Cost of Borrowers' Default Across Countries



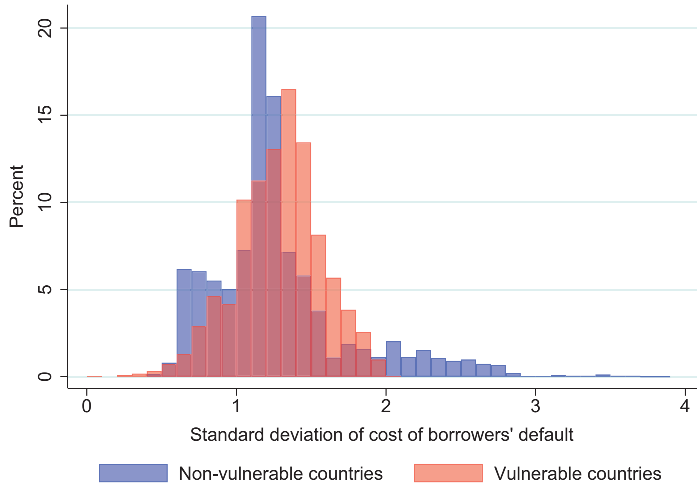
Note: Pooled bank-month observations. "Non-vulnerable countries" include AT, BE, DE, FR, LT, LV, NL, SK. "Vulnerable countries" include IT, ES, GR, IE, PT.

Figure A.IV: Model-Implied Variables: Standard Deviation of Cost of Borrowers' Default



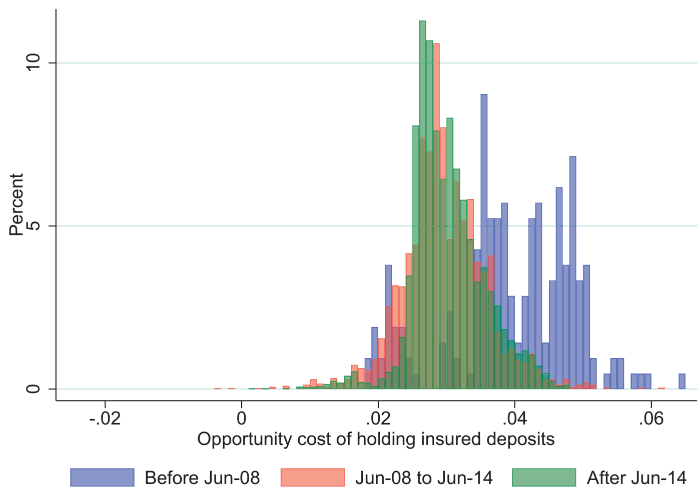
Note: Pooled bank-month observations.

Figure A.V: Model-Implied Variables: Standard Deviation of Cost of Borrowers' Default Across Countries



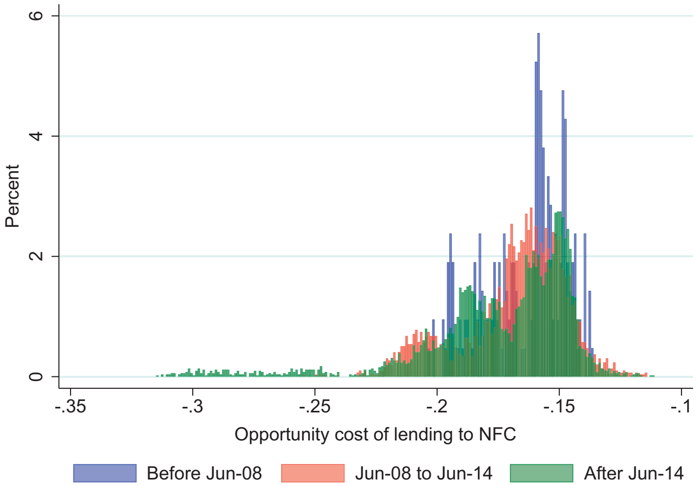
Note: Pooled bank-month observations. "Non-vulnerable countries" include AT, BE, DE, FR, LT, LV, NL, SK. "Vulnerable countries" include IT, ES, GR, IE, PT.

Figure A.VI: Model-Implied Variables: Opportunity Cost of Insured Deposits



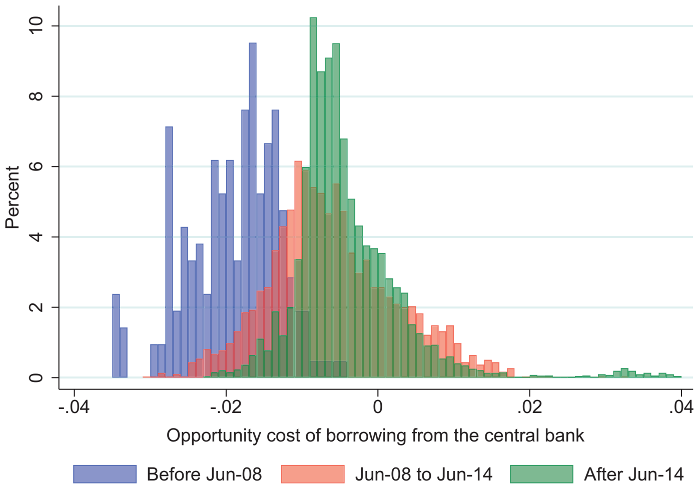
Note: Pooled bank-month observations.

Figure A.VII: Model-Implied Variables: Opportunity Cost of Loans to NFCs



Note: Pooled bank-month observations.

Figure A.VIII: Model-Implied Variables: Opportunity Cost of Central Bank Funding



Note: Pooled bank-month observations.

1.B First Order Conditions and Welfare

First Order Conditions

Note that, as in Egan et al. (2017), we assume that each bank's current decision variables do not affect the continuation value of the bank. This will result in the following two first order conditions for insured and uninsured deposit rates:

$$P_t^{\mathcal{C}} + C_{jmt}^{\mathcal{C}} - (P_{jmt}^{\mathcal{I}} + C_{jmt}^{\mathcal{I}} + (1 - w_{jmt}^{\mathcal{L}})C_{jmt}) = \frac{1}{(1 - S_{jmt}^{\mathcal{I}})\alpha^{\mathcal{I}}}, \tag{17}$$

$$P_t^{\mathcal{C}} + C_{jmt}^{\mathcal{C}} - (P_{jmt}^{\mathcal{N}} + (1 - w_{jmt}^{\mathcal{L}})C_{jmt}) = \frac{1}{(1 - S_{jmt}^{\mathcal{N}})\alpha^{\mathcal{N}}}. \tag{18}$$

From these two equations we can back out $C_{jmt}^{\mathcal{I}}$ as:

$$C_{jmt}^{\mathcal{I}} = \left(P_{jmt}^{\mathcal{N}} + \frac{1}{(1 - S_{jmt}^{\mathcal{N}})\alpha^{\mathcal{N}}}\right) - \left(P_{jmt}^{\mathcal{I}} + \frac{1}{(1 - S_{jmt}^{\mathcal{I}})\alpha^{\mathcal{I}}}\right). \tag{19}$$

We can then invert the survival probability using our measure of bank fragility as follows:

$$F_{jmt} = 1 - \Phi\left(\frac{\overline{C}_{jmt} - \mu_{jmt}}{\sigma_{jmt}}\right) \Rightarrow \frac{\overline{C}_{jmt} - \mu_{jmt}}{\sigma_{jmt}} = \Phi^{-1}(1 - F_{jmt}). \tag{20}$$

The first order conditions for loan interest rates will be the following:

$$w_{jmt}^{\mathcal{L}} C_{jmt} = (1 - \mathcal{X}_{jt}\mathcal{D}_{jmt})\left[1 + P_{jmt}^{\mathcal{H}}\right] - 1 + \frac{1 - \mathcal{X}_{jt}\mathcal{D}_{jmt}^{\mathcal{H}*}}{\left(1 - S_{jmt}^{\mathcal{H}}\right)\alpha^{\mathcal{H}}} - P_t^{\mathcal{C}} - C_{jmt}^{\mathcal{C}}, \tag{21}$$

where:

$$\mathcal{D}_{jmt}^{\mathcal{H}*} = \mathcal{D}_{jmt}\left(1 + \underbrace{(1 - \mathcal{D}_{jmt})\,\beta\left(1 + P_{jmt}^{\mathcal{H}}\right) w_{jmt}^{\mathcal{H}}}_{\text{due to changes in defaults}}\left(1 + \underbrace{\left(1 - w_{jmt}^{\mathcal{H}}\right)\left(1 - S_{jmt}^{\mathcal{H}}\right)\alpha^{\mathcal{H}}\left(P_{jmt}^{\mathcal{H}} - P_{jmt}^{\mathcal{F}}\right)}_{\text{due to compositional changes}}\right)\right), \tag{22}$$

with $\mathcal{H}$ referring to households, and $\mathcal{D}_{jmt}^{\mathcal{F}*}$ is defined symmetrically and refers to firms. This allows us to back out the unobserved extra costs of lending to firms relative to households as:

$$C_{jmt}^{\mathcal{F}} = (1 - \mathcal{X}_{jt}\mathcal{D}_{jmt})\left[P_{jmt}^{\mathcal{F}} - P_{jmt}^{\mathcal{H}}\right] + \frac{1 - \mathcal{X}_{jt}\mathcal{D}_{jmt}^{\mathcal{F}*}}{\left(1 - S_{jmt}^{\mathcal{F}}\right)\alpha^{\mathcal{F}}} - \frac{1 - \mathcal{X}_{jt}\mathcal{D}_{jmt}^{\mathcal{H}*}}{\left(1 - S_{jmt}^{\mathcal{H}}\right)\alpha^{\mathcal{H}}}. \tag{23}$$

Using first order conditions 18 and 23 we can derive the mean of the unobserved costs C_{jmt} as:

$$\begin{aligned}\mu_{jmt} =& \sigma_{jmt}\frac{\phi\left[\Phi^{-1}(1-F_{jmt})\right]}{(1-F_{jmt})} + (1-\mathcal{X}_{jt}\mathcal{D}_{jmt})\left[1+P^{\mathcal{H}}_{jmt}\right] \\ & - 1 + \frac{1-\mathcal{X}_{jt}\mathcal{D}^{\mathcal{H}*}_{jmt}}{\left(1-S^{\mathcal{H}}_{jmt}\right)\alpha^{\mathcal{H}}} - P^{\mathcal{N}}_{jmt} - \frac{1}{(1-S^{\mathcal{N}}_{jmt})\alpha^{\mathcal{N}}}.\end{aligned} \tag{24}$$

Then, from 18 we can back out the costs of borrowing from the central bank as:

$$\begin{aligned}C^{\mathcal{C}}_{jmt} =& w^{\mathcal{L}}_{jmt}P^{\mathcal{N}}_{jmt} + \left(1-w^{\mathcal{L}}_{jmt}\right)\left[(1-\mathcal{X}_{jt}\mathcal{D}_{jmt})\left[1+P^{\mathcal{H}}_{jmt}\right]-1\right] \\ & + w^{\mathcal{L}}_{jmt}\frac{1}{(1-S^{\mathcal{N}}_{jmt})\alpha^{\mathcal{N}}} - (1-w^{\mathcal{L}}_{jmt})\frac{1-\mathcal{X}_{jt}\mathcal{D}^{\mathcal{H}*}_{jmt}}{\left(1-S^{\mathcal{H}}_{jmt}\right)} - P^{\mathcal{C}}_{t}.\end{aligned} \tag{25}$$

Since the only variable part of profits are costs, we can rewrite equation 1.11 as:

$$\underbrace{\Pi_{jmt}\left(\mu_{jmt} - \sigma_{jmt}\lambda\left(-\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)\right)}_{\text{conditional expected profit}} = \frac{F+r}{1+r}M^{*}_{mt}S^{*}_{jmt}\left(\overline{C}_{jmt} - \mu_{jmt} + \sigma_{jmt}\lambda\left(-\frac{\overline{C}_{jmt}-\mu_{jmt}}{\sigma_{jmt}}\right)\right), \tag{26}$$

and then substitute into 19 and 20 to get:

$$\begin{aligned}& -\overbrace{\frac{(1-\mathcal{X}_{jt}\mathcal{D}^{\mathcal{H}*}_{jmt})M^{\mathcal{H}}_{mt}S^{\mathcal{H}}_{jmt}}{\left(1-S^{\mathcal{H}}_{jmt}\right)\alpha^{\mathcal{H}}} - \frac{(1-\mathcal{X}_{jt}\mathcal{D}^{\mathcal{F}*}_{jmt})M^{\mathcal{F}}_{mt}S^{\mathcal{F}}_{jmt}}{\left(1-S^{\mathcal{F}}_{jmt}\right)\alpha^{\mathcal{F}}}}^{\text{exp profit from loans}} + \overbrace{\frac{M^{\mathcal{I}}_{mt}S^{\mathcal{I}}_{jmt}}{(1-S^{\mathcal{I}}_{jmt})\alpha^{\mathcal{I}}} + \frac{M^{\mathcal{N}}_{mt}S^{\mathcal{N}}_{jmt}}{(1-S^{\mathcal{N}}_{jmt})\alpha^{\mathcal{N}}}}^{\text{exp profit from deposits}} \\ & - \underbrace{M^{\mathcal{B}}_{jmt}\left(P^{\mathcal{B}}_{jmt} - C^{\mathcal{C}}_{jmt} - P^{\mathcal{C}}_{t}\right)}_{\text{net exogenous financing cost}} = \frac{F+r}{1+r}\sigma_{jmt}M^{*}_{mt}S^{*}_{jmt}\left[\Phi^{-1}(1-F_{jmt}) + \frac{\phi\left[\Phi^{-1}(1-F_{jmt})\right]}{(1-F_{jmt})}\right].\end{aligned} \tag{27}$$

which determines the standard deviation σ_{jmt} of the unobserved costs C_{jmt}.

Welfare Analysis

Our model has three utility maximizing agents: depositors, borrowers and banks. Both depositors and borrowers have standard linear indirect utillty functions, as defined in (1.1) and (1.3), thus, one can express the welfare of the two agents (in US dollars) respectively as:

$$\begin{aligned} CS_{mt} = \frac{M^{\mathcal{I}}_{mt}}{|\alpha^{\mathcal{I}}|} \ln \left(\sum_i \exp\left(\alpha^{\mathcal{I}} P^{\mathcal{I}}_{imt} + \delta^{\mathcal{I}}_i + \zeta^{\mathcal{I}}_{mt} + \xi^{\mathcal{I}}_{imt}\right) + 1 \right) + \\ \frac{M^{\mathcal{N}}_{mt}}{|\alpha^{\mathcal{N}}|} \ln \left(\sum_j \exp\left(\alpha^{\mathcal{N}} P^{\mathcal{N}}_{jmt} + \gamma^{\mathcal{N}} F_{jmt} + \delta^{\mathcal{N}}_j + \zeta^{\mathcal{N}}_{mt} + \xi^{\mathcal{N}}_{jmt}\right) + 1 \right), \end{aligned} \tag{28}$$

$$\begin{aligned} BS_{mt} = \frac{M^{\mathcal{H}}_{mt}}{|\alpha^{\mathcal{H}}|} \ln \left(\sum_i \exp\left(\alpha^{\mathcal{H}} P^{\mathcal{H}}_{imt} + \delta^{\mathcal{H}}_i + \zeta^{\mathcal{H}}_{mt} + \xi^{\mathcal{H}}_{imt}\right) + 1 \right) + \\ \frac{M^{\mathcal{F}}_{mt}}{|\alpha^{\mathcal{F}}|} \ln \left(\sum_j \exp\left(\alpha^{\mathcal{F}} P^{\mathcal{F}}_{jmt} + \delta^{\mathcal{F}}_j + \zeta^{\mathcal{F}}_{mt} + \xi^{\mathcal{F}}_{jmt}\right) + 1 \right), \end{aligned} \tag{29}$$

where CS stands for depositors' surplus, and BS represents borrowers' surplus. For simplicity, the formulae exclude the Euler-Mascheroni constant, which drops out when we compute changes in welfare. We normalize the utility of the outside option to 0, which justifies the addition of 1 in the expressions.

We measure banks' welfare in terms of their annualized equity value as follows:

$$AEV_{mt} = r \sum_b E_b, \tag{30}$$

where the equity value for each bank E_b is backed out from the default condition (1.11).

When a bank defaults, only a fraction of its assets can be recovered, and the remaining costs are borne by depositors[21] Thus, the expected costs of deposit insurance can be expressed as:

$$EIC_{mt} = 0.6 M^{\mathcal{I}}_{mt} \sum_b F_{bmt} s^{\mathcal{I}}_{bmt}. \tag{31}$$

In the event of bankruptcy not only the insured depositors incur losses, as there may be negative externalities that damage the rest of the economy. We proxy that by introducing a 20% bankruptcy costs, defined as:

$$EBC_{mt} = 0.2 \sum_b F_{bmt} (M^{\mathcal{I}}_{mt} s^{\mathcal{I}}_{bmt} + M^{\mathcal{N}}_{mt} s^{\mathcal{N}}_{bmt}). \tag{32}$$

Finally, we compute the change in total welfare as:

$$\Delta W_{mt} = \Delta CS_{mt} + \Delta BS_{mt} + \Delta AEV_{mt} + \Delta EIC_{mt} + \Delta EBC_{mt}. \tag{33}$$

[21] We assume a 40% recovery rate, in line with Egan et al. (2017).

Note that we do not include central bank's profits in this expression because we do not explicitly model the central bank's objective function.

Table A.IV: Impact of Policy Rate on Welfare Components

The table presents the welfare analysis across different counterfactual policy rates. The unit of observation is a country-year-equilibrium combination. All dependent variables are measured in billions of euros. Banks' franchise value is defined as the annualized equity value net of deposit insurance costs and bank's bankruptcy costs. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

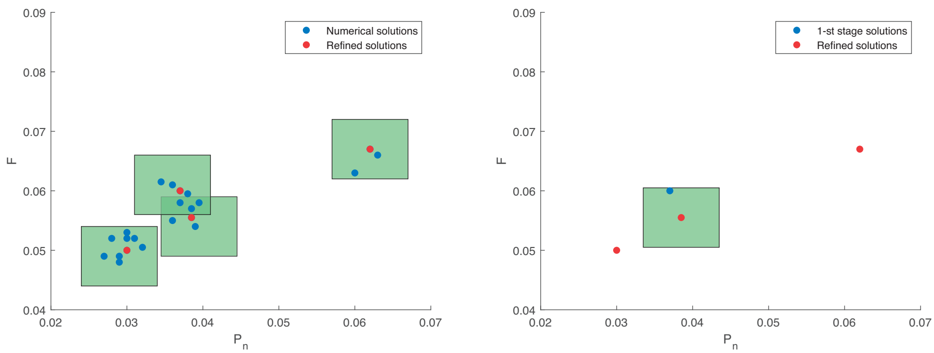
	Depositors' Surplus	**Borrowers' Surplus**	**Banks' Franchise Value**	**Total Welfare**
Policy Rate	23.72***	-0.12	-45.54***	-21.94***
	(3.49)	(0.55)	(8.62)	(4.89)
Country-Year FE	Yes	Yes	Yes	Yes
Observations	5,755	5,755	5,755	5,755
R-squared	0.251	0.117	0.296	0.374

1.C Equilibrium Refinement

As in Egan et al. (2017), we numerically estimate all equilibria. A potential equilibrium is a set of bank-specific interest rates and default probabilities, $\{P^{\mathcal{N}}, P^{\mathcal{I}}, P^{\mathcal{F}}, P^{\mathcal{H}}, F\}$, that satisfies numerical convergence criteria. We look for these equilibria over a large grid of starting values that scales with the number of banks (as the number of equilibria tends to grow combinatorially). Since our estimation procedure is numerical, we often obtain many solutions that are close to each other and potentially capture the same equilibrium. To refine these solutions, we implement similar criteria to Egan et al. (2017). In particular, we cluster all solutions based on their maximum difference. We form a cluster if all the differences across interest rates and default probabilities lie within 0.5%. Within each cluster, we select the unique solution with the smallest objective function value (i.e., function gradient). In practice, these clusters can overlap, so we iterate this procedure twice to remove all remaining solutions that are close to each other. All the solutions that survive this refinement we call model equilibria.

A simplified example of the refinement procedure is depicted in Figure A.IX.

Figure A.IX: Example of equilibrium refinement



(a) 1-st stage **(b) 2-nd stage**

Note: The graphs illustrate a hypothetical refinement procedure based on the smallest objective function value. For brevity, the example only considers two equilibrium variables - uninsured interest rates, $P^{\mathcal{N}}$, and default probabilities, F.

1.D Toy Model

To illustrate the source of multiplicity, we consider an economy with one commercial bank. This bank collects deposits and issues loans. Similarly to Egan et al. (2017), we will abstract from the pricing of loans and introduce an exogenous stochastic return on loans, R^L, that follows a normal distribution with mean μ and volatility σ. To see the connection to our baseline model:

$$R^L \equiv \left(P^L - C\right) \qquad \sim N(\mu, \sigma) \tag{34}$$

where P^L is the loan rate and C is the stochastic cost component. Effectively, we assume that there is a large supply of borrowers that are always willing to borrow at predetermined rates. The bank relies on uninsured deposits to finance the loans. These depositors care about the bank's riskiness and form their exogenous beliefs about its default probability F. The bank competes by offering a deposit rate P^D, and its deposit demand is characterized by logit demand. Namely, the deposit market share is:

$$s^D = \frac{\exp(\alpha^D P^D - \gamma^D F + \delta)}{1 + \exp(\alpha^L P^L - \gamma^D F + \delta)} \tag{35}$$

where α^D is the price sensitivity of depositors, γ^D is the default risk sensitivity, and δ is the utility of borrowing from the bank relative to the outside option. The realized one-period profits can be expressed as:[22]

$$\Pi = s^D \left[R^L - P^D\right] \tag{36}$$

The model is static with an infinite horizon. Since the bank operates under limited liability, the equilibrium equity value of the bank can be expressed as:

$$E = (1-F)\left(\mathbb{E}\left[\tilde{\Pi}|\text{ no default}\right] + \frac{1}{1+r}E\right) = \underbrace{\frac{1+r}{F+r}}_{\text{perpetuity term}} \underbrace{(1-F)\,\mathbb{E}\left[\tilde{\Pi}|\text{ no default}\right]}_{\text{expected profits}} \tag{37}$$

where

$$\mathbb{E}\left[\tilde{\Pi}|\text{ no default}\right] = s^D\left[\mu + \sigma\lambda^R - \tilde{P}^D\right]$$

$$\lambda^R = \frac{\phi(\frac{\overline{R}^L - \mu}{\sigma})}{1 - \Phi(\frac{\overline{R}^L - \mu}{\sigma})} = \frac{\phi(\Phi^{-1}(F))}{1-F}$$

here $\tilde{\Pi}$ and $\tilde{P}^D$ are equilibrium profits and deposit rates, respectively, and λ^R is the inverse Mills ratio evaluated at the (standardized) minimum acceptable loan return, $\overline{R}^L$. As we see from equation (37), the value of the bank depends on default probability F in two key ways. First, it does affect the likelihood of reaching future periods, which is captured

[22] For simplicity, the deposit market size, M^D, is normalized to 1. In addition, we flip the sign of γ^D in line with the notation used in Egan et al. (2017).

by the inter-temporal perpetuity term. Second, it does impact each period's expected profits. These two forces are at the heart of equilibria multiplicity.

Optimization

Since the model is static and beliefs about the bank's default probabilities are exogenous, the bank's optimization simplifies to:

$$E = \max_{P^D} \int_{\overline{R}^L}^{\infty} \left[\Pi + \frac{1}{1+r}E\right] d\omega(R^L) \equiv \max_{P^D} s^D \left[\mu + \sigma\lambda^R - P^D + \frac{1}{1+r}E\right](1-F) \tag{38}$$

Deposit rate FOC:

$$\mu + \sigma\lambda^R - \frac{1}{\alpha^D(1-s^D)} - P^D = 0 \tag{39}$$

Default condition:

$$\begin{aligned} -\mathbb{E}\left[\Pi | R^L = \overline{R}^L\right] &= \frac{1}{1+r}E \\ -s^D\left(\overline{R}^L - P^D\right) &= \frac{1-F}{1+r}s^D\left(\mu + \sigma\lambda^R - \overline{R}^L\right) \\ s^D\left(\mu + \sigma\lambda^R - P^D\right) &= \frac{F+r}{1+r}s^D\left(\mu + \sigma\lambda^R - \overline{R}^L\right) \\ \underbrace{\frac{1+r}{F+r}\mathbb{E}\left[\Pi | \text{ no default}\right]}_{\text{PV of profits}} &= \underbrace{s^D\left(\mu + \sigma\lambda^R - \overline{R}^L\right)}_{\text{max shortfall}} \end{aligned} \tag{40}$$

we can rewrite this condition in terms of F and optimal P^D:

$$\frac{1+r}{F+r}\left(\frac{1}{\alpha^D(1-s^D)}\right) = \sigma\left(\lambda^R - \Phi^{-1}(F)\right) \tag{41}$$

Deposit rates

The deposit FOC establishes a positive relationship between a bank's riskiness and optimal deposit rates. The result also implies that the same pattern holds across model equilibria. To prove this more formally, note that there exists a unique deposit rate P^D that optimizes equity value for each given F. This comes from the continuity and strict monotonicity of the deposit rate FOC in P^D, and the fact that the FOC turns positive (negative) for sufficiently small (large) values of P^D.[23] It is also possible to show this unique P^D is increasing in F, meaning that higher levels of risk are always accompanied by higher deposit rates.[24] There are two economic channels explaining that. First, it is optimal

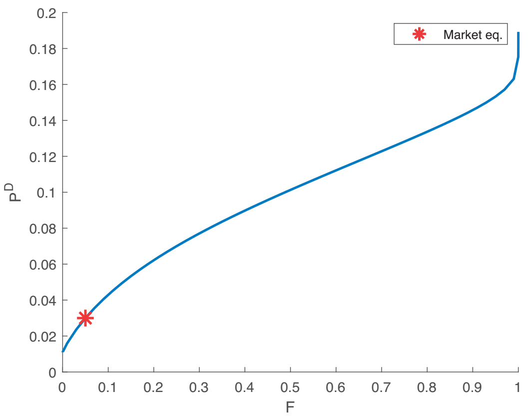
[23] In other words, if we allow for arbitrarily negative deposit rates, the deposit rate FOC satisfies the Intermediate Value Theorem conditions. If the deposit rates are constrained, the Lagrange multiplier of the constraint will ensure this condition.

[24] The proof is stated in the mathematical appendix. If the markups are non-monotonically dependent on interest rates, this conclusion can change. For instance, if higher interest rates translate to the changing profile of depositors or loan borrowers, then there can be multiple price levels satisfying the first-order

for the bank to pass limited liability benefits, $\sigma\lambda^R$, on depositors to capture a larger market share and increase the markups. This stems from the assumption that stochastic revenue affects every loan rather than a fixed amount (as in Hortaçsu et al. (2011)). Second, the bank has to compensate for the depositor's aversion towards its riskiness to sustain its share. For moderate levels of γ^D, the first channel is dominant. A similar monotonic relationship also arises in settings with multiple banks and loan pricing. The only exception is that the risk-taking with loans would result in a negative relationship between loan rates and the bank's risk.

Figure A.X illustrates the connection between a bank's riskiness and deposit rates for one set of parameters.

Figure A.X: Relationship between deposit rates and default probabilities



Note: The graph presents the relationship between deposit rates, P^D and bank's default probabilities F based on the deposit rate FOC. To calibrate the model, we assume that the market equilibrium has $P^D = 0.03$, $F = 0.05$, $S^D = 0.4$, $r = 0.05$, $\gamma^D = 2$, $\alpha^D = 40$.

Default boundary

The default condition guarantees that the fundamental riskiness of the bank, dictated by its profitability and markups, matches depositors' beliefs. Alternatively, equation (40) can also be considered as an incentive compatibility constraint that evaluates the marginal benefit of increasing the default probability after R^L shock realizes. On the one hand, a marginal increase in F translates to losing some of the future conditional profits because the bank lowers its survival probability. On the other hand, an increase in F decreases the (ex-post) maximum loss that the bank has to incur once the shock materializes (i.e., invoking limited liability protection). At the optimum, these two effects should balance out. The marginal increase in F here completely ignores the indirect effects on deposit

conditions for a given level of risk.

market shares. This is the case because the decision to default happens ex-post when the market shares are already determined.

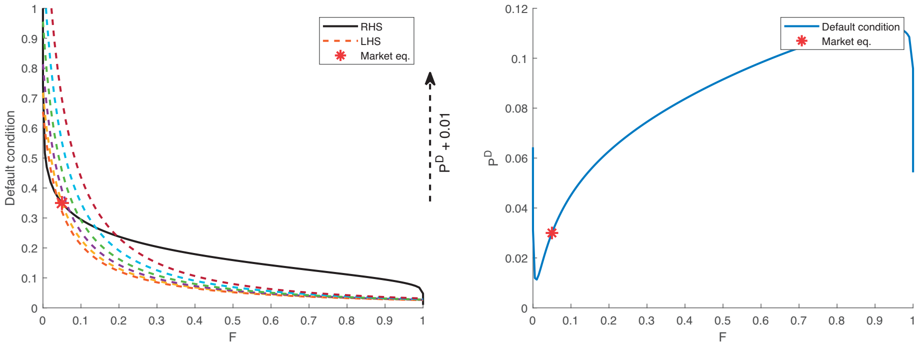
The source of multiplicity stems from the limited liability of banks. If banks had unlimited liability, the default condition would simplify to the participation constraint. Namely, if the bank, on average, earns more on loans than it pays to depositors (i.e., $\mu > P^D$), then the bank never defaults since the current period losses are sunk, and they do not influence future profitability. Otherwise, the bank defaults and leaves the market. In this sense, the model collapses to standard industrial organization models that display a unique equilibrium pinned down by the first-order condition of deposit interest rates (e.g., Berry et al. (1995)). In contrast, limited liability introduces non-linearities in the model and potentially several local optima that can be sustained by the depositors' self-fulfilling beliefs.

To see it more formally, notice the default condition is non-monotonic in the bank's default probability. For illustrative purposes, Figure A.XI (left side) plots the LHS and RHS of equation (41) for different levels of deposit rates. Each intersection point represents the combination of F and P^D such that the bank has no incentives to marginally alter its default probability (the right graph displays all of these combinations).[25] In particular, to sustain extremely small levels of default risk in equilibrium, the bank needs to be highly profitable (LHS), or else it cannot cover its maximal losses of default (RHS). Thus, very low levels of risk are associated with high mark-ups and deposit rates.[26] However, as the default risk increases, the reduction in maximal loss is quickly outweighed by the decreasing future survival probabilities and losses of future profits. This trade-off introduces non-monotonicities, and that is the reason why the curve becomes upward-sloping in the middle section of the graph. Similarly, for extremely high levels of default risk, the limited liability benefits become dominant again, and the curve converges downward towards 0. All of these inflection points can potentially generate a local optimum.

[25] The mathematical appendix shows that the crossing point always exists.

[26] In fact, the bank would never have a 0 default probability unless the mark-ups become infinitely large.

Figure A.XI: Relationship between deposit rates and default probabilities



(a) (b)

Note: The graphs present the relationship between deposit rates P^D and the bank's default probabilities F based on the default condition. The graph on the left depicts the LHS and RHS of equation (41) for different values of P^D, whereas the right graph displays all combinations of F and P^D satisfying the default condition. To calibrate the model, we assume that the market equilibrium has $P^D = 0.03$, $F = 0.05$, $S^D = 0.5$, $r = 0.05$, $\gamma^D = 2$, $\alpha^D = 40$.

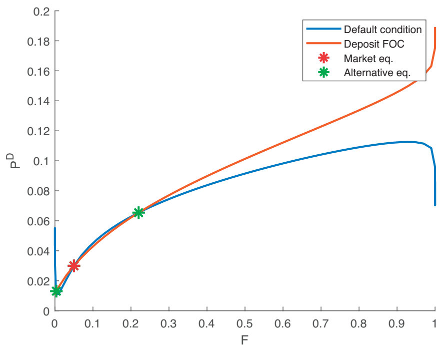
It is worth mentioning that even though this version of the model does not have shareholders' equity, the bank is not entirely a zero-equity entity. As the default condition stipulates, the bank's equity holders are willing to inject capital and cover losses as long as they do not exceed the present value of future equity value. In that sense, the bank has implicitly committed equity which is not directly invested in the bank's productive activities. The key distinction here is that the bank does not have to lose anything additionally if the losses exceed the default threshold. Thus, from this perspective, equity is contingent.

Equilibria multiplicity

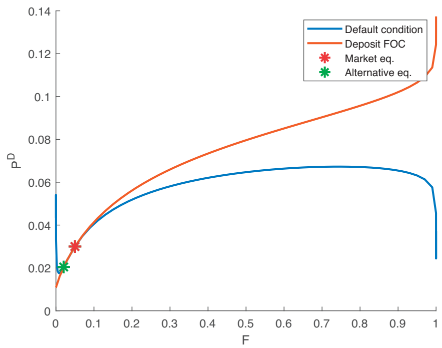
The equilibria encompass all the points where the deposit rate FOC and default condition curves intersect. In this toy example setting, there is always at least one such point. The proof relies on the fact that the deposit rate FOC increases from $\underline{P}^D$ to ∞ when F moves from 0 to 1, where $\underline{P}^D$ is some finite constant. At the same time, the default condition declines from ∞ to 0. Since both curves are continuous, at some point, they have to cross each other. The more challenging question is determining the number of equilibria in this model. Figure A.XII displays three different parametrizations that lead to one, two, and three equilibria. It turns out that the initial size of the bank, driven primarily by the bank-specific taste parameter δ, is crucial in influencing the numerosity of equilibria. Namely, a larger bank is generally far less susceptible to changes in interest rates and default probabilities. The second substantial driver is related to the elasticity to the bank's riskiness, γ^D. When this parameter is smaller, the bank's market shares are more stable, potentially leading to fewer adverse scenarios. A more thorough sensitivity analysis is conducted in the following section.

The maximal number of model equilibria for a single-bank economy is three. Although the rigorous proof of this result is beyond the scope of this analysis, various simulations display that the number of equilibria tends to coincide with the number of non-monotonicity points in the default condition (i.e., as the deposit rate FOC is strictly monotonous in P^D and F). This also suggests that the number of equilibria in a multi-bank economy is potentially bounded by 3^N, where N is the number of banks. In other words, the multiplicity of equilibria tends to grow combinatorially with the number of banks in the economy. Intuitively, the numerosity of equilibria can change once we introduce other non-linearities in the model, such as fixed costs and loan borrower defaults. Yet, theoretically analyzing these cases becomes less and less tractable.

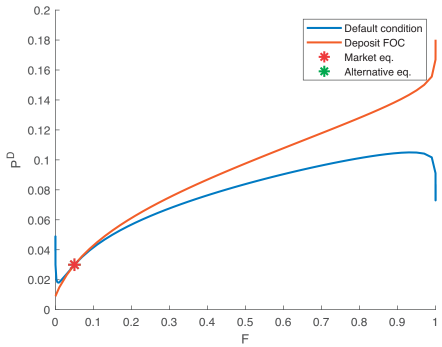
Figure A.XII: Multiplicity of equilibria



(a) $s^D = 0.4, \quad \gamma = 2$



(b) $s^D = 0.5, \quad \gamma = 0$



(c) $s^D = 0.6, \quad \gamma = 2$

Note: The graphs display different types of model parametrizations that lead to one, two, or three equilibria. To calibrate the model, we assume that the market equilibrium has $P^D = 0.03$, $F = 0.05$, $S^D = 0.5$, $r = 0.05$, $\gamma^D = 2$, $\alpha^D = 40$, with the exception of perturbed parameters.

The key feature of the model that sustains the multiplicity of equilibria rests on exogenous depositors' beliefs. We build our model on Egan et al. (2017)'s framework that features self-fulfilling beliefs, as in Diamond & Dybvig (1983). In particular, if depositors expect that the bank is going to default, they will pull their money from it, and that will eventually weaken the fundamentals of the bank. As a result, in our model, we have multiple optima that can be Pareto ranked in terms of the bank's profitability and welfare; however, the bank has no ability to break these bad equilibria.[27] Similarly to Spence (1973)'s signaling game, bad equilibria exist because banks do not know the off-equilibrium beliefs of depositors and their potential competitors' actions. As a result, the banks fear possible bank runs or retaliation if they take an off-equilibrium action.

Sensitivity to parameters

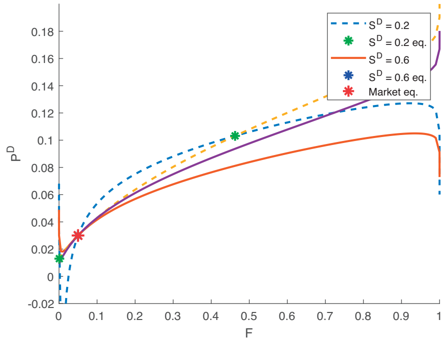
The severity of alternative equilibria does depend on model parameters, such as depositors' interest rate and risk elasticities, as well as changes in bank-specific taste parameter δ and the risk-free rate, r. Figure A.XIII displays the sensitivities to these key parameters. It appears that the size of the bank is a vital dimension that contributes both to the multiplicity and extremeness of alternative equilibria. When the bank is very large (thus, large δ), the high-risk equilibrium is unlikely to realize as the depositors have a strong preference for the bank regardless of its interest rates and default probabilities. As a result, the bank can absorb massive negative shocks. In contrast, an economy populated with small banks (smaller δ's) is much more fragile and tends to have a possibility of bank-run equilibria (as well as highly safe equilibria).

On the other hand, the risk-free rate parameter, r, does not seem to influence the outcomes much. Lower r is associated with more extreme alternative equilibria - the safest equilibrium tends to have lower deposit rates, whereas the risky one is less stable with higher rates. The economic mechanism here is linked to the calibrated σ parameter governing the volatility of stochastic costs. Ceteris paribus, when r is smaller, the bank discounts its future profits less. Thus, for any given default probability in a market equilibrium, the volatility has to increase. Consequently, that tightens the relationship between deposit rates and bank's risk in alternative equilibria.

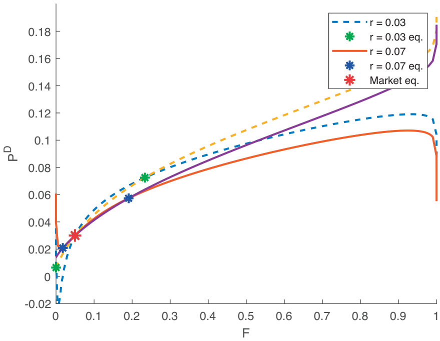
Two key parameters in the model are price and default probability elasticities. We can notice that lower deposit rate sensitivities lead to significantly more responsive deposit rates to changes in default probabilities. Intuitively, when α^D is lower, the bank must act much more aggressively since the depositors do not react as much to changes in prices. Similarly, when γ increases, the bank needs to adjust the rates more dramatically to preserve market shares and pricing power when the risk rises. In both cases, we have an economy with a more severe alternative bank-run equilibrium.

[27] Adding more information structure to the model can lead to the uniqueness of equilibrium, but that is beyond the scope of this paper.

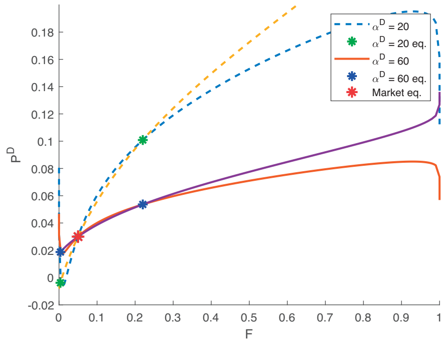
Figure A.XIII: Parameter sensitivity



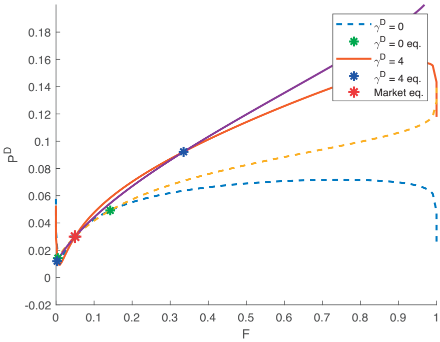
(a) Sensitivity to $s^D(\delta)$



(b) Sensitivity to r



(c) Sensitivity to α^D



(d) Sensitivity to γ^D

Note: The graphs above present the sensitivity of alternative equilibria to model parameters, i.e., δ, r, α^D, γ^D. To calibrate the model, we assume that the market equilibrium has $P^D = 0.03$, $F = 0.05$, $S^D = 0.5$, $r = 0.05$, $\gamma^D = 2$, $\alpha^D = 40$, with the exception of perturbed parameters.

1.E Mathematical Appendix

Deposit rate FOC

Deposit rate FOC wrt P^D

Let's take the derivative of equation (39) wrt P^D:

$$\frac{\partial FOC}{\partial P^D} = -\frac{\alpha^D s^D(1-s^D)}{\alpha^D(1-s^D)^2} - 1 = -\frac{s^D}{1-s^D} - 1 < 0$$

Here we use the result that $\frac{\partial s^D}{\partial P^D} = \alpha^D s^D(1-s^D)$ and that market shares are bounded between 0 and 1. This proves the monotonicity of the deposit rate FOC in P^D.

Deposit rate FOC wrt F

Now, let's take the derivative of equation (39) wrt F:

$$\frac{\partial FOC}{\partial F} = \sigma\frac{\partial \lambda^R}{\partial F} - \frac{\partial\left(\frac{1}{\alpha^D(1-s^D)}\right)}{\partial F} = \sigma\underbrace{\frac{\partial \lambda^R}{\partial F}}_{>0} + \underbrace{\frac{\gamma^D s^D}{\alpha^D(1-s^D)}}_{>0} > 0$$

for $F \in (0,1)$. It is trivial to show that the second term is always positive. The only question remains if the first term is positive too. Note that we can apply the following substitution:

$$\lambda^R = \frac{\phi(\Phi^{-1}(F))}{1-F} = \frac{\phi(x)}{1-\Phi(x)} \qquad \text{, where } x \equiv \Phi^{-1}(F)$$

This implies that:

$$\frac{\partial \lambda^R}{\partial F} = \frac{\partial \lambda^R}{\partial x}\frac{\partial x}{\partial F} = \left(\frac{-x\phi(x)\left(1-\Phi(x)\right)+\phi(x)^2}{\left(1-\Phi(x)\right)^2}\right)\frac{\partial x}{\partial F} = \frac{\phi(x)}{(1-\Phi(x))}\left(-x+\frac{\phi(x)}{(1-\Phi(x))}\right)\frac{\partial x}{\partial F}$$

Given that:

$$\frac{\partial x}{\partial F} = \frac{\partial \Phi^{-1}(F)}{\partial F} = \left(\frac{\partial \Phi(x)}{\partial x}\right)^{-1} = \frac{1}{\phi(x)}$$

We arrive at:

$$\frac{\partial \lambda^R}{\partial F} = \frac{1}{1-F}\left(-\Phi^{-1}(F)+\lambda^R\right)$$

This derivative is positive for $F \in (0,1)$. To prove that, it suffices to show that:

$$-x + \frac{\phi(x)}{(1-\Phi(x))} > 0 \qquad \text{, for } x \in \mathbb{R}$$

Note that for $x \leq 0$ this inequality is immediate. Thus, the only non-obvious case to consider is when $x > 0$. We know that the complementary cumulative distribution function has some bounds:

$$1 - \Phi(x) < \frac{\phi(x)}{x} \qquad , \text{for } x > 0$$

This follows from the integration by parts:

$$1 - \Phi(x) = \int_x^{\infty} \phi(u)du = -\int_x^{\infty} \frac{\phi'(u)}{u}du = \frac{\phi(x)}{x} + \int_x^{\infty} \frac{\phi'(u)}{u^3}du = \frac{\phi(x)}{x} - \underbrace{\int_x^{\infty} \frac{\phi(u)}{u^2}du}_{>0}$$

As a result:

$$\frac{\phi(x)}{1 - \Phi(x)} > x$$

which means that $\frac{\partial \lambda^R}{\partial F} > 0$.

Relationship between P^D and F

Since $\frac{\partial FOC}{\partial P^D} < 0$ and $\frac{\partial FOC}{\partial F} > 0$, we get that P^D is increasing in F. This follows from the Implicit Function Theorem:

$$\frac{\partial P^D}{\partial F} = -\frac{\frac{\partial FOC}{\partial F}}{\frac{\partial FOC}{\partial P^D}} > 0.$$

Default condition

Derivation

The step-by-step derivation of the default condition are presented below:

$$\begin{aligned}
-\mathbb{E}\left[\Pi | R^L = \overline{R}^L\right] &= \frac{1}{1+r}E \\
-\mathbb{E}\left[\Pi | R^L = \overline{R}^L\right] &= \frac{1-F}{1+r}\left(\mathbb{E}\left[\Pi | R^L > \overline{R}^L\right] + \frac{1}{1+r}E\right) \\
-\mathbb{E}\left[\Pi | R^L = \overline{R}^L\right] &= \frac{1-F}{1+r}\left(\mathbb{E}\left[\Pi | R^L > \overline{R}^L\right] - \mathbb{E}\left[\Pi | R^L = \overline{R}^L\right]\right) \\
-s^D\left(\overline{R}^L - P^D\right) &= \frac{1-F}{1+r}s^D\left(\mu + \sigma\lambda^R - \overline{R}^L\right) \\
s^D\left(\mu + \sigma\lambda^R - P^D\right) &= \frac{F+r}{1+r}s^D\left(\mu + \sigma\lambda^R - \overline{R}^L\right) \\
\underbrace{\frac{1+r}{F+r}\mathbb{E}\left[\Pi | \text{ no default}\right]}_{\text{PV of profits}} &= \underbrace{s^D\left(\mu + \sigma\lambda^R - \overline{R}^L\right)}_{\text{max unexpected shortfall}}
\end{aligned}$$

Limiting cases

The RHS of equation (40) has the following limits:

$$\lim_{F\to 0}\left(\frac{\phi(\Phi^{-1}(F))}{1-F}-\Phi^{-1}(F)\right)=\lim_{F\to 0}\left(0-\Phi^{-1}(F)\right)=\infty$$
$$\lim_{F\to 1}\left(\frac{\phi(\Phi^{-1}(F))}{1-F}-\Phi^{-1}(F)\right)=\lim_{F\to 1}\left(\frac{\phi(\Phi^{-1}(F))-(1-F)\,\Phi^{-1}(F)}{1-F}\right)$$

For a normal distribution, $\lim_{x\to\infty}(1-\Phi(x))\,x=0$. This implies that we can apply L'Hôpital's rule to the limit above:

$$\lim_{F\to 1}\left(\frac{\phi(\Phi^{-1}(F))-(1-F)\,\Phi^{-1}(F)}{1-F}\right)\overbrace{=}^{x=\Phi^{-1}(F)}\lim_{x\to\infty}\frac{\phi(x)-(1-\Phi(x))\,x}{1-\Phi(x)}=$$
$$\lim_{x\to\infty}\frac{-x\phi(x)+x\phi(x)-(1-\Phi(x))}{\phi(x)}=\lim_{x\to\infty}\frac{-(1-\Phi(x))}{\phi(x)}=\lim_{x\to\infty}\frac{\phi(x)}{-x\phi(x)}=0$$

Existence of solution

It is possible to show that there exists a unique level of P^D for each F that satisfies the default condition. To illustrate this, move all terms related to F in equation (41) to the RHS and take limits of that:

$$\frac{1}{\alpha^D(1-s^D)}=\frac{F+r}{1+r}\sigma\left(\lambda^R-\Phi^{-1}(F)\right)$$

$$\lim_{F\to 1}(F+r)\left(\frac{\phi(\Phi^{-1}(F))}{1-F}-\Phi^{-1}(F)\right)=0$$
$$\lim_{F\to 0}(F+r)\left(\frac{\phi(\Phi^{-1}(F))}{1-F}-\Phi^{-1}(F)\right)=\infty$$

Since the mark-ups are always strictly positive and bounded, there is always a crossing point between the LHS and RHS. Moreover, since the markups are strictly monotonic in P^D, this implies that this intersection point is unique. In fact, one can even solve algebraically for P^D:

$$P^D=\mu+\sigma\lambda^R-\frac{F+r}{1+r}\sigma\left(\lambda^R-\Phi^{-1}(F)\right)$$

Alternatively:

$$P^D=\frac{1}{\alpha^D}\left(\ln\left(\alpha^D\frac{F+r}{1+r}\sigma\left(\lambda^R-\Phi^{-1}(F)\right)-1\right)+\gamma^D F-\delta\right)$$

Note that the proof hinges on the positivity of markups. If the model had fixed costs, there could be scenarios where this default condition is never satisfied.

Chapter 2

The Term Structure of Corporate Bond Risk Premia*

*I would like to cordially thank my supervisors, Joost Driessen and Lieven Baele, for all the support during this project. I am also grateful to Stefano Giglio, Frank de Jong, Julio Crego, Daniel Karpati, Zhaneta Tancheva, and other participants at Tilburg University and Yale School of Management seminars, and at the 2022 FMA Annual Meeting. Lastly, I would like to thank CentER for the financial support during this period.

2.1 Introduction

Understanding how long- and short-term projects are priced is one of the fundamental questions in finance. The term structure of risk premia allows us to perform NPV calculations, test asset pricing models, and potentially explain the source of many cross-sectional anomalies. This paper constructs a forward-looking implied estimate of the term structure of risk premia in the corporate bond market. A fixed schedule of promised cash flow payments makes the bond market an ideal environment for recovering the risk premia term structure for a wide range of maturities. While there have been recent attempts to measure this term structure in equities, the corporate bond market is left largely unexplored.[1] This paper fills in the gap.

The pricing of corporate debt carries substantial importance to firms and investors. As of 2021, the US corporate bond market stands at roughly $15 trillion, making it the second largest component of a public firm's balance sheet. Adding to that, corporate bonds remain one of the primary sources of new capital for many firms: the annual issuance of public debt ($1.45 trillion) over the last decade has far exceeded the issuance of equity ($252 billion).[2] For these reasons, bond prices and trading potentially carry a great deal of information about the expected returns on firm's assets and its cost of financing.

Analyzing corporate bonds allows us to deepen the understanding of institutional investor risk preferences, which are often at the heart of the equity term structure debate. The corporate bond market offers much more capacity to capture the average institutional investor preferences as less than 10% of the bonds are owned by households (this number reaches nearly 40% for equities). This distinction can be crucial since many explanations behind the duration term premium originate from the need to hedge long-term investment risk and pension liabilities (Gormsen (2021); Gonçalves (2021)), and such motives should be easily detectable in corporate bond markets. Moreover, rich cross-sectional variation in bond characteristics allows to control for clientele effects as the institutional investors display a great of heterogeneity in terms of their regulatory or funding constraints (Bali, Subrahmanyam & Wen (2021)). Such patterns can be tested while controlling, for instance, for the corporate bond credit rating.

The studies focusing on equity markets face a number of challenges ranging from the relatively short historical samples to the difficulty of pinning down the exact duration of a risky asset. In my paper, I address these hurdles by looking at the yields in the corporate bond market. The forward-looking nature of yields combined with the rich literature on expected default probabilities (Campbell, Hilscher & Szilagyi (2008); van Zundert & Driessen (2017); Feldhütter & Schaefer (2018)) allows extracting expected returns that

[1] The empirical measurement of the equity term structure can be found in van Binsbergen, Brandt & Koijen (2012), van Binsbergen, Hueskes, Koijen & Vrugt (2013), van Binsbergen & Koijen (2017), Weber (2018), Gormsen (2021), among many others.

[2] Source: Table L.213 and L.223 in the Federal Reserve Board Z1 Flow of Funds, Balance Sheets, and Integrated Macroeconomic Accounts, as of the fourth quarter of 2020. The bond issuance data is provided by SIFMA and Refinitiv.

are not reliant on historical price information. This offers powerful empirical advantages as in short historical samples, such as the last 20 years, the realized returns may be driven by a few extreme periods of recessions (Bansal, Miller, Song & Yaron (2021)), structural shifts in the risk-free rate (van Binsbergen & Schwert (2021)) or time-variation in risk premia (this paper) and, therefore, deliver biased estimates of short and long duration returns. Furthermore, the fixed maturity of promised cash flows enables me to estimate both the aggregate term structure and the cross-sectional patterns without the need to model asset duration separately.

The first empirical finding is that the implied term structure of risk premia is upward-sloping at the aggregate market level and in various cross-sectional sorts. The results are obtained using two approaches to estimate default probabilities: Merton's style models of default, such as in Feldhütter & Schaefer (2018), and historical default frequencies from Moody's Investors Service (2021). Based on the structural models, the long-short duration portfolios earn roughly 1.3% more, whereas the estimates based on historical probabilities deliver much more conservative estimates of roughly 0.2%.[3] Given that the former method is sensitive to recessions (as it reacts to the market values of leverage) while the latter is devoid of time-variation, the two approaches establish reasonable upper and lower bounds for the slope of the risk premium term structure. Even though the returns may seem small, the risk premia slope can be economically substantial as it constitutes up to 20% of the long-term bond yields and up to 30% of total expected returns.

The flexibility of structural models allows us to investigate various cross-sectional patterns. The upward-sloping term structures of risk premia are widely detected among credit rating, leverage, issuer's size, and book-to-market ratio sorts. An exception is present among the lowest credit rating firms that exhibit no slope, a pattern that might be consistent with the high exposure to short-term risk (such as crash risk). Interestingly, the expected return proxies detect a sizeable credit risk premium of 2-3%, measured as the difference between speculative and highest investment grade bond returns, as well as some support for size and value factors in corporate bonds (in line with Bai, Bali & Wen (2019) and Bartram, Grinblatt & Nozawa (2021)). Yet, most of this risk premium is generated by the shortest-term bonds and slowly dissipates with longer duration. These dynamics are consistent with the firm cycle theories where firms over time, if they survive, converge to similar steady states and risk exposures. This observation is indirectly supported by Yara, Boons & Tamoni (2020) who find narrowing factor premia over longer holding period horizons.

In contrast, the realized returns over the last 20 years paint a very different picture of the term structure of risk premia. The average excess returns deliver a negative but statistically insignificant slope of 1-2%. The negative slope is predominantly driven by exceptionally high short-term duration asset returns. These shortest-duration portfolios are also the only portfolios that yield statistically different estimates from the structural model predictions. However, once I adjust returns by controlling for the market-wide bond

[3]To put it in perspective, the duration gap between the long and short legs of duration portfolios is roughly 12 years.

portfolio return, I find that the risk-adjusted returns display a strong negative relation to duration. The cross-sectional sorts do not alter these conclusions, in line with the work of Gormsen & Lazarus (2019). While the realized returns are not entirely comparable to the holding-to-maturity implied returns, with the help of an affine model, I show that the holding period horizon should not be the vital dimension creating the gap with the implied measures. In this aspect, the implied expected returns contradict the observed trends in the realized returns.

A great deal of this discrepancy is explained by variation in the risk premia slope. Using a parsimonious affine term structure model, I quantify the time variation in short-maturity corporate bond risk premium, and I show that it has been declining over the last 20 years. Even though a one-factor model cannot reconcile all the differences between implied expected and realized returns, the current results confirm that changes in risk premia introduced a substantial bias in realized return measures (up to 30%). My findings mirror the intuition of Fama & French (2002), arguing that the equity risk premium has been trending downwards over the last several decades, creating significant capital gains that confound the measurement of expected returns. In a similar vein, van Binsbergen & Schwert (2021) document that the recent shifts in the US government bond yields have largely affected the cross-section of corporate bond returns. Similarly, I find that the risk premia term structure after the financial crisis steepened, generating positive capital gains on the short-duration assets by up to 0.7%, and suppressing the long-term bond returns by nearly 1.7%. Such structural trends potentially were driven by the drastic changes in the monetary policy; however, these explanations require deeper examination.

This paper contributes to and builds on rapidly growing literature on the equity term structure by shedding new light on corporate bond market valuations. In their seminal work, van Binsbergen et al. (2012) and van Binsbergen et al. (2013) documented abnormally high returns on the short-maturity option-implied dividend assets and dividend strips, a phenomenon I also detect in the corporate bond market returns. However, I discover that the implied measures favor the unconditionally upward (or somewhat flat) term structure of risk premia, in agreement with the leading theoretical asset pricing models (Campbell & Cochrane (1999); Bansal & Yaron (2004); Gabaix (2012)). The gap between the realized and implied proxies of returns is partly driven by the high frequency of recessions, and structural shifts in the following years (in line with Bansal et al. (2021)'s reasoning).

My work also connects to the vast empirical research on the cross-section of corporate bond returns. I document various cross-sectional anomalies, such as credit risk, issuer's size, book-to-market ratios using the implied measures of expected returns (in light of Campello, Chen & Zhang (2008)). In contrast to Gormsen & Lazarus (2019), I find little evidence supporting the claim that duration alone subsumes a wide range of factor premia. On the other hand, most of the factor premia are earned on the short-duration bonds, suggesting that the investors are particularly averse to the short- and intermediate-horizon risks. Interestingly, Giglio, Kelly & Kozak (2020) and Baele, Driessen & Jankauskas (2022) highlighted similar insights into equities.[4]

Lastly, the paper's methodology closely relates to the literature on structural default models, such as the Merton (1974) model. To my knowledge, my paper is the first to construct implied measures of bond risk premia by combining the approaches of Campello et al. (2008) with Feldhütter & Schaefer (2018) that implement the Black-Cox adaptation of the Merton's model. In a number of robustness tests, I also examine alternative specifications with stochastic volatility (van Zundert & Driessen (2017)) and the original Merton's model. Qualitatively, these modifications do not alter the results much, and that supports Huang & Huang (2012) claims that most structural models deliver similar predictions.

The rest of the paper is structured in five sections. The second section presents the methodology and discusses the advantages of using implied measures over realized returns. The subsequent section highlights the data sources and summary statistics. Later, section IV covers the baseline results with the expected returns implied from yields. Section V introduces the puzzle of realized returns. Finally, section VI attempts to reconcile the puzzle.

2.2 Methodology

The majority of the cross-sectional bond returns literature relies on the realized returns to estimate risk premia.[5] The approach banks on the idea that over historical samples the average return will approach the expectation. Equivalently, the average excess return over the treasury bond will yield an unbiased risk premium estimate. This method has a few key requirements: a long history of data, and no major structural shifts driving the term structure of risk-free rates and risk premia. Yet, as argued before, both of those are difficult to satisfy in the TRACE sample (2002-2020).

To address these challenges, I measure the maturity-specific risk premium extracted from forward-looking yields. Following Campello et al. (2008), one can express the (annualized) expected holding-to-maturity returns as:

$$\mathbb{E}_t(r_{j,t,t+m}) = \left[(1 + y_{j,m,t})^m (1 - L_m \pi_{j,m,t}) \right]^{1/m} = (1 + y_{j,m,t}) (1 - L_m \pi_{j,m,t})^{1/m} \quad (2.1)$$

where $\mathbb{E}_t(r_{j,t,t+m})$ stands for the expected holding-period return of bond j of maturity m at time t, $y_{j,m,t}$ - the corresponding bond yield, L - loss given default, and π denotes the duration-matched default probability. Since the returns here are holding-to-maturity, π measures the cumulative default probability that depends on issuer- and issue-specific

[4]Both this paper and Baele et al. (2022) aim to estimate the risk premia without the usage of realized returns. Qualitatively, the results are similar: the aggregate term structure has a positive slope and rich cross-sectional patterns that tend to coincide in both equity and bond markets. In order to connect both papers, one potentially needs a structural model both for equity and debt. Regarding the quantitative magnitude of estimates, the bond market exhibits much lower levels of risk premia and the slope tends to be flatter. This can be justified by the fact the bonds are generally much safer.

[5]More recent work includes Bai et al. (2019), Bartram et al. (2021), among many others.

information. Intuitively, the equation states that the expected return equals the yield corrected for the expected default losses. The expression is exact for zero coupon bonds with defaults at maturity. For coupon-paying bonds, it serves as an approximation where m is the duration of a bond.[6] It is important to mention that the expected return measure from (2.1), like yields, effectively averages the risk premia over the remaining periods ahead, and only in the case of flat term structure exactly matches the next period expected return. However, it is easy to show that the holding-to-maturity return contains all the necessary information to learn about the shape of the term structure.[7] Thus, for simplicity, that remains my preferred measurement of expected returns.

One can rewrite this expression and subtract the maturity-matched risk-free rate to find the risk premium on the bond:

$$\mathbb{E}_t(rp_{j,m,t+m}) = \underbrace{y_{j,m,t} - y_{g,m,t}}_{\text{credit spread}} - \underbrace{(1 + y_{j,m,t})\left(1 - (1 - L_m \pi_{j,m,t})^{1/m}\right)}_{\text{annualized expected losses}} \tag{2.2}$$

where $\mathbb{E}_t(rp)$ stands for the risk premium, and $y_{g,m,t}$ is the yield on a synthetic government bond with the same cash flows. Effectively, this is the expected return on a CDS contract or a portfolio invested in corporate bonds and shorting the respective maturity government bond. The risk premium stems from the cash flow risk. Since corporate bonds are prone to default, the investors demand compensation for default risk and uncertainty about the fluctuations in default probabilities over time. Note that since the default probabilities are non-negative, the maximum risk-premium is bounded by the size of the credit spread.

The key advantage of this approach is that it relies on forward-looking information rather than just on a history of prices. This comes at the cost of estimating the default probabilities for each maturity. Based on the literature, I proxy such expectations in two ways: 1) with the long-term historical default frequencies from Moody's; 2) by calibrating them based on the Merton (1974) model. Both approaches have their pros and cons. The former relies on historical averages and does not vary over time. Yet, with the issuer's credit rating alone, you can attain the whole term structure of default probabilities. This substantially expands the sample, albeit this method limits the time-variation analysis. The second approach needs more modeling assumptions and inputs (i.e. measurement of the capital structure, default threshold and firm's asset returns). On the upside, the methodology delivers forward-looking default probabilities that account for time-varying and firm-specific characteristics.

The implementation of historical default probabilities is straightforward. Moody's Investor Service every year publishes the historical default frequencies of different credit rating issuers over 1-20-year maturities. As Feldhütter & Schaefer (2018) argue, corporate bankruptcies are highly clustered events, therefore, having a long sample of observations is

[6] In the Appendix 2.E, I show that this approximation serves well to measure scenarios where the coupon-bearing bonds can default early or the reinvestment rate differs from the bond yield.

[7] A similar analogy can be drawn between the par-yield and forward-yield curves.

crucial for the measurement. For this reason, I use historical averages over the 1920-2020 sample period.

Regarding the second approach, the structure of Merton's model allows to project the default probabilities just with a few parameters, such as the asset return and volatility. In this paper, I adopt the Feldhütter & Schaefer (2018) implementation of the Black & Cox (1976) model, which appends Merton's model with one additional assumption of early defaults. As a result, the default probability can be written as:

$$\pi_t = N\left(-\frac{ln(\frac{V_t}{dD_T}) + \tilde{\mu}_{A,t}T}{\sigma_{A,t}\sqrt{T}}\right) + exp\left(\frac{-2ln(\frac{V_t}{dD_T})\tilde{\mu}_{A,t}}{\sigma^2_{A,t}}\right) N\left(\frac{-ln(\frac{V_t}{dD_T}) + \tilde{\mu}_{A,t}T}{\sigma_{A,t}\sqrt{T}}\right), \tag{2.3}$$
$$\tilde{\mu}_{A,t} = \mu_{A,t} - \delta_t - 0.5\sigma^2_{A,t}$$

where V_t is the value of firm's assets, D_T is the face value of debt that matures at T, $\mu_{A,t}$ and $\sigma_{A,t}$ are the expected return and volatility of assets, δ_t is the payout rate to all company's stakeholders, and $N(.)$ is the CDF of standard normal distribution. When the default boundary, d, is equal to 1, the first term is just the cumulative probability that the asset value falls below the face value of debt at maturity, as in the Merton's model. In turn, the second term corrects for the early arrival of defaults. The parameter d rescales the total level of debt to reflect the actual default threshold. In order to determine d, I closely follow Feldhütter & Schaefer (2018) and calibrate it to match the average level of historical default frequencies in each credit rating group. In principle, this means that the method partly relies on past information. Nevertheless, the calibration focuses on the unconditional means only and has limited effect on the cross-sectional, cross-maturity and time variation of defaults.

As we can see from (2.3), the structural estimation requires the knowledge of expected asset returns and volatility (i.e. μ_A and σ_A). Following van Zundert & Driessen (2017) and Feldhütter & Schaefer (2018), I assume that the risk premium on assets is equal to a constant price of risk (θ) times the volatility of the firm's assets ($\sigma_{A,t}$), where $\theta = 0.22$ (based on Chen, Collin-Dufresne & Goldstein (2009) estimates). In other words, this assumption reduces the two-parameter model to a single-parameter specification. Lastly, since equity is just a function of firm's assets (in Merton's model, that is just a call option), its volatility can help to identify σ_A:

$$\sigma_{A,t} = \frac{E_t}{D_T + E_t}\sigma_{E,t}c\,(.) = (1 - LEV_t)\sigma_{E,t}c\,(.) \tag{2.4}$$

where c(.) is a function of firm's leverage and $\sigma_{E,t}$ is estimated using daily returns in the previous month (more calibration details are provided in Feldhütter & Schaefer (2018) and Appendix C).[8]

[8] Following Feldhütter & Schaefer (2018), $c(.)$ attains a value of 1 if $LEV_t < 0.25$, 1.05 if $0.25 < LEV_t \leq 0.35$, 1.10 if $0.35 < LEV_t \leq 0.45$, 1.20 if $0.45 < LEV_t \leq 0.55$, 1.40 if $0.55 < LEV_t \leq 0.75$ and 1.80 if $0.75 \geq LEV_t$. The choice of this stepwise function is purely for computational reasons. The exact representation: $c(.) = 1/N(d_1)$, where d_1 is defined in equation 15.

The simplest version of Merton's model assumes that every firm has a single maturity zero coupon debt whose face value determines the default threshold. That is clearly a strong approximation of the reality where a median issuer has 3 outstanding bonds with different maturities. To accommodate that, I will calibrate the firm-level parameters based on a weighted-average debt maturity of all company's issues. However, once I need to project the bond-level default probabilities, I will use the issue-level maturity.

It is important to note that the literature on structural default models is rich and contains many extensions that deliver different formulations of equation (2.3). In the baseline, I follow Feldhütter & Schaefer (2018)'s approach as it is shown to match the average credit spreads well. Yet, in the robustness section, I also consider alternative specifications, such as the original Merton (1974) and van Zundert & Driessen (2017)'s version with stochastic volatility. Qualitatively, these modifications do not alter the results much, supporting Huang & Huang (2012)'s claims that most structural models deliver similar predictions.

One potential criticism of this approach points to the credit spread puzzle. Mirroring the equity risk premium puzzle, many structural models fail to match the level of credit spreads, given the observed historical default frequencies (Huang & Huang (2012)). Feldhütter & Schaefer (2018) address this challenge by calibrating their model to a much longer historical sample, starting in 1920, which substantially increases the level of credit spreads.[9] The intuition hinges on the fact that defaults are highly correlated events and we need long history to estimate empirical means accurately. Bhamra, Kuehn & Strebulaev (2010) tackle the same challenge by building a dynamic leverage model with endogenous defaults and different economic regimes. Both studies point to the direction that it is crucial to make adjustments to the default boundaries or introduce dynamic capital structure to bring these models closer to the data. Finally, recent research highlights the importance of the secondary market liquidity, especially in the international dimension (Huang, Nozawa & Shi (2022)). Even though in this paper I do not explicitly model liquidity, in the extensive robustness tests I investigate the effects of liquidity on the results.

Finally, in line with the existent literature, I sort bonds into duration deciles formed every June (monthly rebalancing delivers similar results). To isolate the pricing of the longest maturity bonds, I also form 10 maturity buckets with 2, 3.5, 5..., 13+ year maturities. For simplicity, all portfolio returns are equally weighted.

2.3 Data

Bonds

The bond transactions data comes from TRACE enhanced (2002-2020). In order to eliminate errors, corrections and trade reversals, I follow Dick-Nielsen (2013). Similarly as in Jostova, Nikolova, Philipov & Stahel (2013), I proxy the end-of-the-month prices

[9] The result holds at the aggregate and in different credit rating sorts. In the future research, it remains to be investigated how well this is matched in other characteristic sorts, such as value, size, etc.

using the most recent price in the last 5 trading days of the month, and compute the daily prices by volume-weighting all the trades on that day. If the bond is not traded during those 5 last trading days of the month, that bond-month observation is excluded. For all other filters, I closely follow Bai et al. (2019). In particular, I exclude from the sample all bonds with less than 1-year time-to-maturity because major corporate bond indices, e.g. the Barclays Capital Corporate Bond Index, automatically delist them. In addition to that, I remove from the sample bonds with maturities over 30 years as the number of firms able to issue such debt is small, and the determination of the appropriate risk-free rate is challenging.

The information regarding bond characteristics, such as coupon, maturity, credit rating, is obtained from FISD. This dataset is appended using the WRDS TRACE masterfile. In cases where there are mismatches between FISD and the WRDS TRACE masterfile, I use FISD as the primary source of information.

Firm-level characteristics

The data on firm-level debt, book-equity, as well as dividends, repurchases and interest paid, come from Compustat. In addition, the stock prices, common shares outstanding, monthly and daily returns are taken from CRSP. All the financial statement information is lagged by 3 months. The book-equity is constructed following Fama & French (2002). The market caps are defined as the product of the end-of-the-month stock price (prc) times the number of common shares outstanding (shrout). If the firm has more than one type of common equity shares traded, I sum all the individual-stock market caps to determine the firm-level market cap. In order to calculate the book-to-market ratios, I divide the most recent available book-equity measure by the most recent market cap[10].

Risk-free rate

The US government yields data comes from Gürkaynak, Sack & Wright (2007). The authors fit a Nelson-Siegel-Svensson curve to all of the off-the-run government bonds ranging from 4-month to 30-year maturity, and provide the daily estimates of this parametrization. For robustness, I also consider interest rate swaps as an alternative measure of risk free rates. This data is available on Bloomberg.

Credit ratings and historical default estimates

The issue-level credit rating data comes from FISD. I use any of S&P, Moody's and Fitch ratings (in this order of importance) to determine the riskiness of the issue. While most issuers have a rating, a large share of individual bonds does not. In those cases, I assign to the bond the issuer's credit rating that comes from Compustat.[11] In case Compustat does not provide data on the issuer's rating, I proxy the missing bond's rating with the

[10]Since I form portfolios in June, the vast majority of firms are sorted based on the book-equity from the previous financial year ending in December, as in Fama & French (2002).

[11]Compustat ceased to update the information on ratings in 2017.

median credit rating across all traded bonds issued by the same company. Consequently, only 83,127 (7.2%) bond-month observations remain unassigned to any group. Lastly, I consider that the bond defaults if the credit rating downgrade issued by the rating agency falls below C (for Moody's, it equals C).[12]

The historical data on the historical defaults (1920-2020) and recovery rates (1983-2020) in each rating group across maturities come from Moody's Investors Service (2021). For both series, I use the longest samples available to estimate the average default probability and recovery rate term structure. Finally, Moody's Investors Service (2021) provides the recovery rate term structure for maturities of 1-5 years, whereas the default estimates are extended to 20 years.[13] Thus, in order to forecast expected default losses for longer horizons, I presume that the recovery rate term structure (as well as conditional default probability) is flat after 5 (20) years.

Summary statistics

The final sample consists of roughly 500,000 bond-month observations. Most of the sample size restrictions come from the requirements of market equity and balance sheet data for the estimation of default probabilities. In general, the sample statistics closely resemble those of Bai et al. (2019). The average bond has a mean return of 0.54% per month and a yield of 4.64%. The median bond has a credit rating of BBB and the principal of $400 million, which is economically substantial. An average issuer has at least 3 different bonds outstanding in a given month, and a leverage ratio of 0.35.[14] Given that roughly 60-70% of total firm's debt is publicly traded, the corporate bonds constitute roughly 25% of the public firm's balance sheet (note that for some firms this may reach over 60%).

Regarding the maturity and duration distribution, we can see that most of the bonds outstanding are of 5-year maturity and duration. There is a wide array of bonds that have long (over 10-year) maturities, however, rarely a bond exhibits duration beyond 20 years. That is party explained by the fact that only the highest creditworthiness issuers are able to issue such long-term debt, and that yields on such debt are generally higher (as we see in the table, the correlation is -0.16 and 0.25, respectively). Nevertheless, there is still a sizable fraction of long-maturity bonds with duration reaching 10 or more years.

[12] The realized defaults occur very rarely, so the end results are not affected by the inclusion or exclusion of these observations.

[13] As presented in Table A.I, there are very few historical defaults to estimate the recovery rates for the Aaa rating category. To alleviate the sample biases, I presume that the recovery rates for Aaa are the same as for Aa.

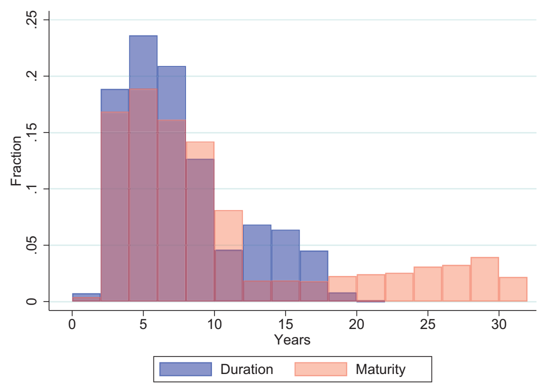
[14] A few issuers have an extreme number of bonds outstanding, however, the final results are not sensitive to the exclusion of those observation.

Table I: Historical samples

The table below summarizes the sample statistics (2002-2020) of monthly observations at the bond- and issuer-level.

				Percentiles					Corr with
Bond-level	Mean	Median	SD	1	5	95	99	N	maturity
bond return	0.54	0.40	3.68	-7.37	-2.85	4.16	9.27	440,340	0.03
yield	4.64	4.28	2.56	0.81	1.48	8.65	17.03	504,779	0.25
spread	2.23	1.53	2.31	0.22	0.45	6.25	13.98	504,779	0.00
coupon	5.33	5.35	1.80	1.45	2.43	8.38	9.88	505,019	0.15
rating	8.37	8.00	2.94	2.00	4.00	14.00	16.00	490,485	-0.16
principle (mln)	563	400	610	2	7	1500	3000	502,481	0.02
Issuer-level									
bonds issued	8	3	25	1	1	25	99	111,169	
LEV	0.35	0.31	0.22	0.03	0.08	0.79	0.94	111,169	
mcap (bln)	19.3	6.0	47.0	0.1	0.4	77.5	220.6	111,169	
total debt (bln)	11.3	2.5	50.0	0.2	0.4	28.9	241.8	111,169	

Figure I: Distribution of bond-month observations across maturity and duration



Note: The graph above presents the distribution of TRACE sample (2002-2020) monthly observations. The maturity and duration of a bond is rounded up (e.g., 2.3-year bond is treated as a 3-year bond). Bonds with maturities less than 1 year or above 30 years are excluded.

2.4 Empirical Findings

2.4.1 Full sample

I will start my analysis by discussing the full sample results over the period 2002-2020. Figure II presents the decomposition of yields into the risk premium, credit loss, and risk-free rate components across different duration portfolios. One can notice that the average term structure of yields is upward-sloping, and most of this effect comes from the upward-sloping term structure of risk-free rates (depicted in gray). On the other hand, the credit spreads, i.e. the sum of blue and green areas, are somewhat hump-shaped. The subsequent analysis decomposes these spreads using the Merton's model, and historical Moody's default probabilities.

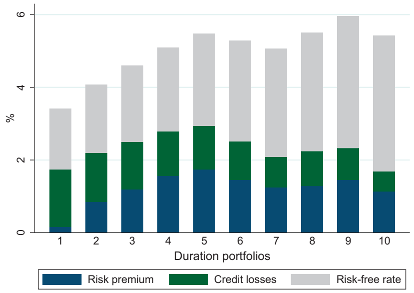
The key finding is that the term structure of the risk premia is somewhere between being flat or upward sloping (depicted in blue). Since the average risk-free rate term structure has a positive slope, the total expected returns are strongly upward-sloping, yielding roughly a 2.3-3.4% term premium. Most of this term premium is driven by the slope of the risk-free rate (2.1%), however, the risk premia slope contributes another 0.2-1.3%. As a result, the risk premia slope can be economically substantial as it constitutes up to 20% of the long-term bond yields, and up to 30% of total expected returns. The positive risk premia slope is in line with the classical models of habit and long-run risk that link the maturity of the asset with higher risk. However, based on the historical probability estimates, we cannot fully reject the validity of the rare disasters model that delivers a flat term structure.[15]

The absolute size of risk premia is well within the reasonable economic magnitude. The short-term bonds, of 1-2 year duration, carry a premium ranging from 0 to 1.4%.[16] Although there is a lot of heterogeneity in the composition of bonds, the average short-term risk premium seems to be relatively modest, indicating that investors do not treat corporate bonds strikingly differently from short-term government bonds. The long-term (12+ year) risk premium levels off at a much narrower interval of 1.2-1.5%, sometimes peaking at 2.2% in the intermediate horizons. The results are qualitatively invariant to the exact sorting strategy, i.e. using deciles vs fixed duration bins. At first sight, the long-term risk premium may appear to be small, especially if one compares these to the equity term structure estimates of 10-20% presented in Weber (2018). Some of these differences are driven by the fact that here the bond returns are net of the maturity-matched risk-free rate. Moreover, equity risk premium arguably pertains to the upside risks that bonds have limited exposure to. These explanations justify economically smaller but still substantial corporate bond risk premia.

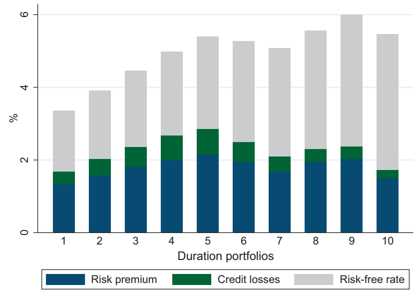
[15] The intuition behind this result relies on the constant probability of rare disasters (Gabaix (2012)). The asset pricing model predictions are summarized in Table A.IX. A good overview of the theoretical asset pricing models effect on the equity term structure slope can be found in Gormsen (2021).

[16] The average duration associated with each duration decile is presented in Appendix Table A.II.

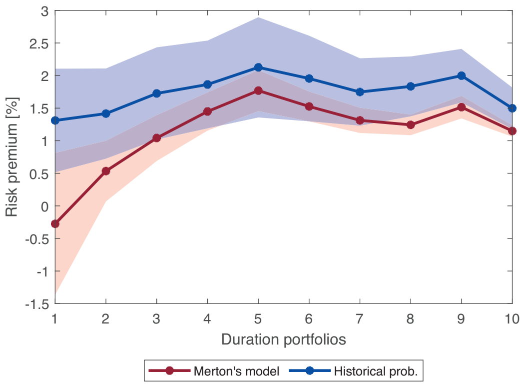
Figure II: Decomposition of yields and term structure of risk premia



(a) Merton's model



(b) Historical prob.



(c) Risk premia term structure

Note: The figures above present the average yields, credit spreads, and decomposition of credit spreads into the risk premia and credit losses components. The top left figure displays the decomposition based on the Merton's model predictions, whereas the right one relies on the historical default frequencies from Moody's Investors Service (2021). The bottom graph depicts the risk premia term structures based on both methods, including the 95% confidence intervals based on the Newey-West standard errors with a 36-month lag.

Table II: Risk premia across duration deciles

The table reports the annual risk premia estimates in each duration portfolio. The numbers in the brackets are standard errors based on the Newey-West method with a 36-month lag. The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Merton's model	-0.26	0.55**	1.03***	1.41***	1.78***	1.52***	1.30***	1.22***	1.47***	1.14***	1.36***
	(0.56)	(0.22)	(0.17)	(0.15)	(0.15)	(0.11)	(0.09)	(0.07)	(0.08)	(0.04)	(0.51)
Hist prob.	1.31***	1.42***	1.74***	1.87***	2.18***	1.98***	1.76***	1.83***	2.08***	1.51***	0.23
	(0.42)	(0.36)	(0.39)	(0.35)	(0.42)	(0.35)	(0.28)	(0.26)	(0.25)	(0.17)	(0.24)
N	188	198	198	198	198	198	198	198	198	198	196

Regarding the statistical inference, one can see that the long-short portfolios earn a statistically significant risk premium based on the Merton's model, yet the historical probabilities deliver a slope that is too flat to be significant (Table II). Given that the estimates are persistent, all the standard errors are estimated with the Newey-West procedure using 36 lags.[17] Merton's model delivers narrower confidence intervals. The predicted default probabilities positively correlate with the increase in credit spreads which reduces the estimation noise, whereas the historical averages do not allow for any time variation. In other words, the historical default probabilities attribute all the variation in the spreads to the discount rate news channel, and there is no room to improve the precision of the estimates. Adding to that, the short-term risk premium displays a great deal of time-series variation relative to the long-term risk premia. This is an intuitive result given that the long-term asset returns are averaged over multiple periods ahead and are not that sensitive to the transitory shocks in the investor preferences or economy. In fact, Mueller, Vedolin & Zhou (2019) document that the short-term (government bond) risk premium is more sensitive to transient economic variables, such as implied option variance, whereas the long-term risk premium is more exposed towards slower cyclical indicators. A similar mechanism can be at play here. Having said that, the long-term asset may still have far larger variance of realized returns because small changes are greatly amplified by a much larger duration.

Glossing over the initial findings, it may be surprising why the two methodologies deliver quite a wide range of possible slope values. One explanation for this is that over the last 20 years we have witnessed a large frequency of recession months (Bansal et al. (2021)). The structural estimation of default probabilities is sensitive to large increases in the leverage ratio, a phenomenon prevalent during the financial crisis. On the other side of the spectrum, the default probabilities over the last 100 years have attributed only a small chance to such events. Loosely speaking, one can view the Moody's historical probabilities as a very conservative estimate of the slope, whereas the Merton's model delivers an upper bound. Another interesting observation is that both methodologies deliver a downward-sloping term structure of credit losses. There are two core reasons justifying that. Firstly, the risky issuers tend to issue more short-term debt so the aggregate term structure tallies in more safe firms over long horizons. As a matter of fact, it will be shown later that most of this hump-shape arises due to the bonds with the lowest credit rating. This also suggests that the current estimates of the risk premia slope are downward biased. Secondly, both historical and model-generated default probabilities exhibit concave CDF's, which implies that the forward default probabilities are decreasing with maturity, and more so among the riskiest firms. This insight relies on the idea that, over time, firms mature and, if they survive the distress, they transition into more stable credit ratings.[18] Overall, the discrepancies between the two methodologies establish reasonable bounds for the slope,

[17] 36 months are chosen as the most distant significant autocorrelation in all duration portfolios. In general, this means that the standard errors are potentially too conservative for some portfolios.

[18] The Merton's model features similar relationships between the maturity of the debt and default risk. Except for the safest firms, with maturity the probability of default increases at a slower rate because the asset volatility scales at a slower rate than the asset return drift.

and highlight the need to conduct deeper time-series and cross-sectional analysis.

2.4.2 Cross-sectional patterns

The maturity dimension of a bond may not be sufficient to understand how investors price short- and long-maturity risks in the corporate bond market. In this section, I will investigate if the cross-sectional characteristics that are known to predict returns in the bond market are also related to the slope of the risk premia term structure. The characteristics I consider are: default risk, such as leverage and credit rating (similarly as in Bai et al. (2019)), as well as book-to-market ratio and size factors that are often used to predict returns (e.g. Bartram et al. (2021)). This additional slice of data helps to understand which maturity risks originate the source of factor premia. On top of that, the double sorts control for risk characteristics that may strongly correlate with the issuer's ability to issue different maturity bonds. In order to form diversified portfolios, I sort bonds first on the factor characteristic and then into duration quintiles (fixed bins deliver similar results).

The general finding is that most cross-sectional sorts deliver an upward-sloping term structure of risk premia regardless of the estimation method (Table III). This suggests that the full sample results are not purely driven by self-selection issues. In fact, even historical default probabilities, which are more conservative than Merton's model, most of the time deliver positive and statistically significant slope estimates. This signals that the rather flat aggregate patterns are indeed somewhat downward biased. Based on the credit rating sorts, the slope estimates range between 0.5 to 1.7%, with the exception of the lowest credit rating firms. Namely, the riskiest issuers exhibit a relatively flat term structure, a pattern that is potentially explained by high exposure to the short-term risks (such as the short-term crash risk). All other bonds are priced in a qualitatively consistent way with the leading asset pricing models of habit and long-run risk. Interestingly, the hump-shape observed in the aggregate credit spreads and yields disappear in the higher credit rating sorts, confirming that it is predominantly driven by the issuer composition effects. To strengthen the case, the estimates based on historical probabilities also display that.

The remaining sorts exhibit positive slopes with, at times, inconclusive results. The leverage ratio is a strong proxy for the distance to default, and should correlate with the default risk and credit ratings. On the other hand, it controls for the modeling errors as high-leverage firm default rates are most sensitive to changes in parameter values. Thus, it is not surprising that the term structure slope is positive and very similar under both methodologies, but there are mismatches in the high leverage category. The subsequent analysis shows that most of these misalignments arise due to the financial crisis period.

Table III: Cross-sectional sorts

The tables present the average risk premia in duration and BM, size, leverage and credit rating double sorts. The numbers in the brackets are standard errors based on the Newey-West method with a 36-month lag. The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	High	High-Low
Credit rating sorts						
>BBB+	-0.79	0.18	0.48***	0.80***	1.01***	1.80***
	(0.61)	(0.17)	(0.13)	(0.10)	(0.08)	(0.56)
BBB	0.16	0.62***	0.97***	1.16***	1.40***	1.24***
	(0.29)	(0.11)	(0.09)	(0.05)	(0.04)	(0.28)
<BBB-	2.89***	3.48***	3.33***	3.17***	3.01***	0.11
	(0.32)	(0.32)	(0.23)	(0.25)	(0.19)	(0.26)
Leverage sorts						
Low	0.83***	1.18***	1.53***	1.30***	1.37***	0.54***
	(0.10)	(0.11)	(0.15)	(0.07)	(0.04)	(0.07)
High	-0.48	1.26***	1.80***	1.10***	0.99***	1.48**
	(0.85)	(0.34)	(0.17)	(0.17)	(0.15)	(0.71)
Book-to-market sorts						
Growth	0.65***	1.22***	1.67***	1.26***	1.29***	0.64***
	(0.16)	(0.12)	(0.16)	(0.07)	(0.09)	(0.11)
Value	0.67***	1.60***	1.83***	1.42***	1.35***	0.68***
	(0.23)	(0.16)	(0.14)	(0.07)	(0.05)	(0.20)
Size sorts						
Small	0.73**	1.73***	2.13***	1.53***	1.60***	0.86**
	(0.35)	(0.18)	(0.16)	(0.09)	(0.07)	(0.33)
Big	-0.21	0.39***	0.71***	0.98***	1.14***	1.35***
	(0.33)	(0.07)	(0.07)	(0.06)	(0.05)	(0.30)

(a) Merton's model

	Low	2	3	4	High	High-Low
Credit rating sorts						
>BBB+	1.02***	1.08***	1.38***	1.53***	1.50***	0.48***
	(0.31)	(0.29)	(0.33)	(0.22)	(0.17)	(0.15)
BBB	1.32***	1.54***	1.80***	1.86***	1.98***	0.67***
	(0.34)	(0.37)	(0.35)	(0.31)	(0.21)	(0.16)
<BBB-	2.83***	3.51***	3.14***	3.02***	2.70***	-0.13
	(0.55)	(0.55)	(0.44)	(0.40)	(0.33)	(0.35)
Leverage sorts						
Low	0.98***	1.23***	1.53***	1.42***	1.52***	0.53***
	(0.28)	(0.30)	(0.29)	(0.21)	(0.15)	(0.15)
High	1.92***	2.49***	2.73***	2.27***	2.07***	0.15
	(0.50)	(0.52)	(0.47)	(0.39)	(0.27)	(0.27)
Book-to-market sorts						
Growth	1.06***	1.39***	1.74***	1.47***	1.53***	0.46***
	(0.30)	(0.29)	(0.34)	(0.22)	(0.16)	(0.16)
Value	1.71***	2.20***	2.42***	1.98***	1.82***	0.11
	(0.43)	(0.38)	(0.36)	(0.28)	(0.19)	(0.27)
Size sorts						
Small	1.96***	2.44***	2.65***	2.13***	2.29***	0.34
	(0.49)	(0.48)	(0.48)	(0.40)	(0.31)	(0.27)
Big	0.96***	1.07***	1.28***	1.55***	1.54***	0.59***
	(0.25)	(0.26)	(0.22)	(0.21)	(0.16)	(0.10)

(b) Historical prob.

There is some evidence for the value premium in the corporate bonds in line with Bartram et al. (2021)'s findings. The bonds issued by growth firms exhibit a term premium of roughly 0.4-0.6% per annum, with a moderate level of risk premia. In contrast, the value firms display a higher level of risk premia, however, on this matter, the two methodologies are not entirely consistent. Namely, the value premium is quite large based on the historical probabilities (up to 0.8%, mostly among the shortest duration bonds), but the Merton's model delivers much more modest estimates. This may hint that most of the value premium risks are of short-term nature and over time, current growth firms become similar to the value firms (which is in line with Giglio et al. (2020)). Yet, the results are inconclusive.

On the other hand, there is quite distinct evidence documenting the size premium of 0.4-1.4%. At first sight, this may be an artifact of lower liquidity, which will be investigated later, yet the largest differences arise predominantly among the short-term bonds and narrow down with maturity. Relative to the book-to-market sorts, the size premium seems to be more persistent, a finding that resonates with Yara et al. (2020). Regarding the slope, the big firms seem to have a stronger maturity and risk premia relationship. Once again, this can be aligned with the higher exposure of small firms to the short-term or business cycle risks.

To sum up, most of the cross-sectional sorts support the upward-sloping term structure of risk premia. Only the highest credit risk firms display a flat term structure. In addition, the cross-sectional sorts reveal an economically and statistically significant size and credit risk premia, with some weaker support for the value premium. Interestingly, in all sorts, the gap between different cross-sectional portfolios decays with maturity, supporting the idea that most of the factor premium arises from the short- and intermediate-horizon risks.

2.4.3 Time variation

There is plenty of evidence showing that expected returns in equity markets vary over time.[19] Similarly, the term structure of risk-free rates has dramatically changed since the financial crisis with the introduction of unconventional monetary policy. Following the same logic, there is no reason to believe that the term structure of risk premia has not shifted since 2008. In order to investigate these movements, I split the sample into three economic periods: before the financial crisis (2002-2007), the financial crisis (2007-2009), and after the crisis (2009-2020). Dicing the data into smaller pieces comes at a cost - the long-term historical default probabilities become less and less reliable. For this reason, the time variation analysis is performed using only the Merton's model. These results are presented below in Table IV.

One can immediately detect contemporaneous shifts both in the risk premia and risk-free rate term structures. Both of them were much flatter before the financial crisis compared to the most recent decade. After the crisis, however, both slopes are statistically significant and positive. The risk premia term structure did not experience a strong parallel shift down, but both the short-term and long-term bonds were affected by the steepening of the curve. One can immediately spot that the financial crisis was an exceptional event delivering a number of extreme observations. The negative estimates on the short end may seem counter-intuitive given the classical asset pricing model predictions of countercyclical risk premia. To a large degree, these are the structural default model errors that predicted far larger expected losses given the sudden drop in the equity prices, and a surge in the leverage ratios and market volatilities. This highlights that predicting expected returns during the financial crisis is extremely challenging with structural models, and for this reason, the sample splits exclude this period. Despite these outliers, the overall pattern is quite clear - the term structure of risk premia has a flat or positive slope even at the aggregate level.

This subsample analysis unveils some movements in the term structures that may produce one-off capital gains, in a similar vein as Fama & French (2002). It is an important

[19] Shiller (1981), Li, Ng & Swaminathan (2013) to mention a few.

insight as most of the cross-sectional studies bank on the realized returns to estimate risk premia. Yet, the time series variation is only of the key factors confounding the measurement using the price history alone. Thus, the following section will further analyze these general patterns based on the realized returns, and to what degree they match the implied measures of risk premia.

Table IV: Historical samples

The table presents the average risk premia and risk-free rate term structures in three historical subsamples. The numbers in the brackets are standard errors based on the Newey-West method with a 36-month lag (except for the crisis period where a 6-month lag was used due to the sample size constraints). The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Full sample											
risk premium	-0.26	0.55**	1.03***	1.41***	1.78***	1.52***	1.30***	1.22***	1.47***	1.14***	1.36***
	(0.56)	(0.22)	(0.17)	(0.15)	(0.15)	(0.11)	(0.09)	(0.07)	(0.08)	(0.04)	(0.51)
rf	1.53***	1.77***	2.00***	2.22***	2.46***	2.71***	2.93***	3.19***	3.61***	3.73***	2.12***
	(0.53)	(0.54)	(0.50)	(0.47)	(0.44)	(0.42)	(0.41)	(0.40)	(0.38)	(0.36)	(0.41)
Before financial crisis (2002.06-2007.12)											
risk premium	0.46*	0.80***	0.94***	1.26***	1.57***	1.52***	1.36***	1.10***	1.37***	1.06***	0.55**
	(0.25)	(0.24)	(0.12)	(0.10)	(0.05)	(0.07)	(0.04)	(0.02)	(0.03)	(0.03)	(0.22)
rf	3.45***	3.64***	3.78***	3.92***	4.05***	4.23***	4.41***	4.56***	4.87***	4.97***	1.41*
	(0.71)	(0.58)	(0.46)	(0.37)	(0.30)	(0.22)	(0.15)	(0.12)	(0.04)	(0.06)	(0.72)
After financial crisis (2009.06-2020.06)											
risk premium	0.25	0.72***	1.26***	1.54***	1.85***	1.58***	1.35***	1.33***	1.56***	1.21***	0.97***
	(0.19)	(0.11)	(0.14)	(0.19)	(0.22)	(0.16)	(0.10)	(0.05)	(0.10)	(0.04)	(0.18)
rf	0.82**	0.99***	1.24***	1.49***	1.75***	2.00***	2.23***	2.52***	2.97***	3.14***	2.32***
	(0.34)	(0.31)	(0.27)	(0.22)	(0.18)	(0.15)	(0.14)	(0.16)	(0.20)	(0.21)	(0.48)
Financial crisis (2007.12-2009.06)											
risk premium	-5.61***	-1.42***	-0.39	0.96***	1.86***	1.10***	0.74***	0.79***	1.18***	0.92***	6.53***
	(1.82)	(0.27)	(0.43)	(0.27)	(0.11)	(0.25)	(0.20)	(0.23)	(0.10)	(0.09)	(1.80)
rf	1.36***	1.58***	1.98***	2.29***	2.68***	3.08***	3.44***	3.85***	4.25***	4.19***	2.83***
	(0.36)	(0.34)	(0.31)	(0.31)	(0.26)	(0.22)	(0.19)	(0.15)	(0.17)	(0.19)	(0.21)

2.5 The Puzzle of Realized Returns

To begin with, I compute the (gross) realized returns which are defined in the following way:

$$ret_{i,t,t+1} = \frac{P_{i,t} + AI_{i,t} + c_{i,t}}{P_{i,t-1} + AI_{i,t-1}} - 1 \tag{2.5}$$

where $P_{i,t}$ is the clean price of a bond at the end of month t, AI stands for the accrued interest, and c_t is the coupon payment. Based on these gross returns, one can express the excess returns:

$$eret_{i,t,t+1} = ret_{i,t,t+1} - rf_{i,t,t+1} \tag{2.6}$$

where $r_{f,i}$ is a return on a US nominal government bond with identical cash flows and maturity to the corporate bond (find more details and example calculations in Appendix). Given these returns, I can sort the bonds into different duration or maturity portfolios the same way as before to test if the realized returns line with the expected return measures.

Figure III displays that there is a slight negative relation between the excess realized returns and duration of a bond. The result is invariant to the sorts based on duration deciles or the fixed duration bins. The graph portrays how noisy the realized returns are, and that, except for the shortest duration portfolios, no statistically significant patterns are easily discoverable. The longest duration quintile earns approximately 1-2% per annum less than the shortest duration quintile. Moreover, the inverse relationship is even stronger when one looks at the Sharpe ratios. The long-maturity bonds exhibit much more volatility without generating almost any excess returns. This indicates that over the last 20 years the term structure of excess returns was downward sloping, driven predominantly by the shortest and longest maturity bonds, however, this slope is not statistically significant.

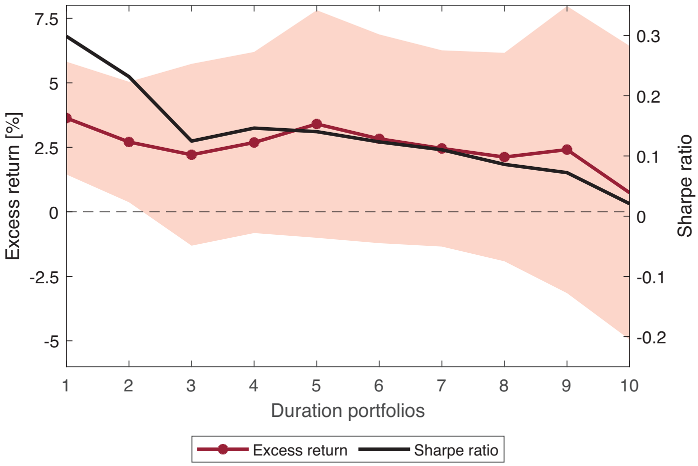
In light of the previous findings, the exceptionally high returns on the short-duration bond portfolios present a puzzle. The implied measures deliver the opposite pattern of results.[20] Paradoxically, the short-term bonds generated so high average excess returns that they far exceeded the average credit spreads in the same period. This is only possible if there were major movements in the expected risk premia or cash flows, and the sample mean deviated from the long-term levels. To further investigate the hypothesis, I correct for the market risk by running CAPM-style regressions where the duration portfolio returns are regressed on the bond market factor. For simplicity, the bond market factor is measured as an equally-weighted mean excess return of all the bonds in a given month.[21] This procedure allows to recover the abnormal patterns that are not explained by the pure correlation with the general market trends.

Table V outlines a stark negative relation between alphas and duration. This result is consistent with Gormsen & Lazarus (2019)'s findings that the realized returns and duration sorts deliver a negative slope both in equities and corporate bonds. Not surprisingly, the bond market betas are increasing with duration. Through the lens of CAPM, assets that bear a higher risk premium must have a larger market beta. However, it is not obvious that the same prediction holds in this setting when the market risk premium exhibits a term structure. Additionally, van Binsbergen et al. (2012) documented that short-maturity dividend assets generate very high excess returns while having low market betas. It appears that the same puzzling pattern reverberates to the corporate bond markets.

[20] To be precise, only the shortest duration portfolios generate statistically significant returns.

[21] The equally weighted returns are chosen since they are much easier to compute, but results are qualitatively very similar to the market-weighted results.

Figure III: Average excess returns



Note: The figure above reports the average excess return and Sharpe ratio in duration deciles. The shaded areas represent the 95% confidence intervals of average excess returns based on the Newey-West standard errors with a 6-month lag.

Table V: Alphas

The table presents the annualized excess and abnormal return estimates in each duration portfolio. The market factor is constructed using an equal average of bond excess returns in a given month. The numbers in the brackets are standard errors based on the Newey-West method with a 6-month lag. The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	5	6	7	8	9	High	High-Low
beta	0.23***	0.43***	0.74***	0.79***	1.07***	1.02***	1.00***	1.09***	1.43***	1.56***	1.33***
	(0.05)	(0.03)	(0.07)	(0.03)	(0.06)	(0.02)	(0.02)	(0.04)	(0.09)	(0.06)	(0.09)
alpha	3.06**	1.73***	0.52	0.88	0.95*	0.48	0.17	-0.37	-0.87	-2.84***	-5.81***
	(0.92)	(0.50)	(0.62)	(0.52)	(0.43)	(0.29)	(0.30)	(0.50)	(0.77)	(0.65)	(1.24)
N	188	198	198	198	198	198	198	198	198	198	196

The initial findings raise a number of questions. Firstly, large alphas indicate that a static one-factor model alone cannot match the risk premia estimates in the corporate bond market, and explain the mismatch with the implied expected return measures. Yet, most of the results are driven by exceptionally high(low) short(long)-duration asset excess and abnormal realized returns. This raises a concern if the patterns survive once we control for the characteristics of the issuers. Secondly, there is a fundamental gap between the risk premia implied from credit spreads and the average monthly returns presented here. The former depicts the holding-to-maturity annualized returns, whereas the latter estimates the term structure over the next month. Averaged over long periods of time, the two objects should deliver the same estimates, but in small samples, there can be a sharp wedge. To inspect how these disparities evolve, I will perform the same realized return analysis with different holding period horizons.

2.5.1 Cross-section of realized returns

The key message of Gormsen & Lazarus (2019) is that duration is a sufficient statistic to explain most cross-sectional anomalies, such as size and value premia. In addition, the positive abnormal returns are generated mostly by the short-duration portfolios lending support to the downward-sloping term structure of risk premia. In the cross-sectional slice of the analysis, I test the validity of these predictions in the corporate bond market.

In a nutshell, the aggregate patterns are reflected in all cross-sectional sorts. As we see in Table VI, the average excess returns display a negative but statistically insignificant slope. It again appears that most of this duration premium comes form the short-term bonds. At the same time, the risk-adjusted returns are strongly downward sloping in all sorts. Economically, they are sizeable too - the long-short portfolios can generate up to 5% risk premia per year, far beyond any credit spreads that we normally observe in the market. Thus, the realized return analysis confirms Gormsen & Lazarus (2019) second claim that the average returns and alphas are downward sloping.

In terms of the level premia, the realized returns share qualitatively similar insights as the implied measures. There is quite a substantial credit risk premium of 3-4.5% which decreases with portfolio duration, however, remains qualitatively large even at long maturities. Hence, I find strong evidence against the credit risk puzzle of Campbell et al. (2008) who documented that low-rated firms earn low average returns. In a similar vein, there is an observable hump-shaped pattern in the average excess returns and alphas which does not necessarily vanish in the credit sorts. Most of these differences are not statistically significant, but their prevalence should not be ignored. Such non-monotonic duration and return relationships are not completely bizarre as it may be a product of the reinvestment and fundamental risk trade-off (Andrews & Gonçalves (2020), Gonçalves (2021)). Namely, the long-term asset may offer hedging benefits driving the total risk premia down. The same pattern can also be seen in the leverage sorts which highly resemble the credit rating results.

Table VI: Cross-sectional sorts

The table presents the annualized excess and abnormal return estimates in each duration portfolio. The market factor is constructed using an equal average of bond excess returns in a given month. For conciseness, the standard errors are omitted. The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	High	High-Low
Credit rating sorts						
>BBB+	1.43**	1.30	1.63	1.37	0.96	-0.47
	(0.68)	(0.96)	(1.47)	(1.72)	(2.70)	(2.32)
BBB	2.01**	2.01	2.77	2.47	1.43	-0.58
	(0.88)	(1.36)	(1.80)	(2.25)	(3.03)	(2.35)
<BBB-	6.52**	4.79	4.77	4.35	4.55	-1.97
	(2.59)	(4.25)	(3.27)	(3.35)	(3.42)	(2.17)
Leverage						
Low LEV	1.40***	1.78	2.13	1.64	1.15	-0.25
	(0.49)	(1.08)	(1.43)	(1.51)	(2.45)	(2.05)
High LEV	3.92**	3.51	4.02	3.27	1.78	-2.14
	(1.76)	(2.90)	(2.89)	(2.98)	(3.77)	(2.48)
Book-to-market sorts						
Growth	1.53***	1.87	2.53	1.86	1.06	-0.48
	(0.53)	(1.13)	(1.83)	(1.69)	(2.56)	(2.17)
Value	4.17**	2.52	3.85	2.51	1.63	-2.55
	(1.80)	(2.40)	(2.63)	(2.18)	(3.05)	(2.03)
Size sorts						
Small	4.10**	3.57	4.00	2.96	2.90	-1.20
	(1.67)	(2.79)	(2.82)	(2.74)	(3.15)	(1.98)
Big	1.32**	1.31	1.72	1.57	0.95	-0.37
	(0.52)	(0.89)	(1.26)	(1.77)	(2.71)	(2.28)

(a) Average excess returns

	Low	2	3	4	High	High-Low
Credit rating sorts						
>BBB+	0.76	0.23	-0.18	-0.79	-2.50***	-3.28***
	(0.54)	(0.36)	(0.48)	(0.50)	(0.78)	(0.89)
BBB	1.36***	0.64	0.88***	0.10	-2.11***	-3.41***
	(0.36)	(0.40)	(0.28)	(0.32)	(0.73)	(0.97)
<BBB-	4.94***	1.13	1.39	0.68	0.75	-3.80**
	(1.59)	(1.84)	(1.04)	(0.92)	(1.01)	(1.88)
Leverage						
Low LEV	0.93***	0.63**	0.45*	-0.24	-1.93***	-2.85***
	(0.23)	(0.30)	(0.26)	(0.30)	(0.67)	(0.74)
High LEV	2.53***	0.83	0.85	0.03	-2.55***	-5.01***
	(0.79)	(0.80)	(0.62)	(0.45)	(0.91)	(1.38)
Book-to-market sorts						
Growth	1.11***	0.63*	0.44	-0.18	-1.99***	-3.01***
	(0.28)	(0.38)	(0.45)	(0.34)	(0.64)	(0.71)
Value	2.98***	0.41	1.13*	0.08	-1.87***	-4.88***
	(1.00)	(0.98)	(0.66)	(0.35)	(0.66)	(1.38)
Size sorts						
Small	3.02***	1.00	1.03*	0.10	-0.55	-3.38***
	(0.92)	(0.90)	(0.53)	(0.37)	(0.82)	(1.08)
Big	0.76***	0.30	0.14	-0.70	-2.46***	-3.25***
	(0.23)	(0.27)	(0.30)	(0.49)	(0.71)	(0.78)

(b) Alphas

Regarding the value and size premia, there is strong evidence that they are priced in bonds. The value firms seem to always generate higher returns, with the exception of the longest duration bonds. That is the reoccurring theme from the yields-based analysis, a confirming sign that the value premium is indeed driven mostly by short-run risks. Another stark observation is that the growth firms' risk premia tend to increase with the maturity of cash flow and peak at around a 5-year duration. Most of this effect comes from the elevated market betas, still one could speculate that the longest-duration bonds issued by growth firms are the most sought-after by the investors. Lastly, the size portfolios point to similar conclusions. The size premium of roughly 2-3% does not easily disappear with maturity, even after the market beta correction. Overall, these findings contradict Gormsen & Lazarus (2019)'s claim that most known anomalies can be subsumed by the duration dimension alone.

2.5.2 Different holding period returns

Different holding period analysis is one way to improve the measurement of risk premia. In particular, some of the bonds are traded infrequently, and if there is any correlation between trading activity, riskiness of a bond and maturity, we may have biased long-term mean and standard deviation estimates. In order to alleviate this challenge, I switch to quarterly, semi-annual and yearly returns. The prime advantage of using longer frequencies is that I can expand the end-of-the-month price window from the last 5 days to a full month while keeping, approximately, the same amount of measurement noise. However, this comes at the price of overlapping observations that mechanically introduce autocorrelation in the bond panel that will be corrected with more lags in the Newey-West method.

Before the presentation of new findings, it is worthwhile to glance at how much more observations are added to the sample once the measurement window is expanded. It turns out that this procedure increases the sample by nearly 30% (pretty uniformly across the duration groups). Appendix Table A.VIII presents the average trading activity of different maturity bonds. Conditional on the fact that a bond is traded, on average, there are 6-7 monthly prices in every given year if one uses a 5-day window. In contrast, this number rises to 9 if a monthly window is used. In other words, an average bond is traded in roughly 58% of all the weeks, and 75% of all the months in a year. Based on the average trading activity, it does not seem that there are stark liquidity issues and some types of bonds are overwhelmingly less traded than others. Table A.VIII does indicate that the shortest maturity, low BM ratio and highest credit worthiness firms are slightly less traded than others. Yet, the disparities, at first sight, are small and alternative horizon returns should not substantially affect the monthly frequency results.

As we see in Table VII, the negative pattern also persists in longer holding period horizons. It has to be mentioned that the slope is gradually decreasing with the holding period horizon. Moreover, the average excess returns become more non-linear on the short leg. However, at a much lower rate than what we need to confidently close the distance between the implied returns and realized ones. At this point, there are indications that the time variation in the risk premia is present, yet the holding periods of up to one year are not enough to fully expose it.

Table VII: Alphas

The table presents the annualized excess and abnormal return estimates in each duration portfolio. The market factor is constructed using an equal average of bond excess returns in a given month. The numbers in the brackets are standard errors based on the Newey-West method. The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Average excess returns											
Quarterly	2.94***	2.52**	2.26	2.43	2.95	2.62	2.32	1.58	2.05	0.56	-2.27
	(1.12)	(1.25)	(1.57)	(1.67)	(2.22)	(2.05)	(1.89)	(2.04)	(2.98)	(2.76)	(2.29)
Semi-Annual	2.98**	2.87**	2.40	2.52	3.16	3.10	2.41	1.35	1.81	0.49	-2.39
	(1.37)	(1.41)	(1.50)	(1.63)	(2.15)	(2.12)	(1.91)	(1.97)	(2.78)	(2.54)	(1.77)
Annual	2.82**	3.10**	2.66*	2.93*	3.43*	3.12	2.19	1.44	1.84	0.83	-1.95
	(1.25)	(1.32)	(1.38)	(1.59)	(1.94)	(1.95)	(1.64)	(1.70)	(2.33)	(2.11)	(1.28)
Alphas											
Quarterly	2.15***	1.42***	0.63	0.73*	0.61**	0.48**	0.32	-0.56	-1.05	-2.51***	-4.63***
	(0.78)	(0.40)	(0.38)	(0.39)	(0.28)	(0.22)	(0.26)	(0.41)	(0.72)	(0.54)	(1.07)
Semi-Annual	2.02***	1.50***	0.67**	0.72**	0.68***	0.77*	0.26	-0.84*	-1.36**	-2.52***	-4.49***
	(0.72)	(0.41)	(0.31)	(0.31)	(0.21)	(0.41)	(0.35)	(0.46)	(0.65)	(0.47)	(0.96)
Annual	1.37**	1.39***	0.80***	0.83***	0.75***	0.60	0.02	-0.76	-1.42***	-2.12***	-3.49***
	(0.58)	(0.45)	(0.20)	(0.26)	(0.20)	(0.38)	(0.30)	(0.58)	(0.53)	(0.35)	(0.66)

2.6 Reconciling the Puzzle

Section 2.4.3 presented evidence that the slope of the risk premia term structure has steepened over time. Such changes should have generated capital gains that pushed the average excess returns up, especially for the shortest-maturity assets. In this section, I will attempt to quantify this effect.

To begin with, the change in the price of a bond can be proxied by the second-order Taylor series expansion:

$$\Delta P \approx \frac{\partial P}{\partial y}\Delta r_f + \frac{\partial P}{\partial y}\Delta rp + \frac{1}{2}\frac{\partial^2 P}{\partial y^2}[(\Delta r_f)^2 + (\Delta rp)^2 + 2\Delta r_f \Delta rp]$$

where Δr_f, Δrp are the changes in the risk-free rate and risk-premia, respectively. Ignoring certain terms that are quantitatively close to zero, one can rewrite this as:

$$\underbrace{\frac{\Delta P}{P} - \frac{\Delta P_f}{P_f}}_{\text{excess return}} \approx \underbrace{-\frac{1}{1+y}\left[D + \frac{1}{2}\Delta D\right]\Delta rp}_{\text{risk premia channel}} \quad \underbrace{-\frac{1}{1+y}\Delta D \Delta r_f}_{\text{convexity due to risk free rate}} \tag{2.7}$$

where P_f is the cash-flow matched synthetic government bond, and D is the bond's duration. The first term on the right-hand side captures the usual first order and convexity terms, whereas the second term corrects for the extra convexity generated by the concurrent changes in the risk-free rate term structure.

The first clear indication that the term structure of risk premia has shifted stems from the sample split analysis with realized returns. Table VIII provides a snapshot at the pre- and after-crisis periods. The main insight is that the term structure slopes both in alphas and average excess returns are much less pronounced before 2008. After the crisis, the excess returns have been heavily elevated for the shortest duration assets (the same picture is reflected in the crisis period too). In fact, the returns are so high that they skew the full sample results towards the negative slope. This may be driven by the surging short-term risk since the last financial crisis, or sizeable one-off capital gains that pushed the realized returns up even though the expected returns moved in the opposite direction.

Table VIII: Subsamples and realized returns

The table presents the annualized excess and abnormal return estimates in each historical period. The market factor is constructed using an equal average of bond excess returns in a given month. The numbers in the brackets are standard errors based on the Newey-West method. The stars indicate standard significance levels: * - 10%, ** - 5%, *** - 1%.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Before financial crisis (2002.06-2007.12)											
excess return	1.76**	1.11	0.70	2.12	2.34	1.35	1.56	0.19	0.83	-1.19	-2.72
	(0.66)	(0.97)	(1.18)	(1.70)	(1.58)	(1.93)	(1.67)	(1.25)	(2.22)	(1.97)	(1.85)
α	1.42**	0.83	0.35	1.62	1.79**	0.61	0.83*	-0.43	-0.26	-2.15**	-3.46***
	(0.58)	(0.53)	(0.59)	(0.97)	(0.71)	(0.65)	(0.45)	(0.34)	(0.95)	(0.84)	(1.10)
After financial crisis (2009.06-2020.06)											
excess return	3.94***	2.96***	2.91**	3.22**	3.79**	3.84**	3.35**	3.44*	3.27	1.80	-2.14
	(1.27)	(1.03)	(1.14)	(1.47)	(1.52)	(1.49)	(1.43)	(1.76)	(2.69)	(2.29)	(2.30)
α	3.22***	1.73***	1.06*	0.80	0.78	0.65**	0.08	-0.43	-2.57***	-3.85***	-7.05***
	(1.02)	(0.53)	(0.60)	(0.68)	(0.48)	(0.28)	(0.34)	(0.49)	(0.89)	(0.90)	(1.52)
Financial crisis (2007.12-2009.06)											
excess return	6.18	5.72	1.86	0.71	3.92	0.23	-1.10	-1.31	1.16	-0.91	-7.09
	(6.51)	(7.87)	(16.03)	(13.68)	(19.57)	(16.45)	(15.65)	(16.39)	(21.66)	(24.79)	(21.03)
α	6.04	5.43**	1.31	0.18	3.16	-0.44	-1.73	-1.98	0.38	-1.86	-7.90
	(4.98)	(2.28)	(3.92)	(1.49)	(2.00)	(1.34)	(1.81)	(2.89)	(3.07)	(2.79)	(6.05)

The capital gains analysis shows that the steepening of the term structure can only partially explain the puzzle of realized returns. As we can see from Table IX, the time variation in the slope generates heightened short-maturity bond returns by up to 0.7%. This is the upper bound of the estimate as the capital gains are averaged over time. On the other hand, the long-term bonds should have experienced a reduction in their excess returns of up to 1.7%, a prediction that is contradicted by the overall increase in the realized returns after the crisis (surprisingly, the abnormal returns concurrently declined). Overall, the variation in the yields-based risk premia is too small to capture the drastic dynamics in the realized returns, and the puzzle remains. However, to investigate these effects more rigorously, I will implement the affine model for risk premia dynamics.

Table IX: Capital gains

The table below summarizes the changes (Δ) in the annualized realized returns and the yields-based expected returns before and after the financial crisis. The capital gains are calculated using (2.7).

	Low	2	3	4	5	6	7	8	9	High	High-Low
Realized returns											
Before	2.14	1.17	0.99	1.84	2.47	1.26	1.62	0.15	0.84	-1.31	-3.48
After	3.95	2.98	2.87	3.21	3.78	3.79	3.34	3.29	3.36	1.78	-2.17
Δ	1.81	1.81	1.89	1.37	1.31	2.53	1.72	3.15	2.52	3.09	1.31
Expected return estimates											
Before	0.90	0.77	1.03	1.31	1.61	1.50	1.36	1.09	1.38	1.06	0.18
After	0.30	0.72	1.27	1.55	1.85	1.58	1.35	1.34	1.56	1.21	0.91
Δ	-0.61	-0.05	0.24	0.24	0.24	0.09	-0.01	0.26	0.18	0.14	0.73
Capital gain	0.74	0.14	-0.70	-0.92	-1.07	-0.41	0.16	-1.77	-1.70	-1.71	-2.45

2.6.1 Matching the data with a one-factor model

In this section, I will more rigorously quantify the effect of the declining risk premium on realized returns by modeling the whole time-series dynamics of risk premia. The realized returns and the baseline implied estimates are not immediately comparable because of the differences in the holding period. The implied measures of expected returns and risk premia are inherently holding-to-maturity. This comes from the fact that yields are derived from the complete schedule of promised cash flows. While many bonds are held for long investment periods, the distinction does not allow to cleanly compare multi-period with one-month ahead expected returns, $E_t(R_{t+1})$. Yet, without further assumptions about the evolution of expected losses and risk-premia, there is no easy way to link these two measures and quantify the effect of risk premia shifts on the realized returns in the sample. One general solution to this challenge is employing a term structure model.

I will start with standard assumptions. Duffie & Singleton (1999) show that, under some technical conditions, defaultable bonds can be valued similarly to default-free bonds.[22] Namely, the price P of an m-maturity risky bond can be written as:

$$P_t = E_t^Q\Big[\exp(-\int_t^{t+m}(r_s + s_s^Q)ds)\Big] \tag{2.8}$$

where r is the instantaneous short rate and s^Q is the instantaneous credit spread. Abstracting from liquidity effects for a moment, s^Q equals the expected default losses, $h_t^Q L$, under the risk-neutral measure Q. So, effectively equation (2.8) replaces the usual pricing equation containing the bond yield with its continuous time counterparts. As a result, by specifying the dynamics of r_t and s_t^Q, one can model the risk premia and realized returns at any frequency.

To do so, one can model s_t^Q directly by specifying its process and later use the historical

[22] The key assumption here is that loss given default is proportional to the market price of the security.

default probabilities to back out the risk premia estimates (e.g., similarly to Driessen (2005)). Yet, this approach does not utilize the information from the structural models discussed in the previous sections. For this reason, I proceed by modeling the risk premium component directly.

We can decompose the credit spreads into the short-term risk premium, rp_t, and the expected losses component under the physical probability measure (P):

$$s_t^Q = h_t^Q L = \underbrace{(h_t^Q - h_t^P)L}_{\text{risk premium, } rp_t} + \underbrace{h_t^P L}_{\text{expected default losses}} \tag{2.9}$$

where h_t^P is the (instantaneous) probability of default, and L stands for loss given default. The decomposition is well-defined when the markets are complete and there is no arbitrage.[23] For simplicity, I assume that the risk-free rate is exogenously determined and the focus is only on the credit spread component.[24] Additionally, in order to directly link the risk-premium estimates to the measures of implied expected returns, I assume that rp_t and $h_t^P L$ dynamics are independent. This conjuncture is not strictly necessary for the implementation of the model, but it greatly facilitates the comparison with the empirical baseline estimates.

To keep the analysis parsimonious, I introduce a one-factor affine term structure model of Cox, Ingersoll & Ross (1985) (CIR). A standard representation of this model includes a single state variable, F_t, that follows a mean-reverting process:

$$dF_t = \kappa (\theta - F_t)\, dt + \sqrt{\beta F_t} dW_t \tag{2.10}$$

In my setup, this state variable can be interpreted as the short-term risk premium, $rp_t = F_t$. Intuitively, the equation tells us that the systematic risk is time-varying, and there is a premium for holding different maturity bonds because they are exposed to unexpected factor shocks. As in de Jong (2000), I assume that the market price of risk, λ_t, is proportional to the level of shock volatility: $\lambda_t = \psi\sqrt{\beta F_t}$. Based on equation (2.10), the variance scales with the level of the factor, and the risk premium generally rises in more volatile times.

Given the assumptions, one can price bonds by separately modeling risk-free rates, risk premia, and expected default losses. This is particularly useful since I can input the expected default losses from my empirical analysis section without the need to specify its dynamics. Moreover, for a standard square root process, the m-maturity risk premium ($rp_{t,m}$) is just a linear function of F_t with maturity-specific constants and loadings (Hull

[23] The assumption mirrors Jarrow, Lando & Yu (2000)'s argument that, in the absence of arbitrage opportunities, there exists a risk premium parameter μ_t that maps physical default intensity, h_t^P, to risk-neutral default intensity, h_t^Q.

[24] This effectively means that the correlation between spreads and interest rates is not explicitly modeled. Yet, this assumption can be relaxed, as in Driessen (2005). Similar exogenous interest rate assumptions are also embedded in other frameworks like Lettau & Wachter (2007).

(2009)):

$$rp_{t,m} = \frac{1}{m}\left[A(m) + B(m)F_t\right] \tag{2.11}$$

with

$$A(m) = -\frac{2\kappa^*\theta^*}{\beta}\ln\left(\frac{2\gamma e^{[(\kappa^*+\gamma)/2]m}}{(\kappa^*+\gamma)(e^{\gamma m}-1)+2\gamma}\right)$$
$$B(m) = \frac{2(e^{\gamma m}-1)}{(\kappa^*+\gamma)(e^{\gamma m}-1)+2\gamma},$$

where $\kappa^* = \kappa + \psi\beta, \quad \theta^* = \frac{\kappa}{\kappa^*}\theta, \quad \gamma \equiv \sqrt{(\kappa^*)^2 + 2\beta}$.

These parameters can be estimated using a standard Kalman filter and quasi-maximum likelihood (QMLE) following de Jong (2000). More estimation details are provided in Appendix 2.D.

2.6.2 Quantitative results

In my baseline, I employ the risk premium estimates using the Feldhütter & Schaefer (2018) default probabilities and link them to the realized return dynamics over the last 20 years. The QMLE results are presented in Table X. As we see, the long-term mean of short-term risk premium, θ is roughly 1%, which is close to the average maturity-matched risk premium in the sample. The persistence parameter, κ, of 0.529 translates to annual risk-premium mean-reversion of 0.59 which indicates some persistence but to a lesser degree than in government bonds (i.e., de Jong (2000) finds that to be around 0.94). This means that the risk premium dynamics are more volatile, which is in line with the considerable variation in discount rates documented by Shiller (1981). The variance parameter β is sizable and supports the idea that risk-premium over the sample was subject to extreme shocks, such as the financial crisis of 2008-2009. Lastly, the price of risk, ψ, is close in magnitude to the estimates found in government bonds by de Jong (2000).

Table X: Structural parameter estimates

The table below summarizes the Kalman filter and quasi-maximum likelihood (QMLE) estimates of equation (2.10) parameters (as well as their asymptotic standard errors).

	θ	κ	β	ψ
QMLE	0.010	0.529	0.019	-12.320
se	0.003	0.124	0.003	7.269

Another way to analyze the estimation results is by inspecting the fitted time series for each of the duration portfolios. In general, the filtered series closely follow the data: the correlation between the filtered series and the baseline estimates ranges from 70% to 90% (Appendix Figure A.IV). The main errors arise from the shortest duration portfolios that have substantially more volatile patterns, especially during the financial crisis periods.

This is difficult to match using a parsimonious one-factor model. Naturally, the more persistent longer duration portfolio series are more accurately matched.

Given the estimated parameters, one can construct the instantaneous risk premium term structure. In a one-factor model, that equals to the drift of m-maturity risk premium dynamics specified in equation (2.11):[25]

$$F_t - \psi B(m)\beta F_t \tag{2.12}$$

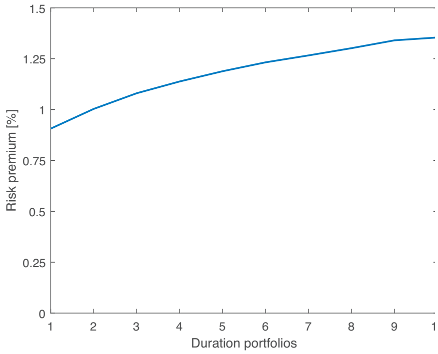
The first term in the equation above represents the short-term risk premium, which in frameworks like Driessen (2005) captures the default event (jump) risk premium. The second component captures the systematic variation in the default event risk premia over time. This drift term is convenient for analyzing the term structure as it brings the holding-to-maturity expected returns to frequencies that are much closer to the monthly level. Moreover, the risk premium expression is analytically parsimonious as all the cross-maturity risk premia differences will be driven by the second term, $\psi B(m)\beta F_t$.

The right panel of Figure IV summarizes the average instantaneous risk premia term structure over the last 20 years. The term structure is upward-sloping which qualitatively matches the slope of holding-to-maturity risk premia. However, the slope is substantially smaller (roughly 45 basis points) than in the baseline, and this comes from the fact that a one-factor model struggles to simultaneously match highly volatile short-term risk and quite persistent long-term risk dynamics. In fact, most of the extreme variation during the 2008-2009 financial crisis is attributed to the measurement noise, which leads to higher filtered short-term risk dynamics. Thus, the affine model confirms the key finding that the risk premia is generally increasing with maturity, and the holding period is likely not driving the difference between implied and realized returns.

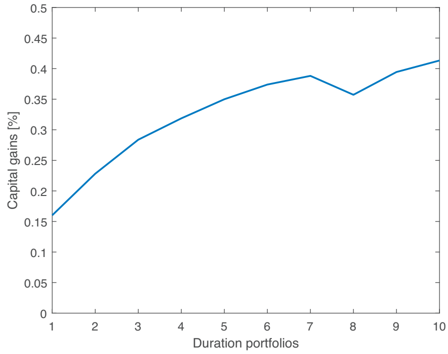
Finally, the affine model estimates allow us to pin down the capital gains that arise purely from the shifts in risk premia over the sample. One way to quantify this effect is by calculating the yearly capital gains driven by the risk premium dynamics alone (i.e., while keeping expected losses and risk-free rates constant). The right plot of Figure IV presents these estimates. Over the sample period, there was a noticeable decline in the short-term risk premium, which generated some capital gains on both short-term (0.2%) and long-term bonds (0.4%). The reduction in the short-term risk premium is greater. Still, the duration magnifies this effect on long-term bonds much more, leading to positive capital gains on all bonds (which qualitatively matches the excess return patterns after the financial crisis). Thus, the one-factor model alone cannot generate downward-sloping capital gains in realized returns. However, the results do support the idea that the realized returns are noisy, and the capital gains can constitute 30% of the actual risk premia (even averaged over 20 years of data). Once we allow for the expected default losses to vary as well, the informativeness of average realized returns can be substantially smaller.

[25] More detailed derivations are provided in Appendix 2.E.2.

Figure IV: Risk premia and capital gains



(a) Instantaneous risk premia term structure



(b) Realized capital gains

Note: The figures above summarizes the instantaneous risk premia term structure (left plot) and the average capital gains driven by the decline in risk premia for each duration portfolio over the 2002-2020 sample (right plot).

2.7 Conclusions

This paper documents a puzzle in corporate bond market returns. On the one hand, the expected returns constructed from yields, in the spirit of Campello et al. (2008), and expected default probability proxies deliver a positive risk premia and duration relationship both at the aggregate and for most cross-sectional sorts. Only the lowest credit rating firms display a flat term structure which indicates their high exposure to short-term risks. In contrast, over the same sample period (2002-2020), the short-duration bond portfolios generated exceptionally high risk-adjusted realized returns, whereas the long-term bonds featured large negative alphas. This dichotomy is also found in the equities where model-based risk premia slope (Bansal et al. (2021); Baele et al. (2022)) often contradicts the realized return patterns (van Binsbergen & Koijen (2017)).

There are indications showing that the risk premia slope is time-varying and potentially creating this wedge between the realizations and expectations. Different sample splits indicate the steepening of the risk premia curve after the financial crisis. At the same time, the realized returns in duration portfolios turned in the opposite direction. However, I show that the capital revaluation gains based on the yields-based estimates are too small to fully explain the stark mismatch in the realized return alphas. In fact, the average excess returns often far exceed the credit spreads, which serve as the upper bound for the holding-to-maturity risk premia, posing a theoretical challenge to match such results. Hence, one has to be cautious in testing asset pricing models based on the last 20 years of data alone.

Despite the disagreement on the slope of the risk premia, the analysis highlights a number of cross-sectional patterns. Both the yields-based and realized returns detect the credit risk, size and value risk premia. Moreover, this factor premium is not subsumed by the duration dimension alone, as argued by Gormsen & Lazarus (2019). Most of the abnormal returns are generated by the short- and intermediate-maturity bonds, hinting at the origin of the factor premia and the need to deeper analyze such horizon risks.

References

Andrews, S. & Gonçalves, A. (2020), 'The bond, equity, and real estate term structures', *Kenan Institute of Private Enterprise Research Paper Forthcoming* .

Baele, L., Driessen, J. & Jankauskas, T. (2022), The implied equity term structure. Available on SSRN.

Bai, J., Bali, T. G. & Wen, Q. (2019), 'Common risk factors in the cross-section of corporate bond returns', *Journal of Financial Economics* **131**(3), 619–642.

Bali, T. G., Subrahmanyam, A. & Wen, Q. (2021), 'Long-term reversals in the corporate bond market', *Journal of Financial Economics* **139**(2), 656–677.

Bansal, R., Miller, S., Song, D. & Yaron, A. (2021), 'The term structure of equity risk premia', *Journal of Financial Economics* **142**(3), 1209–1228.

Bansal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The Journal of Finance* **59**(4), 1481–1509.

Bao, J. (2009), Structural models of default and the cross-section of corporate bond yield spreads. Available on SSRN.

Bartram, S. M., Grinblatt, M. & Nozawa, Y. (2021), Book-to-market, mispricing, and the cross-section of corporate bond returns. Available on SSRN.

Bhamra, H. S., Kuehn, L.-A. & Strebulaev, I. A. (2010), 'The aggregate dynamics of capital structure and macroeconomic risk', *The Review of Financial Studies* **23**(12), 4187–4241.

Black, F. & Cox, J. C. (1976), 'Valuing corporate securities: Some effects of bond indenture provisions', *The Journal of Finance* **31**(2), 351–367.

Bongaerts, D., De Jong, F. & Driessen, J. (2017), 'An asset pricing approach to liquidity effects in corporate bond markets', *The Review of Financial Studies* **30**(4), 1229–1269.

Campbell, J. Y. & Cochrane, J. H. (1999), 'By force of habit: A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* **107**(2), 205–251.

Campbell, J. Y., Hilscher, J. & Szilagyi, J. (2008), 'In search of distress risk', *Journal of Finance* **63**(6), 2899–2939.

Campello, M., Chen, L. & Zhang, L. (2008), 'Expected returns, yield spreads, and asset pricing tests', *The Review of Financial Studies* **21**(3), 1297–1338.

Chen, L., Collin-Dufresne, P. & Goldstein, R. (2009), 'On the relation between the credit spread puzzle and the equity premium puzzle', *The Review of Financial Studies* **22**(9), 3367–3409.

Corsi, F. (2009), 'A simple approximate long-memory model of realized volatility', *Journal of Financial Econometrics* **7**(2), 174–196.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–407.

de Jong, F. (2000), 'Time series and cross-section information in affine term-structure models', *Journal of Business and Economic Statistics* **18**(3), 300–314.

Dick-Nielsen, J. (2013), How to clean enhanced trace data. Available on SSRN.

Driessen, J. (2005), 'Is default event risk priced in corporate bonds?', *The Review of Financial Studies* **18**(1), 165–195.

Duffie, D. & Singleton, K. J. (1999), 'Modeling term structures of defaultable bonds', *The Review of Financial Studies* **12**(4), 687–720.

Fama, E. F. & French, K. R. (2002), 'The equity premium', *The Journal of Finance* **57**(2), 637–659.

Feldhütter, P. & Schaefer, S. M. (2018), 'The myth of the credit spread puzzle', *The Review of Financial Studies* **31**(8), 2897–2942.

Gabaix, X. (2012), 'Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance', *The Quarterly Journal of Economics* **127**(2), 645–700.

Giglio, S., Kelly, B. T. & Kozak, S. (2020), Equity term structures without dividend strips data. Available on SSRN.

Gonçalves, A. S. (2021), 'Reinvestment risk and the equity term structure', *The Journal of Finance, forthcoming* .

Gormsen, N. J. (2021), Time variation of the equity term structure.

Gormsen, N. J. & Lazarus, E. (2019), Duration-driven returns. Available on SSRN.

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2007), 'Treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**(8), 2291–2304.

Hasler, M. & Marfe, R. (2016), 'Disaster recovery and the term structure of dividend strips', *Journal of Financial Economics* **122**(1), 116–134.

Huang, J. Z. & Huang, M. (2012), 'How much of the corporate-treasury yield spread is due to credit risk?', *The Review of Asset Pricing Studies* **2**(2), 153–202.

Huang, J.-Z., Nozawa, Y. & Shi, Z. (2022), 'The global credit spread puzzle', *PBCSF-NIFR Research Paper* .

Hull, J. (2009), *Options, futures, and other derivatives*, Pearson Prentice-Hall.

Hull, J. & White, A. (1987), 'The pricing of options on assets with stochastic volatilities', *The Journal of Finance* **42**(2), 281–300.

Jarrow, R., Lando, D. & Yu, F. (2000), 'Default risk and diversification: Theory and applications', *Mathematical Finance* .

Jostova, G., Nikolova, S., Philipov, A. & Stahel, C. W. (2013), 'Momentum in corporate bond returns', *The Review of Financial Studies* **26**(7), 1649–1693.

Lettau, M. & Wachter, J. A. (2007), 'Why is long-horizon equity less risky? a duration-based explanation of the value premium', *The Journal of Finance* **62**(1), 55–92.

Li, Y., Ng, D. T. & Swaminathan, B. (2013), 'Predicting market returns using aggregate implied cost of capital', *Journal of Financial Economics* **110**(2), 419–436.

Merton, R. C. (1974), 'On the pricing of corporate debt: The risk structure of interest rates', *The Journal of Finance* **29**(2), 449–470.

Moody's Investors Service (2021), Annual default study: Following a sharp rise in 2020, corporate defaults will drop in 2021.

Mueller, P., Vedolin, A. & Zhou, H. (2019), 'Short-run bond risk premia', *Quarterly Journal of Finance* **9**(03), 1950011.

Shiller, R. (1981), 'Do stock prices move too much to be justified by subsequent changes in dividends?', *The American Economic Review* **71**(3), 421–436.

van Binsbergen, J., Brandt, M. & Koijen, R. (2012), 'On the timing and pricing of dividends', *American Economic Review* **102**(4), 1596–1618.

van Binsbergen, J. H. & Schwert, M. (2021), Duration-based valuation of corporate bonds. Available on SSRN.

van Binsbergen, J., Hueskes, W., Koijen, R. & Vrugt, E. (2013), 'Equity yields', *Journal of Financial Economics* **110**(3), 503–519.

van Binsbergen, J. & Koijen, R. (2017), 'The term structure of returns: Facts and theory', *Journal of Financial Economics* **124**(1), 1–21.

van Zundert, J. & Driessen, J. (2017), Are stock and corporate bond markets integrated? evidence from expected returns. Available on SSRN.

Weber, M. (2018), 'Cash flow duration and the term structure of equity returns', *Journal of Financial Economics* **128**(3), 486–503.

Yara, F. B., Boons, M. & Tamoni, A. (2020), New and old sorts: Implications for asset pricing. martijn and tamoni, andrea, new and old sorts: Implications for asset pricing. Available on SSRN.

Yu, F. (2002), 'Modeling expected return on defaultable bonds', *The Journal of Fixed Income* **12**(2), 69–81.

Appendices

2.A Additional Tables and Figures

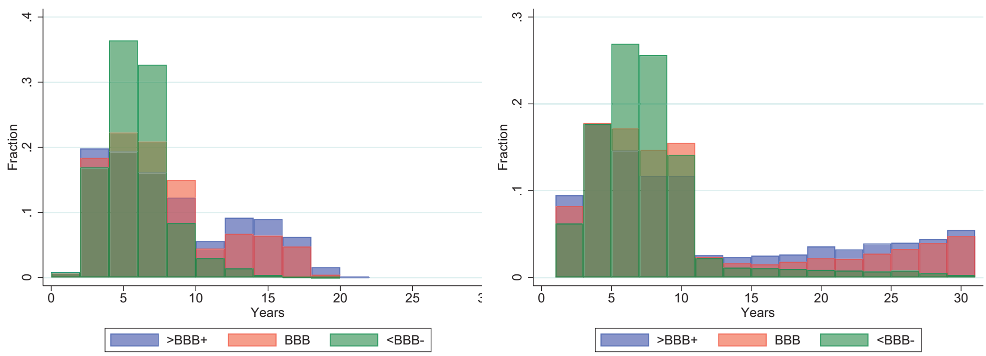
Table A.I: Credit ratings

The table below summarizes the distribution of bond-month observations across credit rating groups and duration.

Rating	N	prc.	Rating group
Aaa	4,255	0.9%	
Aa1	4,967	1.0%	
Aa2	10,044	2.0%	group 1
Aa3	10,284	2.1%	40.9%
A1	31,459	6.4%	
A2	72,200	14.7%	
A3	67,373	13.7%	
Baa1	72,210	14.7%	
Baa2	55,687	11.4%	group 2
Baa3	79,848	16.3%	42.4%
Ba1	19,819	4.0%	
Ba2	16,707	3.4%	
Ba3	16,993	3.5%	
B1	10,977	2.2%	
B2	7,908	1.6%	group 3
B3	5,136	1.0%	16.7%
Caa1	1,910	0.4%	
Caa2	1,164	0.2%	
Caa3	146	0.0%	
Ca	661	0.1%	
C	632	0.1%	
D	105	0.0%	
Total	490,485	100.0%	

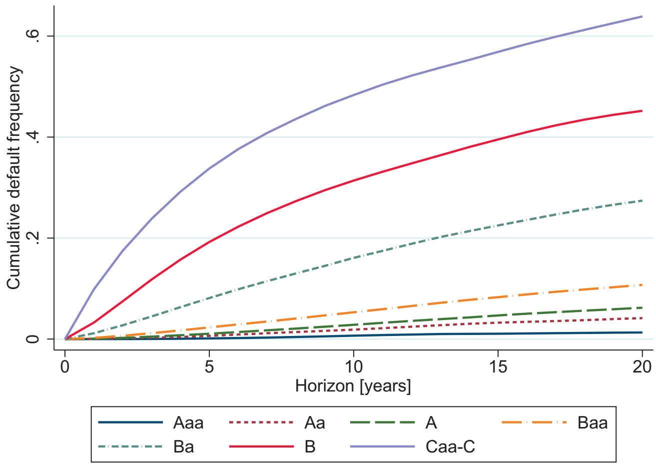
Duration	group 1	group 2	group 3
0	36	2	161
1	1,469	1,420	477
2	19,160	17,675	5,477
3	20,296	20,333	8,271
4	19,781	22,923	13,045
5	19,119	23,283	16,909
6	15,802	21,419	16,028
7	16,598	22,395	10,997
8	15,917	21,600	5,252
9	8,869	9,511	1,598
10	5,088	3,826	1,165
11	5,663	5,001	1,285
12	8,338	6,321	936
13	9,517	7,535	232
14	11,070	7,429	187
15	7,317	6,155	86
16	6,661	6,336	56
17	6,338	3,774	13
18	2,960	844	1
19	405	60	1
20	51	3	-
21	8	-	-
Total	200,463	207,845	82,177

Figure A.I: Bond-month observations across duration and maturity

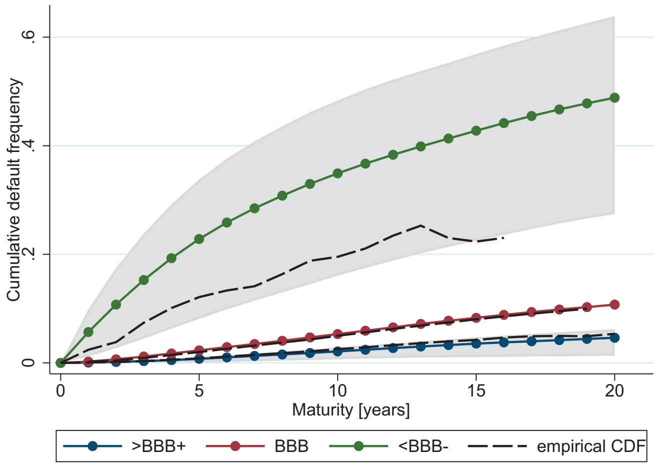


(a) Duration (b) Maturity

Figure A.II: Historical default frequencies in credit groups



(a) Credit rating categories



(b) Aggregated credit rating groups

Note: The figures above present the historical default frequencies based on Moody's Investors Service (2021). The first graph depicts the default frequencies in each credit rating group, whereas the second one aggregates it into the three groups used in this paper. The gray areas denote the minimal and maximal default frequencies in each aggregated groups. The black dashed line depicts the average expected default in the sample for each group.

Table A.II: Duration

The table below summarizes the average bond duration (expressed in years) in each duration decile at the time of portfolio formation.

Full sample (2002.06-2020.06)										
Low	2	3	4	5	6	7	8	9	High	High-Low
1.3	2.4	3.4	4.2	5.1	6.0	7.1	8.8	11.5	14.1	12.7
Before financial crisis (2002.06-2007.12)										
1.4	2.3	3.2	4.0	4.9	5.8	6.7	7.9	10.3	12.7	11.3
After financial crisis (2009.06-2020.06)										
1.3	2.5	3.4	4.3	5.2	6.2	7.3	9.4	12.4	15.1	13.8
Financial crisis (2007.12-2009.06)										
1.5	2.6	3.4	4.1	4.9	5.8	6.7	8.0	10.2	12.4	10.9

Table A.III: Duration of cross-sectional sorts

The table below summarizes the average bond duration (expressed in years) in each duration quintile at the time of portfolio formation.

Book-to-Market sorts						
	Low	2	3	4	High	High-Low
Growth	1.8	3.8	5.5	7.7	12.5	10.6
Value	1.8	3.8	5.5	8.1	12.5	10.7
Size sorts						
Small	1.9	3.8	5.2	6.9	11.0	9.1
Big	1.8	3.8	6.1	9.5	13.5	11.7
Leverage sorts						
Low	1.8	3.8	5.7	8.2	12.9	11.1
High	1.8	3.7	5.4	7.9	12.2	10.4
Credit rating sorts						
>BBB+	1.7	3.9	6.2	9.5	13.6	11.8
=BBB	1.9	3.9	5.8	8.1	12.3	10.5
<BBB-	2.0	3.6	4.7	5.7	8.4	6.4
Liquidity sorts						
Low	2.1	4.2	6.0	8.6	12.6	10.4
High	1.6	3.4	5.0	7.2	12.4	10.8

Table A.IV: Number of bonds

The table below summarizes the average number of bonds in each duration in each duration decile and bin. The duration bins are set to equal to 2, 3.5, 5, .., 15+ years.

	Low	2	3	4	5	6	7	8	9	High
Full sample (2002.06-2020.06)										
deciles	72	137	147	160	175	184	185	149	137	191
bins	78	219	282	278	245	78	46	68	96	145
Before financial crisis (2002.06-2007.12)										
deciles	27	58	60	66	77	91	105	105	77	93
bins	38	99	123	143	159	46	35	58	48	8
After financial crisis (2009.06-2020.06)										
deciles	95	178	193	208	226	231	225	177	171	245
bins	101	285	361	347	290	100	51	71	118	224
Financial crisis (2007.12-2009.06)										
deciles	37	81	92	106	118	129	141	84	85	113
bins	32	122	210	204	189	26	44	76	80	4

Table A.V: Number of bonds in each cross-sectional sort

The table below summarizes the average number of bonds in each duration in each duration quintile and bin. The duration bins are set to equal to 3, 6, 9, 12, and 15+ years.

	Low	2	3	4	High
Book-to-Market sorts					
Growth	86	138	150	174	149
Value	80	124	150	147	144
Size sorts					
Small	115	176	211	206	153
Big	87	126	135	141	185
Leverage sorts					
Low	120	177	196	225	211
High	82	132	157	126	110
Credit rating sorts					
>BBB+	70	103	106	111	142
=BBB	92	137	150	170	152
<BBB-	38	64	89	59	47
Liquidity sorts					
Low	72	133	164	164	142
High	125	170	182	181	180

(a) Duration quintiles

	Low	2	3	4	High
Book-to-Market sorts					
Growth	97	246	195	35	123
Value	88	234	156	53	114
Size sorts					
Small	121	368	242	47	83
Big	95	183	159	55	182
Leverage sorts					
Low	129	305	255	57	184
High	87	246	146	46	81
Credit rating sorts					
>BBB+	79	146	128	40	140
=BBB	99	232	198	48	122
<BBB-	37	170	73	14	3
Liquidity sorts					
Low	74	249	179	69	104
High	139	294	216	30	157

(b) Duration bins

Table A.VI: Credit spreads

The table below summarizes the average credit spread in each duration quintile. The duration bins are set to equal to 3, 5, 8, 12, and 15+ years.

	Low	2	3	4	High	High-Low
Credit rating sorts						
>BBB+	1.04	1.12	1.40	1.63	1.63	0.58
=BBB	1.50	1.76	2.07	2.14	2.29	0.79
<BBB-	4.62	5.55	4.93	4.65	4.21	-0.41
Leverage sorts						
Low	1.19	1.55	1.93	1.71	1.76	0.57
High	2.64	3.52	3.74	2.77	2.36	-0.27
Book-to-Market sorts						
Growth	1.40	1.94	2.38	1.83	1.78	0.38
Value	2.29	3.09	3.30	2.44	2.14	-0.15
Size sorts						
Small	2.74	3.51	3.69	2.72	2.74	0.00
Big	1.06	1.19	1.46	1.72	1.74	0.68
Liquidity sorts						
Low	2.57	2.84	2.77	2.33	2.17	-0.39
High	1.49	2.10	2.84	1.93	1.82	0.34

Table A.VII: Cross-sectional betas

The table below summarizes the average credit spread in each duration quintile. The duration bins are set to equal to 3, 5, 8, 12, and 15+ years.

	Low	2	3	4	High	High-Low
Credit rating sorts						
>BBB+	0.28***	0.47***	0.79***	0.94***	1.51***	1.22***
BBB	0.31***	0.60***	0.82***	1.03***	1.54***	1.23***
<BBB-	0.83***	1.60***	1.47***	1.60***	1.65***	0.80***
Leverage						
Low LEV	0.21***	0.50***	0.73***	0.82***	1.34***	1.13***
High LEV	0.63***	1.17***	1.38***	1.41***	1.88***	1.25***
Book-to-market sorts						
Growth	0.22***	0.54***	0.91***	0.89***	1.33***	1.10***
Value	0.51***	0.92***	1.19***	1.06***	1.52***	1.01***
Size sorts						
Small	0.55***	1.12***	1.29***	1.24***	1.50***	0.95***
Big	0.23***	0.44***	0.69***	0.99***	1.49***	1.25***

Table A.VIII: Trading activity

The table below summarizes the average number of monthly observations.

Duration quintile	1	2	3	4	5
Trading frequency	Months per year traded				
last 5 days	6	6	7	7	8
once a month	7	8	8	9	10
	BM high				
last 5 days	6	7	7	8	8
once a month	8	8	8	9	10
	BM low				
last 5 days	6	7	7	8	8
once a month	7	8	9	9	10
	Size high				
last 5 days	6	7	7	7	8
once a month	7	8	9	9	10
	Size low				
last 5 days	6	7	7	7	8
once a month	7	8	8	9	10
	>BBB+				
last 5 days	6	6	6	7	7
once a month	7	8	8	9	10
	BBB				
last 5 days	6	7	7	7	8
once a month	7	8	8	9	10
	<BBB-				
last 5 days	5	6	7	8	7
once a month	7	8	9	9	9

Table A.IX: Theoretical model predictions for equities

The table below presents the key predictions for equity term structure slope and its cyclicality. The table is adapted from Gormsen (2021).

Theories	Average slope	Cyclicality of Term Premia
Habit (Campbell & Cochrane (1999))	Upward	Countercyclical
Long-run risk (Bansal & Yaron (2004))	Upward	Countercyclical
Rare disasters (Gabaix (2012))	Flat	Constant
Rare disaster with quick recoveries (Hasler & Marfe (2016))	Downward	Procyclical

2.B Data Appendix

2.B.1 Data Filters

To clean the bond transaction data, these filters are applied:

1. Only the USD-nominated corporate bonds are kept.

2. Following Bongaerts, De Jong & Driessen (2017), I drop putable, perpetual and convertible bonds. Similarly, I keep only senior unsecured bonds with no additional guarantees. Furthermore, I select only bonds with fixed coupons and fixed coupon payment frequencies (i.e. bonds with 1, 3, 6, 12 month coupon frequencies).

3. To reduce the influence of potential errors and outliers, I drop all bonds that are traded for less than $5 of their principal or that have negative coupons. Since very short-maturity bonds are traded less often and, in this region, the Gürkaynak et al. (2007) risk-free rate estimates are not very reliable, I also exclude all bonds with shorter than 3-month maturity. Based on similar arguments, I remove from the sample all bonds with maturities exceeding 30 years.

Table A.X: Data filters

The table below summarizes the sample filters applied.

	N
TRACE enhanced (2002-2020)	3,202,692
Merge with FISD & WRDS masterfile	3,133,512
1. corporate bonds USD nom	3,129,540
-putable bonds	67,553
-convertible bonds	50,397
-perpetual bonds	1,194
2. senior unsecured bonds	2,455,874
-guaranteed	409,779
3. standard fixed-coupon bonds	1,617,866
-trim prices at 1 (8.5) and 99% (147.2)	32,269
- missing or negative coupon	10,311
- <4month and >30 year maturities	54,079
Final TRACE sample (quarterly prices)	1,521,207
Final TRACE sample (monthly prices)	1,125,092

2.B.2 Accrued interest and yields

In order to calculate returns and yields, I adjust the clean prices reported in TRACE with accrued interest. Namely, I follow these steps to compute the accrued interest:

- I assume that all bonds follow the US 30/360 day count convention (which is true for the absolute majority of US corporate bonds).
- Based on FISD data on the last coupon payment date (before maturity), I find the most recent coupon date. This exercise requires the knowledge of coupon frequency which is assumed to be semi-annual if such information is missing. Due to potential errors and data irregularities, some of these FISD coupon dates exceed the maturity date. In those cases (as well as for missing observations), I assume that the last coupon is paid at maturity.
- Using this most recent coupon date, I calculate the accrued interest for the last day of the trading month.

For most bonds, FINRA reports bond yields. However, since the exact methodology is unpublished, I calculate the yields based on available information on bond prices, coupons, coupon frequency, time-to-maturity and coupon payment dates. In most cases, the correlation between FINRA reports and my estimates is high, however, there are discrepancies that may arise from different assumptions and prices used (i.e., I aggregate all intra-day prices).

2.C Merton's Model

2.C.1 Modeling assumptions

In the Merton (1974) model, one can express the real world default probabilities in the following way:

$$\pi_t = N\left(-\frac{log(\frac{V_t}{D_T}) + (\mu_{A,t} - \delta_t - 0.5\sigma_{A,t}^2)T}{\sigma_{A,t}\sqrt{T}}\right) \tag{13}$$

where V_t is the value of firm's assets, D_T is the face value of debt that matures at T, $\mu_{A,t}$ and $\sigma_{A,t}$ are the expected return and volatility of assets, δ_t is the payout rate to all company's stakeholders, and $N(.)$ is the CDF of standard normal distribution. The unknown parameters are $\mu_{A,t}$ and $\sigma_{A,t}$, and V_t is endogenously determined by the Black-Scholes formula.

There are many ways to proceed from here and calibrate $\mu_{A,t}$ and $\sigma_{A,t}$. A standard approach is to apply the Ito's lemma and express $\sigma_{A,t}$ in terms of measurable equity volatility $\sigma_{E,t}$[26]:

$$\sigma_{E,t} = \left(\frac{V_t}{E_t}\right) N(d_1)\sigma_{A,t} \tag{14}$$

where E_t is equity value of the firm and d_1 comes from the Black-Scholes formula (r_t is the risk-free rate):

$$d_1 = \frac{log(\frac{V_t}{D_T}) + (r_t - 0.5\sigma_{A,t}^2)T}{\sigma_{A,t}\sqrt{T}} \tag{15}$$

Based on the equations above, one can solve for $\sigma_{A,t}$ and V_t which determine the risk-neutral default probabilities. In order to compute the real world default frequencies, as in (13), we need to specify $\mu_{A,t}$. Following van Zundert & Driessen (2017) (ZD hereafter) and Feldhütter & Schaefer (2018) (FS), I assume that the risk premium on assets equals to a constant price of risk (θ) times the volatility of the firm's assets ($\sigma_{A,t}$), where $\theta = 0.22$ (based on Chen et al. (2009) estimates). In other words, this assumption reduces the two-parameter model to a single-parameter specification. Needless to say, the modeling of $\sigma_{E,t}$ becomes crucial.

Similarly as in van Zundert & Driessen (2017), I model the volatility of equity as a mean reverting process with one common autoregressive coefficient (ρ) across firms, and long-term mean (μ) specific to each credit rating category. This aims to correct for the fact that, empirically, extreme short-term volatilities converge over time to more moderate equilibrium levels. In turn, the reversion captures the pattern that short-term dynamics are less relevant for firms with the long-term rather than short-term debt[27]. In order to

[26]The formula implicitly assumes that the leverage will remain constant.

[27]Strictly speaking, the Merton's model does not embed time variation in its framework. Thus, all the extensions serve as an approximation of the true model. However, under certain independence assumptions, the equity volatility process can be modeled independently of the assets (Hull & White

estimate ρ and μ, I run the following pooled OLS regression:

$$\hat{\sigma}_{E,t,12} = \beta_0 + \beta_1 \hat{\sigma}_{E,t-12,12} + \epsilon_t \tag{16}$$

where $\hat{\sigma}_{E,t,12}$ is an equity volatility at time t estimated with daily returns over the last year and β_1 captures the annualized mean-reversion parameter (ρ). Based on the CRSP daily returns sample over 2002-2020, the estimate of ρ equals 0.749. An alternative specification with shorter time windows of 1 month (Figure A.XI column 1) delivers lower R-square estimates (0.301 vs 0.569). Simultaneously, Corsi (2009) type of models with lagged slow-moving volatilities of 3 and 12 months (columns 2 and 3) offer a slightly higher predictive power. However, since the difference is not substantial, I keep the parsimonious specification (16) as a baseline model.

The long-term mean (μ) is recovered by simply averaging all monthly standard deviations in each credit rating group over the whole sample length. This procedure results in these estimates: $\mu = 0.254$ for firms with ratings A and higher, $\mu = 0.259$ for BBB firms, and $\mu = 0.368$ for firms with rating BB and below.

Table A.XI: Parameter estimates

The table below summarizes the volatility mean-reversion parameter estimates for each credit rating group. The volatilities are winsorized at 2.5 and 97.5% levels.

	(1)	(2)	(3)	(4)
	$\sigma_{t,1}$	$\sigma_{t,1}$	$\sigma_{t,1}$	$\sigma_{t,12}$
$\sigma_{t-12,1}$	0.552^{***}			
$\sigma_{t-1,1}$		0.262^{***}	0.263^{***}	
$\sigma_{t-1,3}$		0.954^{***}	0.554^{***}	
$\sigma_{t-1,12}$			0.900^{***}	
$\sigma_{t-12,12}$				0.749^{***}
N	1171242	1258431	1225113	1132526
R^2	0.301	0.619	0.633	0.569

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Based on this, I can substitute the weighted average variance into (14):

$$\bar{\sigma}_{E,t,T} = \sqrt{\frac{1}{T}\sum_{s}^{T} \sigma_{E,s}^2} \quad \text{, where} \quad \sigma_{E,s} = (1-\rho)\mu + \rho\sigma_{E,s-1} \tag{17}$$

Consequently, I can solve for $\mu_{A,t}$ and $\sigma_{A,t}$. To prevent outliers, $\sigma_{A,t}$ is assumed to be at least 20% of the forecasted equity volatility ($\sigma_{A,T} \geq 0.2\bar{\sigma}_{E,T}$). In order to obtain π_t at the bond level, I substitute the bond maturity into (13).

(1987)).

2.C.2 Black-Cox model

An extension of the Merton's model by Black & Cox (1976) allows for the early defaults that may occur if the firm's asset value falls below a default threshold determined by the debt covenants. Accordingly, the default probability (15) adjusts to:

where d is a parameter capturing the default boundary and all other variables are as in (15) (more details about the Black-Cox model are provided in Bao (2009)).

In order to estimate d and implement this framework empirically, I closely follow Feldhütter & Schaefer (2018). Namely, $\mu_{A,t}$ is modeled exactly the same way as in the Merton's model, whereas $\sigma_{A,t}$ is parameterized as a function of equity volatility, leverage and some adjustment factor:

$$\sigma_{A,t} = \frac{E_t}{D_T + E_t}\sigma_{E,t}c\,(.) = (1 - LEV_t)\sigma_{E,t}c\,(.) \tag{18}$$

where $c(.)$ attains a value of 1 if $LEV_t < 0.25$, 1.05 if $0.25 < LEV_t \leq 0.35$, 1.10 if $0.35 < LEV_t \leq 0.45$, 1.20 if $0.45 < LEV_t \leq 0.55$, 1.40 if $0.55 < LEV_t \leq 0.75$ and 1.80 if $0.75 \geq LEV_t$ (Feldhütter & Schaefer (2018)). Since the step function significantly suppresses the volatility for highly leveraged firms, additionally, I impose that $\sigma_{A,t} \geq 0.2\sigma_{E,t}$.

Later, within each credit rating subcategory (Aaa-Aa, A, Baa, Ba, B, Caa-C), I create a representative firm which takes the median values of $\mu_{A,t}$, $\sigma_{A,t}$, δ_t, LEV_t. This allows me to calibrate d by matching model implied default probabilities from (2.3) to historical default frequencies (over 1920-2020). In other words, I minimize the squared absolute deviation in all credit rating groups (Aaa-Caa) across maturities (1-20):

$$\hat{d}_i = argmin\frac{1}{20}\sum_{t=1}^{20}(\pi_{i,t} - \hat{\pi}_{i,t})^2 \qquad \text{for each } i \in \{\text{Aaa-A,...,Caa-C}\}. \tag{19}$$

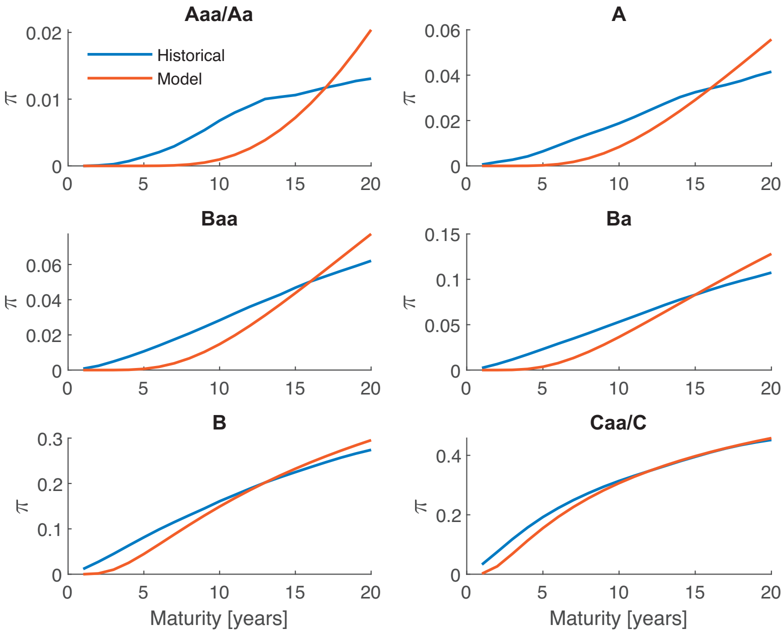
where $\pi_{i,t}$ is a historical Moody's estimate of default probability in rating i and maturity t, and $\hat{\pi}_{i,t}$ is the respective Black-Cox default probability for a representative firm in rating i. The calibration delivers a default barrier estimate ranging from 0.688 to 0.934 which encompasses Feldhütter & Schaefer (2018) estimate of 0.8944. Table A.XII outlines alternative calibration methods (imposing a single default threshold or changing the historical sample) all of which produce fairly similar estimates. Given that the mean model fit error is substantially lower (1.4% vs 2.8%) in the scenario with credit rating specific default thresholds, I adopted this baseline for the calculation of resulting default probabilities. The model fit to the historical probabilities is presented below (Figure A.III).

Table A.XII: Black-Cox model implementation

The table below summarizes the median model statistics for each credit rating group (top panel), as well as the default threshold estimates (bottom panel). Columns a) and c) assume one single default barrier for all firms, whereas columns b) and d) allow for a separate parameter for each credit rating category.

		Median statistics		
Credit rating	LEV	μ	σ	δ
Aaa-Aa	0.479	0.105	0.055	0.050
A	0.406	0.133	0.057	0.044
Baa	0.397	0.128	0.053	0.043
Ba	0.448	0.174	0.064	0.048
B	0.562	0.167	0.062	0.044
Caa-C	0.883	0.136	0.063	0.049
		Default threshold $\hat{d}$		
Credit rating	a) 1920-2020	b) 1920-2020	c) 1983-2020	d) 1983-2020
Aaa-Aa	0.773	0.688	0.781	0.488
A	0.773	0.846	0.781	0.695
Baa	0.773	0.934	0.781	0.880
Ba	0.773	0.697	0.781	0.621
B	0.773	0.860	0.781	0.839
Caa-C	0.773	0.764	0.781	0.782
MSE	0.028	0.014	0.026	0.012

Figure A.III: Model fit



Note: The figures illustrate the Black-Cox model fit to historical Moody's default probabilities (over 1920-2020).

2.C.3 Modeling inputs

Equations (13), (14) and (15) require a number of data inputs which imposes some restrictions on the sample size. Table A.XIII highlights the key components and their data source. Some subjectivity is involved in choosing the default threshold - on e could argue that only the long-term debt that matters. On the other hand, the Merton's model underestimates the default risk by assuming that default occurs only at maturity, as well as treating all firm's assets as perfectly liquid. Thus, having more stringent default thresholds can alleviate this issue. Secondly, there may be concerns that the maturity of bonds may poorly proxy the total firm's debt. However, as we see in Table X, the public debt constitutes the majority of all debt for most firms in the sample.

Table A.XIII: Model inputs

The table below summarizes the choice of input variables in the Merton (1974) model. Except for T and r_t, all empirical choices are made following van Zundert & Driessen (2017) and Feldhütter & Schaefer (2018).

Input	Proxy
E	**Equity value:** most recent market capitalization based on CRSP ($\|prc\|$*$shrout/1000$)
D_T	**Default threshold:** sum of short-term debt and long-term debt (quarterly Compustat items *dlcq* and *dlttq*)
μ	**Asset drift rate:** $r_T + 0.22\sigma$
σ^E	**Equity volatility:** estimated based on daily returns from CRSP and matched to the debt maturity, T, using the autoregressive process with an AR(1) coefficient and long-term means, as in Table A.XI
δ	**Payout rate:** the sum of interest payments, dividends and net stock repurchases. Annual Compustat items *intpn, dv and prstkc.*
r_t	**Risk-free rate:** nominal US government zero coupon yields with duration of T years (from Gürkaynak et al. (2007))
T	**Debt maturity:** principal-weighted average maturity of company's public debt

Table A.XIV: Model inputs

The table below summarizes the distributions of firm-month level inputs of the structural models. The bottom panel presents the implied asset volatility and expected return estimates based on the Feldhütter & Schaefer (2018) Black-Cox implementation (FS2018), van Zundert & Driessen (2017) methodology with stochastic volatility (ZD2017), and the standard Merton's model. The sample spans 2002-2020.

IBES-Compustat 1980-2021						
	p5	median	p95	mean	std	N
Sales	22	583	13846	3597	13800	50712
Book equity	27	314	7102	1894	8054	50712
Market equity	71	773	20633	5578	29186	50712
Sales growth	-0.19	0.10	0.78	0.17	0.29	49036
ROE	-0.44	0.11	0.40	0.08	0.23	48161
B/M	0.08	0.42	1.37	0.55	0.61	50712
E/P	-0.19	0.04	0.13	0.00	0.31	50712
Payout	0.00	0.00	0.86	0.16	0.27	50712
Compustat 1980-2021						
	p5	median	p95	mean	std	N
Sales	6	254	9264	2401	11167	79537
Book equity	8	139	4511	1268	6515	79537
Market equity	19	317	12737	3665	23464	79537
Sales growth	-0.21	0.10	0.78	0.16	0.30	73370
ROE	-0.47	0.10	0.39	0.06	0.24	72207
B/M	0.08	0.46	1.59	0.61	0.77	79537
E/P	-0.24	0.04	0.15	0.00	0.40	79537
Payout	0.00	0.00	0.85	0.15	0.26	79537

2.D Kalman Filter and Quasi-Maximum Likelihood

2.D.1 Kalman Filter implementation

Following de Jong (2000), I translate the dynamics of a short-term risk factor to the state-space representation. There are M duration portfolios observable each month (in my baseline, I use 10 duration deciles). The measurement equation for their risk premia is:

$$rp_t = A + BF_t + e_t, \quad \text{var}(e_t) = H \tag{20}$$

where A and B are solutions based on equation (2.11), e_t is a vector of measurement errors that I assume to be homoscedastic and independent across time and maturities of bonds. As a result, H is a diagonal covariance matrix:

$$H = I_M \sigma_H^2 \tag{21}$$

where I_M is an $M \times M$ identity matrix, and σ_H^2 is the variance of errors. Note that for government bonds the measurement error captures the data noisiness and transient market deviations from the fundamental default-free price. In my setting, this error will capture both the noise in risk-free rates and expected loss estimates.

The transition equation is:

$$F_{t+\Delta t} = \Phi F_t + \eta_t, \quad \text{var}(\eta_t) = Q_t \tag{22}$$

where Φ and Q_t come from the dynamics of state variable F, and $\Delta t = 1/12$ since the returns are measured at a monthly frequency. For a square root process, the dynamics of F are:

$$F_{t+\Delta t} - \theta = e^{-\kappa \Delta t}(F_t - \theta) + \int_0^{\Delta t} e^{-\kappa(\Delta t - s)} \beta F_{t+s} dW_{t+s} \tag{23}$$

which implies that:

$$\begin{aligned} \Phi &= e^{-\kappa \Delta t} \\ Q_t &= \frac{1 - e^{-2\kappa \Delta t}}{2\kappa} \beta\theta + \frac{e^{-\kappa \Delta t} - e^{-2\kappa \Delta t}}{\kappa} \beta (F_t - \theta) \end{aligned} \tag{24}$$

For estimation purposes, it is convenient to define all the dynamics using the demeaned factor $\tilde{F}_t$. Then, one can initiate the Kalman filter at its unconditional mean and variance:[28]

$$\hat{F}_0 = E\left(\tilde{F}_t\right) = 0, \qquad \hat{P}_0 = \text{var}\left(\tilde{F}_t\right) = \frac{\beta\theta}{2\kappa} \tag{25}$$

where $\hat{P}$ is the updated variance of estimated state variable, $\hat{F}$. All other Kalman filter implementation steps are identical to de Jong (2000).

[28] The empirical estimates suggest that the risk premium was trending over time, so this assumption is likely to underestimate the true decline in risk premia. One could treat the initial conditions as a separate model parameter and take into account this more rigorously.

2.D.2 Quasi-Maximum Likelihood

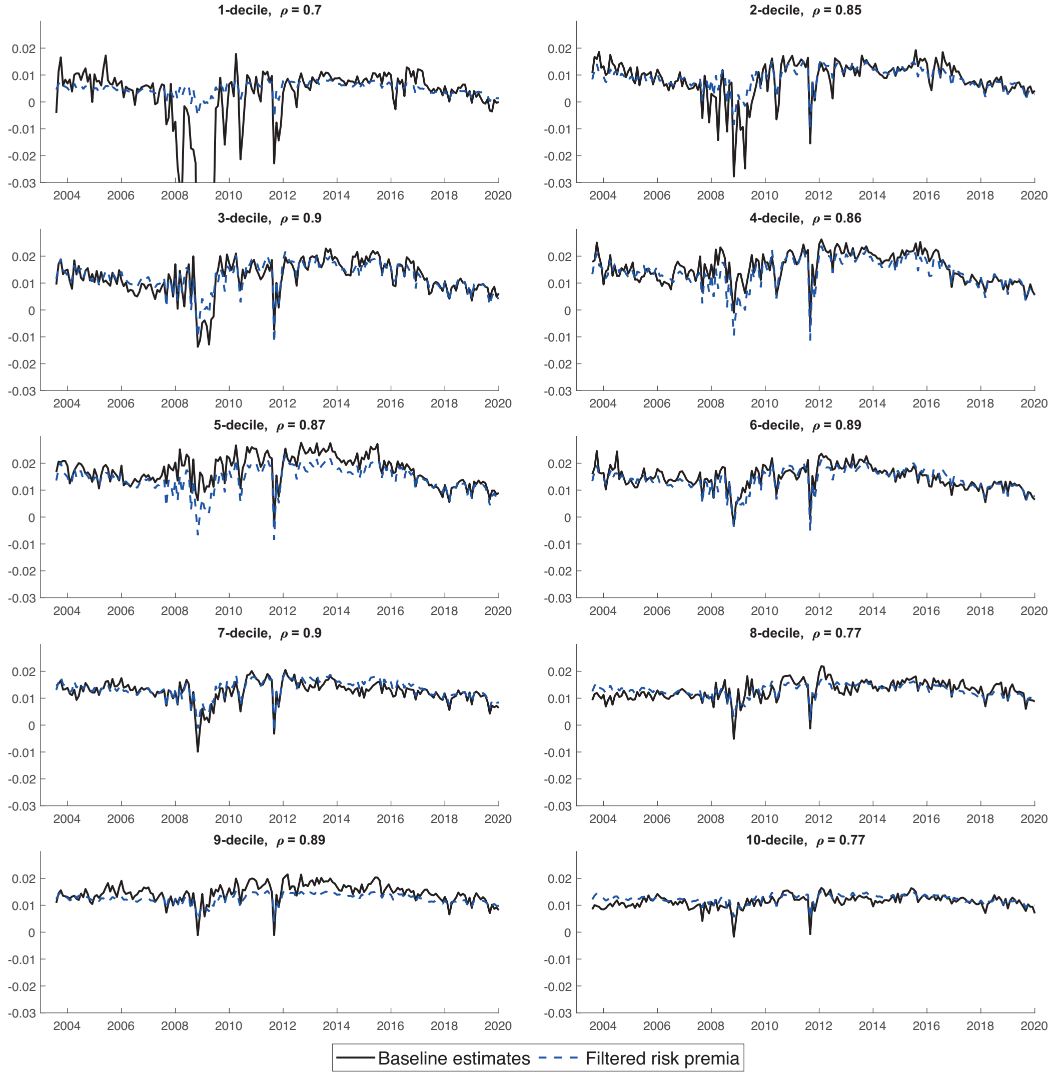
The model has a small number of key parameters $\Theta = [\kappa, \theta, \beta, \sigma_H^2, \psi]$ that are estimated using the quasi-maximum likelihood (QMLE). Namely, the log likelihood contribution at time t is:

$$-2 \ln L_t = \ln 2\pi + \ln |V_t| + u_t' V_t^{-1} u_t \tag{26}$$

where u_t is a fitted measurement error, and V_t is the Kalman filter prediction for the variance of risk premia. Note that this likelihood contribution is just an approximation of the underlying true likelihood since the state variable in (23) is not normally distributed over any discrete interval. In fact, the true distribution is non-centered chi-squared and the state variable F_t always stays positive when the Feller condition $2\kappa\theta > \beta$ is satisfied (Cox et al. (1985)). As de Jong (2000) shows, these approximation errors are quantitatively small, however, it is possible that during the estimation procedure the variance Q_t turns negative. To address, this concern I bound Q_t to be strictly positive. Lastly, since the state dynamics (20) and (22) are non-linear in Θ, there is a possibility of local optima. Therefore, I run the estimation for a large number of starting values.

The filtered series are presented below.

Figure A.IV: Filtered risk premia



Note: The figure displays the filtered dynamics of duration portfolio risk premia based on one-factor CIR model. The baseline risk premia estimates are derived using Feldhütter & Schaefer (2018) default probabilities (solid black line). The dash line is the filtered series based on Kalman filter and QMLE estimation. A statistic ρ summarizes the correlation coefficient between fitted series and duration portfolio risk premia. All risk premia are measured in units.

2.E Other Derivations

2.E.1 Approximation

We can deduce the relationship between the yields and the expected annualized return by calculating the future value (FV) of bonds' cash flows:

$$FV_0 = \sum_{t=1}^{m-1} E_0[c_t](1+y)^{m-t} + E_0[c_m + P_m] \tag{27}$$

where m is the maturity of the bond, y is the yield-to-maturity at time 0, c is a coupon payment, and P is the principal. As common in the literature, the assumption is that the bond holder is able to reinvest future coupons at the same rate as the initial investment.

In order to make use of this expression, one has to specify how defaults affect future cash flows. For instance, one can assume that upon the default the bond is restructured into the same maturity bond with lower principal and coupons payments (determined by the recovery rate). Then, each expectation can be replaced with $E_0[c_t] = (1-\pi_t)c_t + \pi_t L c_t = c(1 - L\pi_t)$ and (27) can be rewritten as:

$$FV_0 = \sum_{t=1}^{m-1} c(1 - L\pi_t)(1+y)^{m-t} + (c+P)(1 - L\pi_m) \tag{28}$$

Here π_t stands for the cumulative default probability at time t and L - loss given default which is set to be maturity invariant. We can easily rearrange this into:

$$\begin{aligned} FV_0 &= (1+y)^m \left[\sum_{t=1}^{m-1} \frac{c(1-L\pi_t)}{(1+y)^t} + \frac{(c+P)(1-L\pi_m)}{(1+y)^m} \right] = \\ &(1+y)^m \left[\underbrace{\sum_{t=1}^{m-1} \frac{c(1-L\pi_m)}{(1+y)^t} + \frac{(c+P)(1-L\pi_t)}{(1+y)^m}}_{P_0(1-L\pi_m)} + \underbrace{\sum_{t=1}^{m-1} \frac{cL(\pi_m - \pi_t)}{(1+y)^t}}_{\bar{FV}_0 > 0} \right] \end{aligned} \tag{29}$$

which means that the annualized expected return is:

$$\left(\frac{FV_0}{P_0} \right)^{1/m} = (1+y) \left[(1 - L\pi_m) + \frac{\bar{FV}_0}{P_0} \right]^{1/m} \tag{30}$$

As we see, the equation is almost identical to (2.1), except for the second term, $\frac{\bar{FV}_0}{P_0}$, which captures the effect of allowing defaults to occur yearly and correcting coupons with lower default probabilities. Even though the term is positive and increases with the default risk and maturity, it remains quantitatively small for most bonds. For example, based on the Moody's historical default frequencies, one can compute that the expected return for the American Airlines 3.5-year and 20-year maturity bonds on Jan 31, 2007 using both methods. The results are presented below.

Table A.XV: Different methods of extracting expected return from yields

Column I shows the baseline case using equation (2.1). Column II corrects for the early default assuming the bond is restructured into a smaller principal and coupon bond, whereas column III assumes that the bond is completely liquidated. Finally, column IV presents the scenario where all the intermediate cash flows are reinvested at the risk free rate of 4.9%.

Maturity	Rating	Yield	I	II	III	IV
3.5	BBB+	5.39%	5.13%	5.16%	5.18%	5.11%
20	BBB+	6.09%	5.73%	5.88%	5.88%	5.45%

Columns I and II present that the corrections leads to an increase of 3 basis points in expected return on a 3.5-year maturity bond, and 15bp on a 20-year bond. If one assumes that upon the default the firm is liquidated and $(1-L)$ is immediately reimbursed, (29) transforms into (31).[29] However, that has almost no further quantitative effect on the analysis. Overall, it seems that the baseline introduces a slight downward bias for the longest maturity bonds, however, those are the same bonds that are most sensitive to the reinvestment assumptions. For comparison, column IV illustrates the scenario where defaults occur early but all the intermediate cash flows are reinvested at the rik-free rate (on Jan 2007 the government yield term structure was quite flat at around 4.9%). This barely alters the short term expected returns but depresses the long-term bond returns by roughly 40bp. Since the Campello et al. (2008) approximation lies somewhere in the middle of all of these estimates, for simplicity, in the baseline I follow their approach.

2.E.2 Instantaneous risk premium

In this section, I will derive the instantaneous risk premium expression for the affine model in Section 2.6.1. Based on equations (2.9), (2.11) and the independence assumption, we can show that the m-maturity credit spread is a function of maturity-matched expected credit losses, $s^P_{t,m}$, and risk premia components:

$$s^Q_{t,m} = s^P_{t,m} + rp_{t,m} \tag{32}$$

Respectively, the bond yield is:

$$y_{t,m} = rf_{t,m} + s^P_{t,m} + rp_{t,m} \tag{33}$$

where $rf_{t,m}$ is the maturity-matched risk-free rate. In affine term structure models (such as CIR), the yields are linear in state variables. Thus, the price of a (zero-coupon)

[29]

$$FV_0 = (1+y)^m \left[\sum_{t=1}^{m-1} \frac{c(1-\pi_t)}{(1+y)^t} + \frac{(c+P)(1-\pi_m)}{(1+y)^m} + \sum_{t=1}^{m} \frac{\pi_{t-1,t}(c+P)(1-L)}{(1+y)^t} \right] \tag{31}$$

bond is:

$$P_t = exp(-my_{t,m}) = exp\left(-m\left[rf_{t,m} + s^P_{t,m}\right] - A(m) - B(m)F_t\right) \tag{34}$$

Note that $m = T - t$ where T is the maturity date. Using the Ito's lemma, one can show that the drift of bond return process under the physical probability measure is (conditional on no default):

$$rf_{t,m} + s^P_{t,m} + F_t - \psi B(m)\beta F_t \tag{35}$$

Following Yu (2002), this drift have to be adjusted for expected default losses, $s^P_{t,m}$, to arrive at the unconditional return drift:

$$\underbrace{rf_{t,m}}_{\text{risk-free rate}} + \underbrace{F_t - \psi B(m)\beta F_t}_{\text{risk premium}} \tag{36}$$

2.E.3 Holding-to-maturity returns

When defaults happen only at maturity, a zero-coupon bond has an expected return holding period return of:

$$\begin{aligned}\mathbb{E}_t(r^H_{j,t,t+m}) &= \pi_{j,m,t}(1 - L_m)\left(1 + y_{j,m,t}\right)^m + (1 - \pi_{j,m,t})(1 + y_{j,m,t})^m \\ &= \left(1 + y_{j,m,t}\right)^m\left(1 - L_m\pi_{j,m,t}\right)\end{aligned}$$

where $\mathbb{E}_t(r^H_{j,t,t+m})$ stands for the expected holding-period return of bond j of maturity m at time t, $y_{j,m,t}$ - the corresponding bond yield, L_m - loss given default, and $\pi_{j,m,t}$ denotes the default probability.

In the paper, I report the annualized version, $\mathbb{E}_t(r_{j,t,t+m}) = \mathbb{E}_t(r^H_{j,t,t+m})^{1/m}$.

Chapter 3

The Implied Equity Term Structure*

Co-authored with Lieven Baele and Joost Driessen

*We thank Stefano Cassella, Stefano Giglio, Benjamin Golez, Christian Julliard, Shane Miller, Michael Weber, Isabel Garrigues, and other seminar participants at the 5th Asset Pricing Conference by LTI@UniTorino, 2022 FMA Annual meeting, 29th Finance Forum at University of Santiago de Compostela, Tilburg University, and Yale University for useful comments.

3.1 Introduction

An active literature studies and measures the expected returns on equity dividend assets with different maturities. The properties of the term structure of equity risk premia have far reaching implications for at least three key reasons. To begin with, it is a direct test of competing theoretical asset pricing models as they mostly imply a flat or positive relationship between the maturity of an asset and its expected return. Second, the risk premium term structure is vital to understand the aggregate cost of capital and NPV calculations which may be driving companies' and investors' decisions at the micro level. Third, there is evidence suggesting that many documented anomalies (Cochrane (2011) 'factor zoo'), such as the value premium, may be captured or tamed by the differences in cash flow duration across firms (Dechow, Sloan & Soliman (2004), Weber (2018), Gormsen & Lazarus (2022)). In light of these arguments, it is crucial to find a precise measurement of the equity term structure.

In our paper, we propose a new methodology which tackles the inherent empirical challenges of measuring equity risk premia across different maturities. The idea of the procedure is straightforward: we project expected cash flows and estimate discount rates that bring the model price of the firm's equity as close as possible to the observed market value. In the implied cost of capital literature this is normally done by identifying one discount rate specific to the firm (e.g. Pástor, Sinha & Swaminathan (2008), Li, Ng & Swaminathan (2013)). In some sense, this is equivalent to assuming no commonality across observations which could further improve the precision of the estimation procedure. Hence, such a firm-level analysis limits the precision of estimates and, most importantly, leaves no degrees of freedom to recover maturity-specific rates. For these reasons, we diverge from the implied cost of capital literature by imposing some similarity across firms based on observable characteristics (e.g. size, book-to-market, credit risk) and use the cross-sectional variation in cash flows and prices to identify maturity specific risk premia. As a starting point, we assume that discount rates are the same across all firms, but vary with maturity. In this way we can focus on the pricing differences purely driven by the timing and size of cash flows. However, we subsequently allow these discount rates to be portfolio-specific and potentially flexibly modelled as any function of firm characteristics as long as the rates are not entirely firm-specific.

The key value of our method is that it delivers equity term structure estimates over the last 40 years which are neither based on (noisy) realized returns nor assume a specific asset pricing model. We achieve this by combining insights from two strands of the literature: the cross-sectional studies on the realized returns of portfolios with different cash flow durations (e.g. Dechow et al. (2004), Weber (2018)) and the implied cost of capital literature (Pástor et al. (2008), Li et al. (2013)). The key advantage of the first branch of literature is that the U.S. cross-section of stock prices and fundamentals span a much longer sample (1980-2021 in our case) as compared to the option contracts and dividend strip futures data used in the seminal work of van Binsbergen, Brandt & Koijen (2012) and van Binsbergen, Hueskes, Koijen & Vrugt (2013). The longer sample is particularly important when studying the dynamics of the equity term structure over the business cycle. Moreover, the liquidity of the dividend derivative markets may pose problems, especially in such short samples (Bansal, Miller, Song & Yaron (2019)). The sparse and only recent coverage of firm-level dividend strips also limits a thorough estimation of the equity term structure in the cross-section [1]. Yet, the realized equity returns used in cross-sectional studies are also subject to the critique of non-stationarity and large degrees of noise (Fama & French (2002)). Also, firm duration may be correlated with other drivers of expected returns, which makes it nontrivial to identify the duration effect. For these reasons, instead of averaging realized returns, we resort to the implied cost of capital literature and imply the expected returns using the constructed cash flow forecasts and observed market prices. In other words, we combine the best of two worlds without explicitly assuming a specific asset pricing model (as in Giglio, Kelly & Kozak (2022)). As shown later, our approach allows us to discover novel cross-sectional patterns.

We now explain our methodology in more detail. We augment the mean-reverting cash flow model of Dechow et al. (2004) with the IBES forecasts data on sales, return on equity (ROE) and earnings-per-share (EPS). In this manner, we are able to include a richer investors' information set and capture more flexible cross-maturity cash flow patterns than standard mean-reversion models alone[2]. Given these forecasts, we are able to estimate the short-, intermediate- and long-maturity discount rates by imposing a parsimonious Nelson-Siegel type functional form often applied in the fixed income literature (Nelson & Siegel (1987)). Specifically, by including just 3 parameters for the level, slope and decay, we are able to capture a wide array of monotonically increasing or decreasing term structures.

Our first main finding is that, when we apply our method to the full sample of stocks, the aggregate equity term structure's slope is positive. However, we also find substantial cross-sectional differences in risk premium term structures. For most portfolios sorted on size, book-to-market and credit risk the term structure remains upward sloping. But for small value firms and small firms with high credit risk, we document a downward sloping term structure. This novel result illuminates the role of cross-sectional characteristics not only for the level of expected returns (Fama & French (1993)), but also for the shape of the term structure. This is a powerful insight because the empirical success of asset pricing models may actually depend on the sample of firms one is looking at. On the one hand, the aggregate market risk premia are increasing with maturity as canonical habit (Campbell & Cochrane (1999)) and long run risk models (Bansal & Yaron (2004)) predict. On the other hand, to our knowledge, there is no equilibrium model which could match

[1] Gormsen & Lazarus (2022) are amongst the first to study the interesting market of single-stock dividend futures. Their sample starts, however, only in 2010 and covers only 180 different stocks.

[2] Cash flow mean-reversion models, applied in Dechow et al. (2004) and Weber (2018), are based on immediate and relatively fast mean-reversion that stifles most of the cross-maturity variation in firms' cash flows. As a consequence, the resulting cash flow duration proxies have small dispersion and estimated expected returns are mostly explained by the book-to-market ratios only. Since the empirical evidence does not support immediate reversion to the steady state across all firms, we incorporate IBES analyst forecasts for the first 2-5 years, in a similar vein as in Gebhardt, Lee & Swaminathan (2001) and Claus & Thomas (2001).

the cross-sectional patterns that we observe (as well as Giglio et al. (2022) document in their recent work). Thus, our findings reveal a new dimension to analyze and explain.

Our empirical analysis also shows that there is substantial time series variation in the equity term structure slope. In recessions, the term structure remains upward sloping but does become flatter, with short-term discount rates rising and long-term discount rates falling. This is consistent with the results of Bansal et al. (2019) and Giglio et al. (2022), who also detect countercylicality of short term risk premia, at least at the annual frequency[3]. On the other hand, our findings challenge Gormsen (2020) who, at the market level, finds a countercyclical term premium implying that most of the variation is driven by changes in long term rates. In contrast, our estimates reveal that most of the time variation comes from the short and intermediate maturity risk premia. Moreover, we do not find evidence that in recessions the variation is so large that it can create an inversion of the term structure slope at the aggregate market level.

As mentioned above, we also perform the estimation for different credit rating groups. We find a positive slope for all investment-grade firms, but a flat term structure for speculative grade firms. Moreover, we also find that the level of expected returns is decreasing monotonically with the creditworthiness of the issuer, a result which is intuitive, but often undetected in the distress risk literature using realized returns (Campbell, Hilscher & Szilagyi (2008)). When we perform double sorts on credit rating and size, we find that the smallest and least creditworthy firms actually exhibit a downward sloping risk premium term structure. This suggests that firms with high exposure towards short-term transient risk may be driving the puzzling negative relationship between the maturity of the cash flow and its expected return. We also apply our method to industry portfolios and find positive slopes for all industries, except for the steel works industry.

The key limitation of our approach is that it relies on projections of firm's future cash flows. We therefore perform a range of robustness checks on these cash flow projections and show that our qualitative findings persist in a wide array of robustness tests and alternative specifications. First, we implement van Binsbergen, Han & Lopez-Lira (2022)'s estimates of the analyst forecast biases to correct for documented optimism in IBES consensus estimates. Second, we consider time-varying long-term cash flow growth rates proxied by the Survey of Professional Forecasters' (SPF) predictions on long-term inflation and real GDP growth rates. Third, we average cash flow forecasts across different models, cap them, and exclude the most extreme size and book-to-market portfolios to alleviate the possible impact of outliers. None of these extensions affect the upward sloping pattern of the aggregate market equity risk premia. Perhaps not surprisingly, there are some indications that the decreasing term structure slope among the smallest value firms is dampened by additional filters and cash flow refinements. This means that some of the extreme patterns indeed come from estimation noise or small subsets of exceptional firms, but are not washed out entirely. On a similar note, we also perform the estimation with perturbed cash flow reversion parameters and different forecasting horizons which reveal

[3]The equity risk premia here is measured holding-to-maturity returns.

another interesting relationship: the slope of the implied term structure does depend on the assumed persistence of cash flow growth rates. For moderate to high levels of mean reversion, which includes our mean reversion estimates, the recovered term structure is upward sloping. In turn, with very slow mean reversion the slope can flip sign. Thus, one should be careful when drawing conclusions without accounting for the uncertainty that shrouds the cash flow persistence estimates. Interestingly, in most regressions our cash flow reversion estimates generate slower reversion compared to the existing literature (Dechow et al. (2004) and Weber (2018)). We provide evidence that cross-sectional characteristics, such as ROE and sales growth, converge slower in the medium to distant horizon than in the first few years.

The paper contributes to the empirical asset pricing literature aiming to measure the shape of the equity term structure. We find empirical support for the negative slope among the smallest value and high-risk firms, however, for larger firms the evidence conforms with Bansal et al. (2019)'s findings. This carries substantial ramifications for the theoretical research since it finds suggestive evidence supporting the classical theoretical asset pricing models of Campbell & Cochrane (1999) and Bansal & Yaron (2004), at least for large firms. On the other hand, our results challenge the current frameworks in the cross-sectional dimension as, to our knowledge, there is no microfounded model that can generate different equity term structures across firms sorted on book-to-market, size and credit rating. Moreover, this paper questions the existence of the documented credit risk anomaly (Campbell et al. (2008)).

The closest study to our paper is Giglio et al. (2022). They also estimate term structures in cross-sectional portfolios. They find that small (value) firms have a flat (downward sloping) risk premia pattern, whereas for the big (growth) firms the slope is upward sloping. Moreover, Giglio et al. (2022)'s findings conform with the idea that the unconditional equity term structure is upward sloping (mildly in their case), but it inverts in recessions. Even though we share similar insights, their paper is based on a fundamentally different approach. They assume a reduced form model for the pricing kernel (in the light of Lettau & Wachter (2007)) and fit its parameters to the principal components of realized returns. In contrast, we make assumptions about cash flows and imply ex-ante expected returns without assuming a specific asset pricing model. In this way, our methodology provides new degrees of freedom and a perspective that contributes to the ongoing debate in the equity term structure literature.

Finally, this paper is different from the studies that rely on the realized returns of duration portfolios (Weber (2018), Gormsen (2020)). Our approach does not suffer from the mechanical correlations of duration with the book-to-market ratio as we focus on the model-implied prices of firms. At the same time, our method is internally consistent - we use the same discount rate curve throughout the optimization.[4]

The structure of the paper is as follows. Section 3.2 describes the identification strategy and methodology. It also expands on the relative advantages and restrictions of the

[4] A common feature of duration portfolio studies is employing a flat (time-invariant) term structure to create duration measures eventually leading to the downward-sloping patterns in realized returns.

proposed framework over the existent studies. Section 3.3 presents the data, whereas section 3.4 elaborates on the estimation techniques applied. Section 3.5 outlines the results and key comparisons with the literature. Section 3.6 contains robustness checks, and section 3.7 contains a summary and final remarks.

3.2 Implying Discount Rates

As outlined, our method relies on three main components - the pricing equation, cash flow forecasts and the functional form for discount rates. In the following sections, we discuss the identifying assumptions in greater detail.

3.2.1 Pricing equation

At time t the value of firm i's equity can be described by this discounted cash flow equation:

$$\sum_{m=1}^{\infty} \frac{E_t[CF_{i,t+m}]}{(1+r_{i,t,m})^m} = P_{i,t} \tag{3.1}$$

where $CF_{i,t+m}$ is the free cash flow accruing to the company's shareholders[5], $P_{i,t}$ is the market value of equity, and $r_{i,t,m}$ is the maturity-specific (zero-coupon) discount rate. We assume that at some year T the firm reaches its steady-state where expected cash flows grow at the rate of g_t, and $r_{i,t,m}$ converges to $r_{i,t,T}$. In other words, both the term structure of cash flow growth rates and discount rates remain flat beyond this maturity. Given this, one can simplify this infinite sum to:[6]

$$\sum_{m=1}^{T} \frac{E_t[CF_{i,t+m}]}{(1+r_{i,t,m})^m} + \frac{E_t[CF_{i,t+T+1}]}{(1+r_{i,t,T})^T(r_{i,t,T}-g_t)} = P_{i,t} \tag{3.2}$$

In the base specification, we follow Dechow et al. (2004) and Weber (2018) and set g_t to 6%, which equals the long-term average nominal GDP growth rate. As shown later, allowing for time variation in g does not affect the results. Similarly, based on the literature, we fix the forecasting horizon T at 15 years.[7] For robustness, alternative horizons are also considered and the qualitative conclusions seem to hold as long as the forecasting horizon is distant enough for firms to reach their steady states. That is, with a more persistent cash growth rate process one may need to increase T to reach stable results.

The key identifying assumption is that the discount rate curve $\{r_{i,t,m}\}_{m=1:T}$ is constant across some set of firms. In the implied cost of capital literature, equation (3.2) delivers

[5] An equivalent formulation of (3.1) is the well-known Gordon growth dividend discount model where $CF_{i,t+m}$ is the dividend payment. Implicitly, we assume that the fundamental value of the listed firms is non-negative - a condition which, except for a small fraction of firms, is satisfied in the data.

[6] More formal derivations are included in Appendix E.

[7] The forecasting horizon of 15 years is used by Dechow et al. (2004), Li et al. (2013), and Weber (2018), amongst others.

a firm-specific discount rate $r_{i,t}$. However, this leaves no degrees of freedom to identify maturity-specific discount rates. To address this challenge, we group firms into characteristic portfolios and impose that all firms in the same group (i.e., portfolio I) share the same discount rates $\{r_{I,t,m}\}_{m=1:T}$. Effectively, the identification of discount rates relies on the cross-maturity variation of cash flows within the portfolio of firms.

3.2.2 Cash flow forecasts

The key ingredients in the estimation procedure are cash flow forecasts. Following Dechow et al. (2004) and others, we deduce the dynamics of CF's from the identity:[8]

$$CF_t = \underbrace{E_t}_{\text{earnings}} - \underbrace{(BE_t - BE_{t-1})}_{\text{reinvestments}} = BE_{t-1} \times (ROE_t - \%BE_t) \tag{3.3}$$

where BE stands for the book value of equity, $\%BE$ denotes its growth rate, and ROE is the return on equity. Since Nissim & Penman (2001) documented that the growth rate of sales is a better predictor for book equity growth than most other accounting variables, for the rest of the paper we proxy $\%BE$ by the growth rate of revenues.

In order to forecast ROE and sales growth, a common approach in the literature is to assume mean-reverting processes and use those projections to forecast cash flows (Dechow et al. (2004), Weber (2018), Chen & Li (2018)). It is a parsimonious but restrictive approach as all variation in cash flows is attributed to initial differences across firms which dissipate at the rate of reversion parameters. If the researcher underestimates the cash flow persistence, all firms quickly converge to the steady state and there are no meaningful differences between, for instance, highly profitable and loss making firms which eventually leads to the poor identification of short-term and long-term risk premia. As this scenario is empirically unrealistic (e.g. loss-making firms often take years to break even), one way to address this is to incorporate analyst forecasts to the first years of cash flow predictions (similarly to Claus & Thomas (2001), Gebhardt et al. (2001), Li et al. (2013)). This allows for more flexible patterns arising from different firm life cycles, and for more accurate depiction of investor beliefs. Needless to say, the analyst projections also alleviate the concern that our results are highly correlated with the initial book-to-market ratios (as in Dechow et al. (2004) and Weber (2018)) and provide additional cross-sectional variation in the estimation of long- versus short-term discount rates. Of course, analyst forecasts may carry some biases or additional noise too, and for this reason we later investigate specifications where cash flow forecasts are averaged across different approaches and find that the main results remain intact.

Based on the arguments mentioned above, we model cash flows in two stages. First, for the 1-5 year horizon, depending on the data availability, we extract information from the median analyst forecasts. Second, starting from the last year of analyst data availability until period T, we follow the same mean-reverting process as in Dechow et al. (2004) and

[8] To be precise, this identify is true when clean surplus accounting holds. To a large extent, this is empirically supported (Dechow et al. (2004)).

Weber (2018). The key modeling choices of these two stages are discussed below.

Stage one: Analyst forecasts

The estimation procedure benefits from the short and medium horizon forecasts available from IBES by introducing more cross-firm variation. However, the empirical challenge of using IBES forecasts is that they are predominantly available for the earnings-per-share (EPS) measure and there is very little information on the ROE side[9]. To circumvent these data deficiencies, we complement the available forecasts with implied ROE forecasts from median EPS forecasts, which almost triples our sample and data range. If we assume that the number of shares remains fixed, we can proxy the ROE forecasts by the median EPS forecasts divided by the forecasted BE values per share. In a similar spirit, we also imply the revenue growth projections from the level of revenue forecasts. The exact steps in which we construct our projections are described in Appendix A.

This procedure results in a wide heterogeneity across firms in their ROE and sales growth patterns (Table I). This is particularly valuable for ROE forecasts because empirically ROE mean reverts more slowly than sales growth meaning that most cross-sectional cash flow differences can be explained by this variable. One potential concern is that these implied forecasts have nothing to do with the true forecasts. However, as shown in Table II, the correlation between implied and actual ROE forecasts is actually high, reaching $70-80\%$ at shorter horizons, and it remains at roughly 50% for 5 year maturities. Moreover, when comparing the distribution of ROE forecasts with implied ones (Table I), one can notice similar means and medians with slightly larger standard deviations and tails of the distribution. Overall, it seems that the procedure delivers a good approximation, however, tighter filters and forecast averaging across alternative models may help to purge some noise and extremities from the estimation (see Section 3.6).

Stage two: Mean-reverting process

Starting from year 6 or the last available horizon of IBES forecasts, we assume similar mean-reverting processes for ROE and sales growth as in Dechow et al. (2004) and Weber (2018):

$$\begin{aligned} ROE_t &= 0.12 + 0.81(ROE_{t-1} - 0.12) \\ BE_t &= 0.06 + 0.41(BE_{t-1} - 0.06) \end{aligned} \tag{3.4}$$

Both AR(1) coefficients (0.41 for BE and 0.81 for ROE) are chosen based on the pooled OLS regressions using 3-year lags in the full Compustat sample over 1964-2019 (the third row of Table III). The equations state that ROE converges to the long-run cost of capital of 12%; $\%BE$ is proxied by sales growth, which also reverts to the assumed long-term

[9]The availability of ROE forecasts is limited to 2001-2014 period with relatively few instances in the early years. The level of revenue forecasts are accessible starting from 1992 with much larger coverage across firms. Yet, EPS forecasts date back to 1976 with the largest coverage of around 60,000 observations.

nominal growth rate of 6% (g). Both 12% and 6% are chosen based on the full sample historical averages, as in Dechow et al. (2004) and Weber (2018).

Table I: Analyst forecasts

The table below summarizes the implied sales growth, ROE and actual ROE forecasts in our merged Compustat-CRSP-I/B/E/S sample over 1980-2021. The implied sales growth forecasts are calculated as the growth rate in median I/B/E/S sales forecasts. The implied ROE forecasts are constructed by dividing median EPS forecasts by projected book value of equity (assuming that it grows at a sales growth rate). Horizon specifies the future financial year for which the forecasts are recorded. To alleviate outliers, all forecasts are winsorized at a 2.5 and 97.5% level.

Implied sales growth forecasts						
Horizon	p5	p50	p95	mean	sd	N
1	0.00	0.10	0.57	0.15	0.15	29577
2	-0.01	0.09	0.64	0.16	0.18	15211
3	0.00	0.10	1.07	0.22	0.28	6648
4	0.00	0.10	1.31	0.25	0.36	4991
ROE forecasts						
Horizon	p5	p50	p95	mean	sd	N
1	-0.04	0.15	0.41	0.16	0.12	12495
2	0.00	0.16	0.41	0.18	0.11	12313
3	0.03	0.17	0.53	0.20	0.13	7064
4	0.07	0.18	0.41	0.20	0.12	469
5	0.08	0.18	0.51	0.23	0.18	279
Implied ROE forecasts						
Horizon	p5	p50	p95	mean	sd	N
1	-0.24	0.11	0.54	0.13	0.17	50696
2	-0.09	0.13	0.63	0.17	0.16	47685
3	0.01	0.15	0.75	0.20	0.18	36051
4	0.02	0.17	0.86	0.23	0.21	31952
5	0.02	0.18	0.97	0.25	0.24	30940

Table II: Correlation between actual and implied forecasts

The table below summarizes the correlation between the actual ROE forecasts and implied ones for different forecasting horizons. The ROE forecasts are available only for 2001-2014.

Years	1	2	3	4	5
Corr	0.79	0.72	0.66	0.58	0.51
N	12479	11843	6798	454	269

The key difference here with previous research is that we employ different regressions that lead to much higher reversion coefficients (in other words, much slower mean reversion). Dechow et al. (2004) find a coefficient for ROE equal to 0.57, and 0.24 for sales growth; Weber (2018) obtains 0.41 for ROE and 0.24 for sales growth. Both studies estimate the annual AR(1) coefficients using annual pooled OLS regressions over the full sample starting in 1964. Yet, these estimates are sensitive to filters applied and, as illustrated in Figure C.I for the ROE and sales growth quintile portfolios, they are biased downwards due to the very quick initial mean-reversion and slow decay afterwards. In fact, similar pooled regressions where we use more distant lags of 3 to 5 years to estimate the AR(1) coefficient or tighter winsorization of the sample reveal that the true mean reversion for ROE (sales growth) is likely to be in the 0.74-0.84 (0.41-0.53) range (Table III) [10]. Since we aim to predict the long term behaviour of firm's cash flows (and short-term patterns should be mostly captured by the analyst forecasts), we conclude that slower mean reversion is more empirically plausible, especially for longer horizons.

Table III: AR coefficient estimates

The table presents the estimated AR coefficients from different pooled OLS regressions. The first five rows represent estimates using the full Compustat sample over 1964-2021, whereas the last 3 rows - merged Compustat-CRSP-I/B/E/S sample spanning 1980-2021. The second and third rows describe regressions where both ROE and sales growth are winsorized at 5% and 95% levels and ROE is capped at -100% (some firms incur losses exceeding their book value of equity). Rows four, five, seven and eight present specifications where the first order autoregressive coefficient is estimated using more distant lags. All AR coefficients are standardized to 1-year maturity. The standard errors (SE) are clustered by year. For longer lag regressions, the Delta theorem is applied to compute the AR(1) coefficient standard errors.

	ROE				Sales growth			
	AR(1)	SE	R2	N	AR(1)	SE	R2	N
full Compustat sample	0.009	0.009	0.003	190261	0.000	0.000	0.000	234721
winsorize (5% and 95%)	0.727	0.017	0.431	190261	0.226	0.015	0.055	234721
cap ROE at -100%	0.720	0.012	0.473	190261				
3 years lag	0.810	0.009	0.241	151445	0.413	0.017	0.006	193541
5 years lag	0.846	0.007	0.158	123611	0.536	0.025	0.002	161360
merged sample	0.693	0.020	0.475	37267	0.284	0.034	0.094	37934
3 years lag	0.778	0.017	0.217	28230	0.502	0.028	0.022	29019
5 years lag	0.825	0.012	0.144	22459	0.592	0.034	0.008	23164

In the robustness section we also consider alternative autoregressive coefficient magnitudes. While small perturbations do not affect the results much, larger shifts in mean reversion can indeed alter the qualitative conclusions.

3.2.3 Functional form for equity risk premia

Our framework is flexible in modeling different types of heterogeneity in firm discount rates. Nevertheless, since the m period cash flow is not expected to be much different from the $m+1$ period cash flow, the discount rates for similar maturities will be strongly correlated, and estimating them separately would lead to large standard errors and noisy estimates. For this reason, we choose the Nelson-Siegel (NS) functional form (Nelson &

[10] We also explored specifications with multiple lags and moving averages, however, the additional terms often were not statistically significant. For this reason, we keep the parsimonious AR(1) model as our base specification.

Siegel (1987)), often used to model fixed income yield curves, as it is both sufficiently rich in capturing different discount rate curve patterns and parsimonious in the number of parameters:[11]

$$r_{I,t,m} = r^f_{t,m} + \beta_{1,I} - \beta_{2,I}\,[\frac{1 - exp(-m/\lambda)}{m/\lambda}] \tag{3.5}$$

As we see in (3.5), we take into account the term structure of risk-free rates, $r^f_{t,m}$, by proxying it with the US zero coupon bond yield curve. This allows us to directly focus on the level of risk premia rather than total discount rates, and introduces aggregate time variation in our base scenario case. Importantly, for maturities $m > T$, both the risk premia and total discount rates are assumed to be constant, i.e. for all $m > T$ we have $r_{t,m} = r_{t,T}$.

Rather than reporting the parameter estimates, in the analysis below we mainly report two key aspects of the risk premium term structure: the *level* of the 15-year risk premium and the *slope*, defined as the difference between the 15-year and 5-year risk premium. As discussed below, given the typical duration levels of US firms, identifying the risk premiums at the very short end (1-4 years) is difficult, which is why we define the slope using the 15-year and 5-year maturity points.

The parameter λ captures the curvature of the equity term structure and is notorious for being difficult to estimate (e.g. you can always simultaneously increase the short rate and decay to achieve a similar overall effect). Thus, for the rest of analysis we will fix it to economically plausible values (i.e. we set $\lambda = 4$ as it introduces some curvature for the smooth convergence to the long-run levels). In Appendix C, we verify that our results remain similar when we use other values for λ (Figure C.III).

3.3 Data

The analysis is performed using US data over the 1980-2021 period. Even though all required variables are observable as of 1976, we start in 1980 since from that year onwards we have a cross-section of more than 500 firms per year. This is important since our approach requires substantial cross-sectional variation, especially since we also perform analyses where we split the cross-section in several portfolios. The balance sheet information comes at the yearly frequency. Stock price data and balance sheet data, as well as historical credit ratings and S&P500 index composition, are obtained from Compustat-CRSP. Supplementary book equity data is collected from the Kenneth French website. The median analyst forecasts are from Thomson Reuters I/B/E/S summary database, whereas historical constant maturity (zero-coupon) yields are from Gürkaynak, Sack & Wright (2007). Historical GDP and NBER recession periods can be accessed from St. Louis FED, and the long-term inflation and real GDP growth expectations come from the

[11] In this version, we do not include the hump-term as it is known to be highly correlated with other parameters. Thus, the term structure of risk premiums is either monotonically increasing or decreasing - an assumption which can be relaxed in future work.

survey of professional forecasters collected by FED of Philadelphia.

We construct our book equity measure following Davis, Fama & French (2000).[12] We measure book equity (and other historical accounting variables) using the most recent financial year results prior to the valuation month (June). Similarly, the market cap is calculated as the price times the number of shares outstanding at the end of the valuation month (June). Both prices and shares outstanding come from CRSP. June is chosen as the valuation month because it guarantees that most of the financial statement information has already arrived to the market, and analyst forecasts are mostly covering the upcoming financial years. To reduce potential extremities, we keep only those I/B/E/S median forecasts that were generated by more than 1 analyst. Adding to that, in order to alleviate liquidity concerns, we exclude the smallest 20% market cap firms on a year by year basis. Following Weber (2018), utilities ($4900 <=$ SIC<5000) and financials ($6000<=$SIC<7000) are dropped from the sample, as well as firms that are not listed on one of the major stock exchanges (NYSE, AMEX and Nasdaq). Finally, since we value the total equity size of each firm, we exclude all mismatches larger than 5% between Compustat and CRSP database market caps[13]. These filters leave us with 79,537 Compustat-CRSP firm-year observations over the 1950-2021 period. After merging with I/B/E/S, the final sample consists of 50,712 firm-year observations over the period 1980-2021. More details on the filters and sample construction are provided in the Appendix.

In terms of summary statistics, the final sample is representative of the broader Compustat-CRSP dataset. As reflected in Table IV, the estimation sample consists of larger firms (with median market cap of USD 722 mln. as compared to USD 260 mln. in Compustat-CRSP) as larger firms tend to be more frequently covered by analysts. Nevertheless, financial ratios, such as ROE or dividend payouts, have almost identical distributions in both samples. The only exception is the BM ratio which is slightly lower in the final sample (median of 0.43 vs 0.49) because larger firms tend to have higher valuation ratios. Since most financial ratios have large outliers due to small denominators, we winsorize ROE (including sales growth) at the 5% and 95% level on a yearly basis.

[12] Observations with negative book equity are excluded.

[13] Differences in end of the year market caps between Compustat and CRSP records may lead to a severe mispricing in our model. For this reason, we exclude all observations with large discrepancies that potentially indicate multiple classes of traded or non-traded shares.

Table IV: Summary statistics

The table compares the Compustat-CRSP-I/B/E/S sample used for the analysis with a broader Compustat sample obtained after standard filtering procedures (as in Weber (2018)). Sales, book and market equity are presented in millions of USD. ROE, sales growth, book-to-market, earnings-per-share and dividend payout ratio are all presented in units. The dividend payout ratio is calculated similarly as in Li et al. (2013) - the total dividend paid divided by net income (in case it is positive). The ROE is the defined as current earnings before extraordinary items divided by lagged book equity. To alleviate outliers, ROE and sales growth are winsorized at a 2.5 and 97.5% level.

IBES-Compustat 1980-2021						
	p5	median	p95	mean	std	N
Sales	22	583	13846	3597	13800	50712
Book equity	27	314	7102	1894	8054	50712
Market equity	71	773	20633	5578	29186	50712
Sales growth	-0.19	0.10	0.78	0.17	0.29	49036
ROE	-0.44	0.11	0.40	0.08	0.23	48161
B/M	0.08	0.42	1.37	0.55	0.61	50712
E/P	-0.19	0.04	0.13	0.00	0.31	50712
Payout	0.00	0.00	0.86	0.16	0.27	50712
Compustat 1980-2021						
	p5	median	p95	mean	std	N
Sales	6	254	9264	2401	11167	79537
Book equity	8	139	4511	1268	6515	79537
Market equity	19	317	12737	3665	23464	79537
Sales growth	-0.21	0.10	0.78	0.16	0.30	73370
ROE	-0.47	0.10	0.39	0.06	0.24	72207
B/M	0.08	0.46	1.59	0.61	0.77	79537
E/P	-0.24	0.04	0.15	0.00	0.40	79537
Payout	0.00	0.00	0.85	0.15	0.26	79537

3.4 Estimation

We estimate model parameters β_1 and β_2 by minimizing the squared relative pricing error between the firm's total market capitalization and model valuation. Specifically, the squared pricing error for firm i at time t can be expressed as:

$$\phi_{i,t} = \Big(\sum_{m=1}^{T} \frac{E_t[CF_{i,t,t+m}]/P_{i,t}}{(1+r_{t,m})^m} + \frac{E_t[CF_{i,t,t+T+1}]/P_{i,t}}{(1+r_{t,T})^T (r_{t,T} - g_t)} - 1 \Big)^2 \tag{3.6}$$

where $r_{t,m}$ is a function of β_1 and β_2 as in equation (3.5).

In the baseline scenario, we estimate β_1 and β_2 by minimizing the sum of the squared pricing errors over the whole sample:

$$\underset{\beta_1,\beta_2}{argmin} \sum_t \sum_i w_{i,t} \phi_{i,t} \tag{3.7}$$

where $w_{i,t}$ is the weight assigned to firm i in year t.

The main criterion is thus how well the chosen risk premia parameters fit the cross-section of stock prices. This estimation procedure can be seen as a joint calibration of model values to market prices where observations with the largest weight or highest pricing deviations attain more importance. We choose to use a value-weighting scheme for the estimation, to focus the estimates towards economically sizable firms and to prevent that large pricing errors for small firms dominate the estimation procedure. We do impose that each observation year has the same weight in the estimation procedure, to prevent that later years with higher market values dominate the results. Later we provide results based on size sorts to analyze differences between small and large firms.

Clearly, differences in cash flow term structures across firms are a necessary condition for identification. In Figure C.1 of Appendix C we graph the distribution of firm's equity durations, where these durations are calculated using our benchmark risk premium estimates. The graph shows that there is indeed substantial cross-firm variation in durations, with a substantial amount of firms having durations below 10 years or above 20 years, respectively.

Finally, to alleviate concerns that there are firms whose valuation our model cannot capture well, we winsorize the sample at 2.5% and 97.5% based on the model implied internal rate of return of a stock. This step effectively excludes firms with extremely high or low valuation ratios (for instance, companies in severe financial distress) that are challenging to value with standard fundamental value models. In this way, without too much loss of generality (most stocks excluded are smaller ones with extreme ROE and sales growth rates), we reduce the impact of outliers.

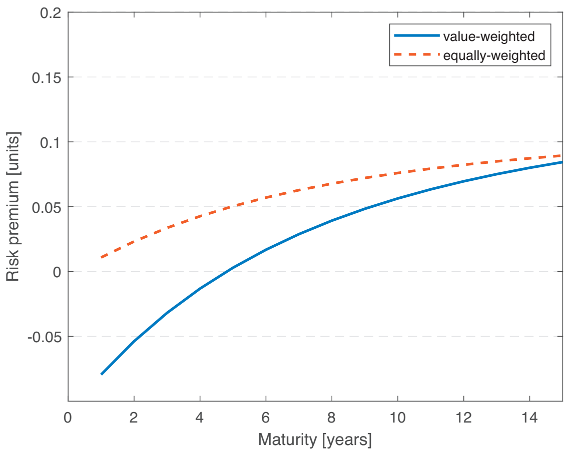
3.5 Results

3.5.1 Market-wide estimates

We start with the panel of firms over the full sample period, 1980-2021. We allow for a time-varying risk-free rate term structure and assume the same equity risk premium term structure for all firm-year observations. Later, we extend the setup by allowing for cross-sectional and time-series variation.

The baseline results unveil a clear pattern - the aggregate stock market has an upward sloping term structure of risk premia. This can be seen from Figure I. The long-term risk premia flatten out and approach a constant level of roughly 7-8%. Given that most value of the market is concentrated in the intermediate and long-term horizon, this should represent the unconditional equity risk premium to some extent. Figure I also shows that risk premium estimates are below zero for maturities of 4 years and shorter. As discussed above, given that equity durations are typically above 5 years, identifying the very short end of the risk premium term structure is difficult and the negative premiums at the short end effectively follow from the extrapolation that is inherent in the Nelson-Siegel specification. We therefore mainly focus on the difference between the 5-year and 15-year maturity points in our discussion. Figure I shows that our benchmark estimates thus generate a slope difference of about 8%, which is substantial. The qualitative finding of a positive slope is confirmed in many specifications below, meaning that, everything else equal, investors have a preference for shorter maturity cash flows. We also present results for an estimation where we equally weight all squared percentage pricing errors of firms. In this case, smaller firms likely dominate the estimation results. Figure I shows that we still obtain an upward sloping term structure, but with a smaller slope and higher short-term premiums. The size sorts presented below confirm the size effect.

Figure I: Full-sample term structure of equity risk premia



Note: The graph depicts the estimated equity risk premium term structure over the full sample for the benchmark value-weighted case, and for a case where we equally weight all firms in the estimation. Compustat-CRSP-I/B/E/S sample: 1980-2021.

Our finding that the equity term structure is upward sloping is consistent with the predictions of well-known theoretical asset pricing models, such as habit (Campbell & Cochrane (1999)) and long-run risk (Bansal & Yaron (2004)). Following their logic, the long-term cash flows are more sensitive to discount rate shocks and contain more uncertainty, therefore, the short maturity risk premium should be lower.

At first sight, our findings seem to go against those of van Binsbergen et al. (2012) and subsequent studies using dividend asset returns (van Binsbergen et al. (2013) and Gormsen (2020)). van Binsbergen et al. (2012) are the first to document the unusually high realized returns on short-maturity dividend strips relative to the total S&P 500. To compare our predictions as closely as possible, we build a sample by tracking down the firms which were constituents of the S&P 500 index in the valuation month. Adding to that, we limit our sample to 1996-2008, as in van Binsbergen et al. (2012).[14] However, we

[14]The match between samples is not perfect since, on average, we have only 250-300 S&P500 firms in our sample. The main reason is that we exclude financials and utilities that constitute roughly 150 firms of the index.

discover a very different pattern: consistent with our full-sample results the implied term structure is upward sloping (Appendix Figure C.IV). This seems to contradict the findings of van Binsbergen et al. (2012). However, in the latter study, the short-maturity rates are measured using options of maturities up to 2 years. Yet, this is exactly the range where our estimates are the least precise and mostly extrapolated from the intermediate maturities. Thus, it is possible to have a U-shaped term structure which could accommodate both features - the downward sloping pattern for very short maturities and an upward-sloping curve starting from intermediate years. Based on the information embedded in the cross-section of stock prices and fundamentals alone, we cannot reliably test this hypothesis.

Other studies use realized stock portfolio returns to identify the term structure. Weber (2018) sorts stocks on their duration and finds a negative term structure of returns. Giglio et al. (2022) use an asset pricing model to describe the cross-section and dynamics of stock returns and find that, on average, the term structure of risk premiums is close to flat. The difference with our ex-ante return patterns could for example be due to trends or structural changes in the term structure of risk premiums, in which case realized returns will differ from ex-ante returns.

3.5.2 Size and book-to-market sorts

We then proceed by estimating the term structure for quintile portfolios, sorted on size or book-to-market (Table V). For the size sort, we find a positive slope for all quintiles, with the largest firms having the steepest term structure. Remarkably, the level estimates, which represent the 15-year risk premium, are all around the 8-9% mark without any clear correlation with the quintile of the size distribution. It thus appears that small firms have much higher short-term risk premia than large firms, whereas the long-term rates are almost indistinguishable. It is important to note that the size of a firm is not a parameter in our model affecting the forecasted cash flow growth rates which means that observed patterns are much more likely to capture investor preferences rather than our modeling choices alone. Our results are thus consistent with the extensive literature that documents a size effect in realized returns.[15] Based on our estimates, all of this premium comes from the short- and intermediate-maturity cash flows (our baseline estimates suggest that roughly 50% of firm value comes from the forecasting horizon cash flows). Quantifying the realized return premium in our sample requires careful modeling of time-series variation in risk premia and the proper weighing of cash flows.

[15] Giglio et al. (2022) report that the size premium over the last 50 years was 0.8%.

Table V: Size and book-to-market portfolios

The table shows the estimation results within size and book-to-market quintiles formed on a yearly basis. The first row always represents the level (i.e. 15-year risk premium) and the second row - slope estimates (15 minus 5-year risk premium spread). Size is defined as price times the number of shares outstanding at the end of valuation month (June). Compustat-CRSP-I/B/E/S sample: 1980-2021. Risk premia are measured in annual percentages.

Quintiles		Size	BM
Low	level	0.087	0.088
	slope	0.052	0.150
2	level	0.085	0.088
	slope	0.109	0.130
3	level	0.081	0.091
	slope	0.118	0.108
4	level	0.084	0.086
	slope	0.145	0.083
High	level	0.088	0.086
	slope	0.117	0.036

Our results are also consistent with the findings of Giglio et al. (2022), who report an upward sloping term structure of (forward) risk premia for large firms and a much flatter term structure for small firms. Similar to our findings, large firms have the lowest short-term discount rates according to their estimates.

It thus seems that investors have a preference for short-maturity cash flows generated by large firms. This could be consistent with the intuition that the majority of small cap stocks are young firms that are exposed to short-term fluctuations priced by investors. Yet, if the firm survives long enough, it matures and evolves similarly as large firms whose dynamics are largely explained by the long-run risks. Even though we do not explicitly model this transition process and the associated risks, additional tests with business cycles and credit risk support this story.

We then turn to the sorts on the book-to-market ratio. The recent literature finds contradicting evidence regarding growth and value stocks. Even though the value anomaly is empirically sizable (2% based on Giglio et al. (2022)), there is still much debate on whether the value premium is a compensation for risk, and if so, what the nature of these risks may be. Recent studies find that the value anomaly may be captured by the duration of firms' cash flows which combined with a downward-sloping term structure would give rise to the value premium (Lettau & Wachter (2007), Weber (2018), Gormsen & Lazarus (2022)). In order to test this aspect, we sort stocks every year in 5 quintiles and re-estimate the model for each of these portfolios. The key finding is that for all portfolios we find an upward sloping term structure (Table V). One can notice that long-maturity rates are similar at 8% to 9% per year. Yet, similarly as with size, there is a strong monotonic relationship between the slope and BM quintiles. High book-to-market (value) stocks have the smallest slope estimate and thus the highest short-term risk premiums, suggesting that these firms are more exposed to (systematic) short-term risk than low book-to-market (growth) stocks. These results are also in line with Giglio et al. (2022). Similarly to size, the value premium comes from the short- and intermediate-maturity cash flows. The level estimates suggest only a 0.2% difference in long-term risk premia. Such a slight difference can also be justified by firms' tendency to revert in their characteristics over time. After 15 years, most growth firms are likely to be very similar to the value firms.

Our findings suggest that the value premium arises from the pricing of short maturity cash flows. The sheer fact that the value premium is detected is not surprising given that there is far more variation in cross-sectional BM ratios relative to observable economic variables affecting the fundamental value of a firm. This is often the key reason why the implied cost of capital estimations generate a value premium. What we add to this is that this value effect is mainly driven by pricing of short term cash flows.

In order to isolate the pure size and BM effects, we analyze double sorts where first observations are assigned to three size and then to three BM portfolios. The obtained results are presented in Table VI. In this case, we find substantial positive slopes for all portfolios except for the smallest value stocks, which exhibit a substantial negatively sloped term structure. The results potentially can be linked to transient short-term risks that distressed firms face if one assumes that the smallest and most extreme value firms are those close to default. Overall, we see that both larger firms and firms in the value domain have more positively sloped term structures, where the effect is largest in the book-to-market direction. Looking at the level estimates (the 15-year risk premium), we again see that these are fairly stable across portfolios.

Table VI: BM-Size double sorts

The table shows the estimation results within 3 book-to-market (BM) and size portfolios formed over the full sample. The data is first sorted on the BM ratios and then on the size. The first row always represents the level (i.e., 15-year risk premium) and the second row - slope estimates (15 minus 5-year risk premium spread). Compustat-CRSP-I/B/E/S sample: 1980-2021. Risk premia are measured in units.

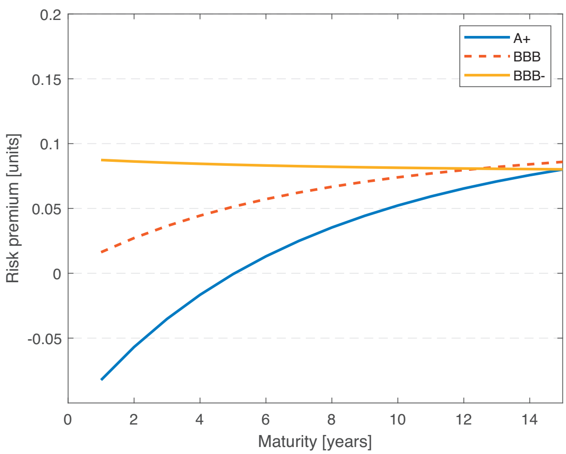
		Size terciles		
BM terciles		Low	2	High
Low	level	0.077	0.076	0.091
	slope	0.121	0.161	0.158
2	level	0.075	0.075	0.092
	slope	0.093	0.115	0.121
High	level	0.091	0.088	0.090
	slope	-0.071	0.037	0.052

3.5.3 Credit risk

Many articles have analyzed how realized equity returns differ across firms with different credit risk. Our approach allows to analyze the ex-ante returns for firms with different credit risk, and to see how the term structure of equity is related to credit risk. Thus, building on this existing work, we proxy the credit risk by the S&P issuer's long-term credit rating. Due to data availability, the sample shrinks substantially. We therefore aggregate all issuers into 3 broader categories: above BBB+ (3,865 observations), between BBB+ and BBB-(4,166), below BBB-(6,790). The first two are investment grade ratings, whereas the last one is classified as speculative.

The results in Figure II show that the term structure strongly depends on the credit rating. Firms with higher credit rating have, on average, lower short- and intermediate-maturity discount rates and a positively sloped term structure, while speculative grade firms have essentially a flat term structure of risk premia. This again corresponds to the intuition that such firms are more exposed to short-term risks. The finding that risk premia increase as credit risk is higher is not trivial as many studies cannot detect this pattern using realized returns (Campbell et al. (2008)).

Figure II: Credit ratings



Note: The figures present the estimated equity risk premia term structure in different S&P credit rating portfolios. A+ group contains all credit ratings above and including an $A-$ rating. BBB group includes all $BBB+$, BBB and $BBB-$ rated firms. Finally, BBB- includes all non-investment grade ratings. Compustat-CRSP-I/B/E/S subsample: 1985-2018.

We then perform double sorts, first sorting on the credit rating and then on size (Figure C.V). We see that small stocks with a speculative credit rating have a downward sloping term structure. Given that it is only the small stocks exhibiting this pattern, this supports the company's life cycle story where small, low-rated companies have high short-term risk, and either default or survive and become larger and less risky.

3.5.4 Time-variation

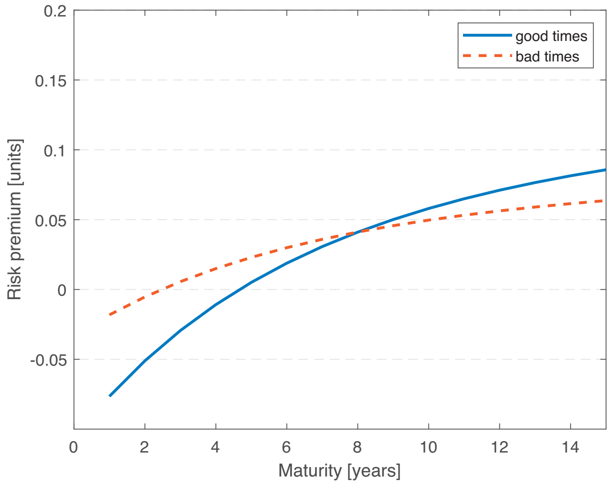
Recent empirical research has analyzed the variation of the equity term structure using realized returns. Gormsen (2020) finds that the slope of the curve is countercyclical - in bad times, the long maturity premium is higher than short maturity, and vice versa. On the other side, Bansal et al. (2019) base their analysis on holding-to-maturity returns and argue that the slope is procyclical. One potential cause for these contradicting findings is that the realized returns are known to be very noisy and one may need far longer samples than the data allows to detect changes[16]. For this reason, we shed our perspective based on ex-ante expected returns.

To inspect the variation of equity risk premia over time and across business cycles, we use the NBER recession dummy which equals 1 if at least one month in the valuation year is classified as the NBER recession. Note that the ex post nature of the NBER recessions does not affect the qualitative implications of our results because, if there is any substantial change in the shape of the term structure, this should have occurred in the most severe episodes.

Figure III depicts the risk premia estimates in both NBER recessions and expansions. Periods with high economic prosperity are concurrent with very small or even negative short maturity premia estimates and relatively high long maturity rates. In contrast, the recessions are associated by rising short term rates and slightly decreasing long maturity premia. Still, in both recessions and expansions the term structure remains upward sloping. Our results are thus closest to Bansal et al. (2019), since we both find the most positive slope in expansions, though we do not find evidence for a negative slope in recessions as in Bansal et al. (2019).

[16]In fact, both realized one period and holding-to-maturity returns in short samples may struggle to distinguish time variation in the level of the equity risk premium from time variation in the slope. However, this concern in our methodology is less of an issue because we identify short- and long- maturity returns based on the relative changes in the pricing of short and long cash flow duration firms, including the fact that our Nelson-Siegel specification allows for time-varying levels of risk premiums. Adding to that, we look at relatively long samples starting from 1985 that should, to a substantial degree, attenuate measurement errors.

Figure III: Good and bad times



Note: The graphs depict the estimated equity risk premia term structure in NBER recessions (dashed red line) and normal times (solid blue line). The year is classified as an NBER recession if at least 1 month in that year is considered to be a recession by the NBER. Compustat-CRSP-I/B/E/S sample: 1980-2021.

3.5.5 Market risk exposures

We investigate if exposures to the market are somewhat related to the upward-sloping term structure that we document. For instance, if the analysts form their expectations using the CAPM, our measure of cash flow duration is likely correlated with the market beta. This argument relies on the assumption that analyst both correctly price and forecast firm's cash flows. Alternatively, different horizon cash flows can merely reflect increases in the quantity of systematic risk, even though the price of risk remains constant. To test that prediction, we construct a model-based cash flow duration, $Dur_{i,t}$, based on the projected cash flows, our baseline discount rate term structure estimates, and the model price of a firm ($\tilde{P}_{i,t}$):

$$\begin{aligned} Dur_{i,t} &= \sum_{m=1}^{\infty} m \frac{E_t[CF_{i,t+m}]}{(1+r_{I,t,m})^m} \frac{1}{\tilde{P}_{i,t}} \\ &= \sum_{m=1}^{T} m \frac{E_t[CF_{i,t+m}]}{(1+r_{I,t,m})^m} \frac{1}{\tilde{P}_{i,t}} + \left(1 - \sum_{m=1}^{T} \frac{E_t[CF_{i,t+m}]}{(1+r_{I,t,m})^m} \frac{1}{\tilde{P}_{i,t}}\right)\left(T + \frac{1+r_{I,t,T}}{r_{I,t,T}-g}\right) \end{aligned} \tag{3.8}$$

This model of cash flow duration parallels the one used in Weber (2018) and Chen & Li (2018), with the exception that we use the model price of a firm and the estimated term structure for discounting. Then, we sort firms into duration quintiles each June and form value-weighted portfolios.[17] Given their returns, we are able to calculate portfolio

[17]The average firm in our estimation sample displays a cash flow duration of roughly 15 years. The

β^{CAPM}. The results are displayed in Table VII.

Table VII: Betas of duration portfolios

The table presents the (monthly) realized return betas for duration quintiles. To sort firms into portfolios, we use the value weighted term structure estimates from our baseline to compute the cash flow duration for each firm. Each June the portfolios are rebalanced based on the most recent duration estimate. The standard errors (in the brackets) are computed using the Newey-West procedure with a 6-month lag. Compustat-CRSP-I/B/E/S sample: 1980-2021.

	Low	2	3	4	High
α	0.006***	0.006***	0.005***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β	0.957***	0.997***	1.038***	0.948***	0.979***
	(0.023)	(0.033)	(0.025)	(0.035)	(0.038)

We can notice that there is no clear pattern in market betas across the duration quintiles. β^{CAPM} increases from 0.96 to 1.04, moving from the lowest to middle duration portfolio. Afterwards, the pattern disappears.[18] Thus, such differences in market betas cannot explain our term structure results since even a 0.1 increase in market beta would realistically correspond to a 0.4-0.6% slope in the term structure. In future research, we will examine if other documented anomalies display a stronger link to these duration portfolio returns.

3.6 Robustness

3.6.1 Biases in cash flow expectations

The literature documents that the analyst forecasts entail biases which can, in turn, affect our estimates of risk premia. Recent empirical work highlights that the biases tend to increase with maturity (van Binsbergen et al. (2022), Cassella, Golez, Gulen & Kelly (2023)), at least for the first few years. Moreover, there is some evidence that analysts overreact/underreact to information contributing to known cross-sectional anomalies (Bordalo, Gennaioli, Porta & Shleifer (2019)). All these errors can distort our estimates if the investors have different expectations from the analyst (or introduce a wedge between subjective and objective risk premia). In order to take this into account, we implement van Binsbergen et al. (2022)'s estimates of analyst biases in our cash flow forecasts. Specifically, we correct all ROE forecasts for the biases in EPS forecasts.[19] However, since van Binsbergen et al. (2022) focus only on the first 2 years of forecasts, we make a simplifying

median duration in each duration quintile is 7.5, 10, 12, 15.5, and 28.5 years, respectively. For an average firm, the present value of the first 15 years of cash flows makes up roughly 55% of the total model firm value.

[18]Most market betas fall below 1 as our sample is constituted of predominantly larger firms (the average β^{CAMP} is 0.96).

[19]In that sense, we investigate the most extreme scenario of analyst biases. If there was the same-direction bias in sales growth, the bias would cancel out. We can see that from equation (3.3).

assumption that for years 2-5 the bias stays constant. Afterward, it dwindles at the speed of cash flow reversion.

The summary statistics of analyst biases are presented in Table C.I, where each EPS bias is normalized by the firm's stock price. On average, the bias is positive and increasing with maturity. One-year cash flows display a bias of roughly 1%, whereas the two-year bias stands at close to 2.5%. However, it appears that the bias is particularly extreme in recessive economic times when the errors jump to 2.3% for one-year and to 5% for two-year forecasts. Thus, van Binsbergen et al. (2022)'s estimates suggest that the analyst are overly optimistic about the severity or duration of recession periods.

The results with corrected cash flow forecasts are displayed in Figure C.VI. The correction has not altered the quantitative conclusions much - the slope remains positive both for value and equal weighting schemes. On the other hand, quantitatively, the short-maturity risk premia become even smaller. It is an expected result, given that a positive bias would mean that cash flows and discount rates should decrease to deliver similar valuations. Even though the bias is increasing for the first few years, in our specification, the bias eventually dissipates. In order to reverse the slope, the bias has to keep steeply increasing with maturity, even at very long maturities. To our knowledge, nobody has provided such empirical estimates of long-term biases, and minor cash flow distortions do not seem to affect our qualitative conclusions.

3.6.2 Time-varying g

The cash flow modeling hinges on the assumption that with time cash flows converge to some steady state g. As outlined in the general version of the model, g_t can be made time-varying. This would take into account changes in the growth expectations and inflation - two key macroeconomic variables that may capture phenomena such as structural shifts or secular stagnation. One possibility is to follow Li et al. (2013) who proxy g_t with the average sample nominal GDP rate up to the valuation year. As shown in Appendix C, the benchmark result of an upward sloping aggregate term structure is very robust to using this time-varying g. We also exploit the market expectations about the long-term inflation and real GDP trends. Specifically, we incorporate the FED survey of professional forecasters data which dates back to 1982 for long-term inflation and to 1992 for real GDP. Mixing inflation forecasts with the average realized real GDP growth rate or real GDP inflation forecasts does not affect the full sample estimates: we again find an upward sloping term structure. Overall, it appears that differences in the steady state growth rate do not generate shifts in the term structure as much as changes in the cross-section of stock prices.

3.6.3 Cash flow forecast averaging

Considering that our findings may be driven by particular filters or noisy cash flow forecasts, we investigate alternative procedures. The generic insight is that, regardless of additional refinements, we always find upward sloping aggregate term structure estimates with in most cases an even more positive slope compared to the benchmark results (Table C.II).

First, we average cash flow forecasts between our model and the simple mean-reversion approach, as in Dechow et al. (2004) and Weber (2018), which should reduce the impact of outliers. The findings again show a strong positive slope at the aggregate market level. We also introduce a more intricate weighting of these two different models where each forecast i is assigned a score based on the likelihood of such cash flow to materialize:

$$w_i = \frac{\phi(forecast_i; \mu, \sigma)}{\sum_{k=1}^{2} \phi(forecast_k; \mu, \sigma)} \tag{3.9}$$

In order to determine the likelihood, we assume that cash flows are normally distributed with the mean of 0.75% and standard deviation of 19.7% (empirical sample statistics). For instance, if one of the forecasts is very far from the prior mean of 0.75%, then all the weight is attributed to the more conservative forecast. On top of that, we cap the maximum (minimum) cash flow at 100% (-100%) of firm's value. All these steps should smooth out the extremities. As we see from the second row, all this averaging does not affect much our baseline results. The same holds if we winsorize the sample at 5% and 95% percentiles of model internal rate of return (the third, fourth, fifth rows). Finally, we exclude the most extreme size and book-to-market portfolios, and again find a positively sloped term structure.

3.6.4 Persistence of cash flows

Concerns about the methodology may also point at the cash-flow reversion process itself. Given the estimates we employ, most of the firms reach their steady state growth rate in 6-10 years. In order to test if the reversion is driving the qualitative results, we rerun our base estimation with perturbed ROE and sales growth reversion parameters (Table VIII). The general tendency is that for small and moderately high mean reversion coefficients the term structure slope is positive. However, very large coefficients can lead to a downward sloping curve. The switch in the slope is an interesting finding that outlines the importance of taking into account the uncertainty surrounding these estimates. One can even speculate that strong biases in investor beliefs may give a rise to the puzzling relationships documented in van Binsbergen et al. (2012). However, in the range of our mean reversion estimates (0.74-0.84) the estimated term structure is upward sloping.

The persistence of cash flow reversion matters in two significant ways. When the firm's transition to the steady state is very short, all firms behave similarly irrespective of their initial differences in growth rates and profitability. As a result, the pricing differences across firms would be largely affected by the initial book-to-market ratios across firms, as in Dechow et al. (2004) and Weber (2018). Thus, the identification of short- versus long-run risk premia would be weaker and estimates are naturally much more noisy and extreme. On the other hand, the persistence can affect which firms earn a higher expected return. Largely profitable firms tend to have most of their value in the immediately and intermediate future since all firms later converge to the same steady state. A decrease in the speed of cash flow mean reversion (higher AR coefficient) would unilaterally increase their expected return. On the other hand, the unprofitable firms (the flip side of profitable firms) may experience a decrease in expected return due to much more persistent losses (the relationship is not entirely clear if there are non-monotinicities). This can potentially shed some light on why the term structure ultimately depends on the persistence parameters.

To conduct a similar experiment, we also alter the forecasting horizon to 10 and 20 years, and assume different steady state growth rates for g. As outlined in Table IX, this again generates two insights. Firstly, it appears that as long as the forecasting period is long enough for the convergence to occur, the slope sign remains unchanged. Although the magnitude of estimates does depend on the forecasting horizon (as more and more value of the firm is attributed to the terminal horizon), the slope of the equity term structure does not. Second, the steady state rate only mildly affects the slope estimates, except for the high growth and short horizon combination. Yet, a long-term cash flow growth rate of 9% is both rather exceptional and convergence at 10 years horizon is unlikely to fully occur. Overall, the qualitative results seem quite robust to the selection of g and T.

Table VIII: Sensitivity of NS estimates

The table presents the sensitivity of slope and level estimates to changes in ROE and sales growth reversion parameters governing the cash flow dynamics. In all specifications below, the equilibrium cash flow growth rate g and forecasting horizon T are set to 6% and 15, respectively. For brevity, the level estimates are excluded from the table. The negative slope estimates are shaded in dark grey. Compustat-CRSP-I/B/E/S sample: 1980-2021.

slope		AR(1) ROE					
		0.4	0.5	0.6	0.7	0.8	0.9
AR(1) sales	0.2	0.20	0.19	0.17	0.15	0.10	0.01
	0.3	0.19	0.18	0.17	0.14	0.10	0.01
	0.4	0.19	0.18	0.16	0.14	0.09	-0.01
	0.5	0.17	0.16	0.15	0.12	0.08	-0.03
	0.6	0.15	0.14	0.13	0.10	0.05	-0.06

level		AR(1) ROE					
		0.4	0.5	0.6	0.7	0.8	0.9
AR(1) sales	0.2	0.09	0.09	0.08	0.08	0.08	0.09
	0.3	0.08	0.08	0.08	0.08	0.08	0.09
	0.4	0.08	0.08	0.08	0.08	0.08	0.09
	0.5	0.08	0.08	0.08	0.08	0.08	0.09
	0.6	0.08	0.08	0.08	0.08	0.08	0.09

Table IX: Sensitivity of NS estimates

The table presents the sensitivity of slope and level estimates to changes the equilibrium cash flow growth rate g and forecasting horizon T. In all specifications below, the ROE and sales growth mean reversion parameters are set to 0.81 and 0.41, respectively. For brevity, the level estimates are excluded from the table. The negative slope estimates are shaded in dark grey. Compustat-CRSP-I/B/E/S sample: 1980-2021.

slope		T			level		T		
		10	15	20			10	15	20
	0.03	0.05	0.07	0.08		0.03	0.07	0.05	0.05
g	0.06	0.03	0.08	0.10	g	0.06	0.09	0.08	0.09
	0.09	-0.02	0.08	0.11		0.09	0.12	0.12	0.12

3.6.5 Industry portfolios

As a final robustness check we perform another cross-sectional sort, where we sort stocks into the 15 Fama-French industries. Because, arguably, the technology sector has changed most substantially over our sample period, we separate technology firms into a separate category following Barron, Byard, Kile & Riedl (2002). In Table C.III the term structure estimates are shown. Again we find largely positively sloped term structures. Only for the steel works industry the slope estimate is negative. In addition, the level estimates (the 15-year risk premium) are fairly stable across industries. While almost all slope estimates are positive, we do see substantial variation across industries in the size of the slope estimate. It seems that consumer-oriented industries have the steepest term structures. At the same time, this variation may be partially due to estimation noise, since we split the cross-section into 16 portfolios in this analysis.

3.7 Conclusion

The empirical literature documenting the properties of the term structure of expected equity returns is divided. The prediction of the seminal asset pricing models that longer maturity assets contain more systematic risk was put to test by a series of papers by van Binsbergen et al. (2012), van Binsbergen et al. (2013), van Binsbergen & Koijen (2017). Specifically, the price data in options and dividend strips markets seem to embed a downward-sloping equity term structure meaning that most established models are not in line with empirics. Nevertheless, the evidence is not without controversy due to relatively short historical samples, containing two extreme crashes in 2001 and 2007-2008, which makes the realized returns noisy proxies of the unconditional average expected returns. The challenge is further exacerbated by liquidity effects in derivative markets (Bansal et al. (2019)).

This paper proposes a new methodology to estimate the term structure of equity risk premia based on ex-ante returns. We do so by exploiting the cross-sectional heterogeneity across firm cash flow forecasts and combining them with the implied cost of capital approach. Rather than sorting into duration portfolios and looking at noisy realized returns, we estimate the risk premium term structure by minimizing errors between the discounted cash flow model price of the firm's equity (in the spirit of Dechow et al. (2004)) and the observed market prices. To incorporate as much identifying variation as possible, we augment Dechow et al. (2004) with the information extracted from the IBES analyst forecasts. Our approach does not require prices of dividend assets and can thus be implemented for longer sample periods, 1985 – 2019 in our estimation for the U.S. market. Also, our approach allows to investigate cross-sectional and time-series variation in risk premium term structures.

Our baseline results show that the equity term structure is upward sloping at the aggregate level, consistent with predictions of classical asset pricing theories such as Campbell & Cochrane (1999) and Bansal & Yaron (2004). However, we also find substantial undocumented heterogeneity between small and large, value and growth firms, with small value firms exhibiting a downward sloping term structure.

We do find evidence that the term structure becomes flatter during recessions, but it remains upward sloping also in these periods. Adding to that, in this paper we cast some light on how credit risk is priced in firms. The implied discount rates are monotonically increasing with credit risk, proxied by the credit ratings, and are particularly high for speculative grade firms. In addition, for the latter firms the term structure is flat or slightly downward sloping for small firms, indeed capturing the idea that these firms either go bust in the upcoming years or survive long enough and transition into higher rated firms. All in all, our ex-ante measure of returns does not support the credit risk anomaly detected in realized returns (Campbell et al. (2008)).

In sum, our key contribution is to develop and implement a coherent framework that merges the cross-section of stock prices and implied cost of capital approach to recover the term structure of equity risk premia. For the modeling cash flows we closely follow the original framework of Dechow et al. (2004) and Weber (2018), and amend it with the information extracted from the analysts forecasts. In doing so, we incorporate more heterogeneity in cash flow expectations, necessary to identify separate horizon discount rates, while at the same time keeping the parsimonious structure of the expected cash flow process. However, we diverge from Dechow et al. (2004) and Weber (2018) by not looking at the realized returns of different duration portfolios, but instead implying the expected returns. As shown, this step allows to flexibly measure ex-ante returns and inspect the cross-sectional and time-varying properties of the equity term structure.

References

Bansal, R., Miller, S., Song, D. & Yaron, A. (2019), The term structure of equity risk premia. Working Paper.

Bansal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The Journal of Finance* **59**(4), 1481–1509.

Barron, O. E., Byard, D., Kile, C. & Riedl, E. J. (2002), 'High-technology intangibles and analysts' forecasts', *Journal of Accounting Research* **40**(2), 289–312.

Bordalo, P., Gennaioli, N., Porta, R. L. & Shleifer, A. (2019), 'Diagnostic expectations and stock returns', *The Journal of Finance* **74**(6), 2839–2874.

Campbell, J. Y. & Cochrane, J. H. (1999), 'By force of habit: A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* **107**(2), 205–251.

Campbell, J. Y., Hilscher, J. & Szilagyi, J. (2008), 'In search of distress risk', *Journal of Finance* **63**(6), 2899–2939.

Cassella, S., Golez, B., Gulen, H. & Kelly, P. (2023), 'Horizon bias and the term structure of equity returns', *The Review of Financial Studies* **36**(3), 1253–1288.

Chen, S. & Li, T. (2018), 'A unified duration-based explanation of the value, profitability, and investment anomalies', *Profitability, and Investment Anomalies* .

Claus, J. & Thomas, J. (2001), 'Equity premia as low as three percent? evidence from analysts' earnings forecasts for domestic and international stock markets', *The Journal of Finance* **56**(5), 1629–1666.

Cochrane, J. H. (2011), 'Presidential address: Discount rates', *Journal of Finance* **66**(4), 1047–1108.

Davis, J. L., Fama, E. F. & French, K. R. (2000), 'Characteristics, covariances, and average returns: 1929 to 1997', *The Journal of Finance* **55**(1), 389–406.

Dechow, M. P., Sloan, R. G. & Soliman, M. T. (2004), 'Implied equity duration: A new measure of equity risk', *Review of Accounting Studies* **9**(2-3), 197–228.

Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds. journal of financial economics', *Journal of financial economicse* **33**(1), 3–56.

Fama, E. F. & French, K. R. (2002), 'The equity premium', *The Journal of Finance* **57**(2), 637–659.

Gebhardt, W. R., Lee, C. M. & Swaminathan, B. (2001), 'Toward an implied cost of capital', *Journal of accounting research* **39**(1), 135–176.

Giglio, S., Kelly, B. T. & Kozak, S. (2022), Equity term structures without dividend strips data. Available on SSRN.

Gormsen, N. J. (2020), Time variation of the equity term structure. Available on SSRN.

Gormsen, N. J. & Lazarus, E. (2022), 'Duration-driven returns', *Journal of Finance* .

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2007), 'Treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**(8), 2291–2304.

Lettau, M. & Wachter, J. A. (2007), 'Why is long-horizon equity less risky? a duration-based explanation of the value premium', *The Journal of Finance* **62**(1), 55–92.

Li, Y., Ng, D. T. & Swaminathan, B. (2013), 'Predicting market returns using aggregate implied cost of capital', *Journal of Financial Economics* **110**(2), 419–436.

Nelson, C. R. & Siegel, A. F. (1987), 'Parsimonious modeling of yield curves', *Journal of business* pp. 473–489.

Nissim, D. & Penman, S. H. (2001), 'Ratio analysis and equity valuation: From research to practice', *Review of accounting studies* **6**(1), 109–154.

Pástor, L., Sinha, M. & Swaminathan, B. (2008), 'Estimating the intertemporal risk–return tradeoff using the implied cost of capital', *The Journal of Finance* **63**(6), 2859–2897.

van Binsbergen, J., Brandt, M. & Koijen, R. (2012), 'On the timing and pricing of dividends', *American Economic Review* **102**(4), 1596–1618.

van Binsbergen, J. H., Han, X. & Lopez-Lira, A. (2022), 'Man vs. machine learning: The term structure of earnings expectations and conditional biases', *The Review of Financial Studies* .

van Binsbergen, J., Hueskes, W., Koijen, R. & Vrugt, E. (2013), 'Equity yields', *Journal of Financial Economics* **110**(3), 503–519.

van Binsbergen, J. & Koijen, R. (2017), 'The term structure of returns: Facts and theory', *Journal of Financial Economics* **124**(1), 1–21.

Weber, M. (2018), 'Cash flow duration and the term structure of equity returns', *Journal of Financial Economics* **128**(3), 486–503.

Appendices

3.A Cash Flow Forecasts

1. Integrating analyst forecasts

We employ and modify the analyst data in the following way:

1. as in Li et al. (2013), we construct the constant maturity median forecasts for revenues, ROE and EPS as a weighted average between year $t+m$ and $t+m+1$ forecasts. This step accounts for the differences in financial years across firms creating discrepancies in timing of cash flows and their discounting. More details on this procedure are provided in Appendix A.2.

2. if a firm has revenue forecasts, we used them for the 1-5 year growth rate of book equity projections (depending on the availability). Namely, we calculate the m'th year sales growth rates as $forecastSALE_{t+m}/forecastSALE_{t+m-1} - 1$. If a firm has a non-missing IBES long-term growth rate forecast for sales, we extrapolate all remaining missing values for the 2-5 year horizon using this rate. All resulting sales growth rates are additionally winsorized at 5 and 95% level.

3. if a firm has ROE forecasts, we used them for years 1-5 (depending on the availability). If the forecasts are not available, we proceed with step 4. All ROE forecasts are winsorized at 5 and 95% level.

4. if EPS forecasts are available when ROE are not, we perform the following exercise. For years 1-4, we projected the book value of equity using the mean-reverting sales growth process with an AR coefficient of 0.41 and long-term mean of 6% (à la Dechow et al. (2004) and Weber (2018)). Again, if a firm has a non-missing long-term growth rate forecast for EPS and the last forecast value was positive, we extrapolated all remaining EPS missing values for the 2-5 year horizon using this rate. Then, we implied the ROE forecasts for maturity m using this formula: $EPS_{t+m} * \#shares_t / BE_{t+m-1}$.[20]. All the implied forecasts are winsorized at 5 and 95% level.

2. Constant maturity forecast

Forecasts are interpolated using this principle:

$$12monthForecast = w * FY1 + (1 - w) * FY2 \tag{10}$$

where $FY1$ and $FY2$ are the current and subsequent financial year I/B/E/S forecasts and w stands for the number of months left until the realization of current financial year cash flow divided by 12. Since we have forecasts for up to 5 years, we repeat the same procedure for 24, 36, and 48 month forecasts.

[20] The historical number of shares outstanding is assumed to remain fixed in the future.

Following Li et al. (2013), for the last available year of forecasts we apply this rule to extrapolate the constant maturity forecast:

$$24monthForecast = 12monthForecast * g \tag{11}$$

where g is the implicit growth between $FY1$ and $FY2$ forecasts. To reduce the effect of outliers, we winsorize this g at 5 and 95% level. Clearly, if $FY1$ and $FY12$ switch signs, the growth rates become ill-defined. Therefore, we do not apply this rule for those cases.

The underlying assumption is that firm cash flows realize at the end of the firm's financial year. Because most firms have their financial year ending on Dec 31, we choose June as a valuation month.

As a result of this procedure, the sample shrinks from roughly 65,000 to 53,000 observations. The main reason is that this method requires at least 2 years of forecast data which is not available for all firms in the IBES database.

Data Filters

There are 3 key dates in the I/B/E/S summary database:

- $statpers$ - the date on which the forecasts are summarized in this database (e.g. 18 Dec 2014)
- $fpedats$ - the end date of the forecasting period (e.g. 31 Dec 2014)
- *anndatsact* - the date when the actuals, i.e. financial statements, for that forecasting period were announced (e.g. 30 Jan 2015)

Filters:

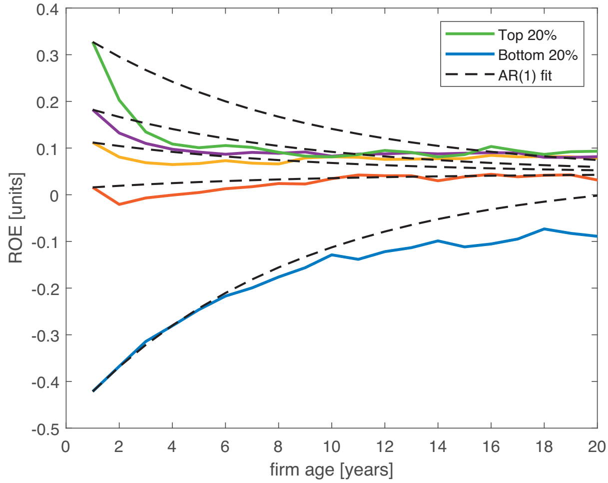
1. The valuation date is June. Thus, only forecasts available in this month are used.
2. We account for potential mistakes by: i) dropping all observations where the forecast record day (*statpers*) is later than the announcement day of actuals (*anndatsact*); ii) dropping all observations where the forecast record day (*statpers*) exceeds the corresponding forecasting period end date ($fpedats$) by more than 4 months (122 days)[21]. The idea is that forecasts announced after the actual financial statements are no longer forecasts. Adding to that, it should be impossible to observe a 2-year forecast on 18 Jun 2014 for a financial year ending on 31 Jun 2014.
3. We keep firms that have forecasts denominated in US dollars. This filter deletes double entries and eliminates concerns that the reporting currency in I/B/E/S does not match the balance sheet data in Compustat (which is important for measures like EPS, but not for ROE though).

[21] The filter is also applied to all mismatches between 1-5 year actual and implied forecasting period end dates calculated as $statpers + forecastinghorizon$.

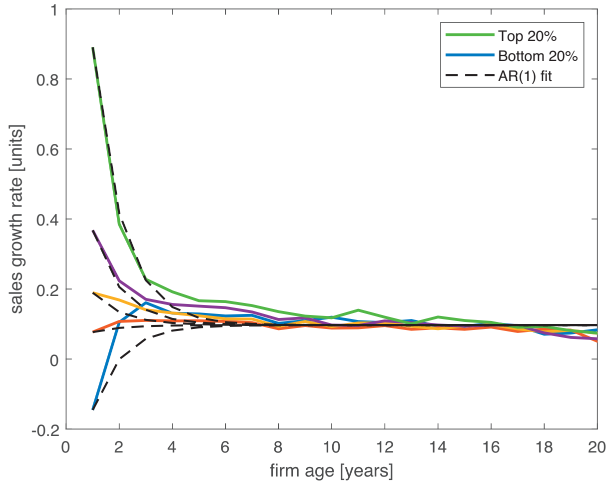
4. In case more than one forecast measure is used, we follow the conservative approach and i) always choose the latest announcement date of actuals across measures, ii) delete all observations where the forecasting period end dates do not match across measures.

3.B Additional Tables and Figures

Figure C.I: ROE and Sales growth quintile portfolios



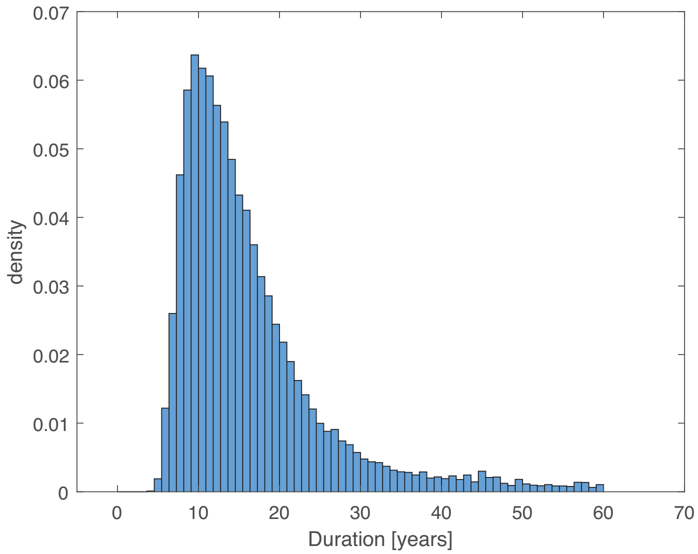
(a) ROE quintiles



(b) Sales growth quintiles

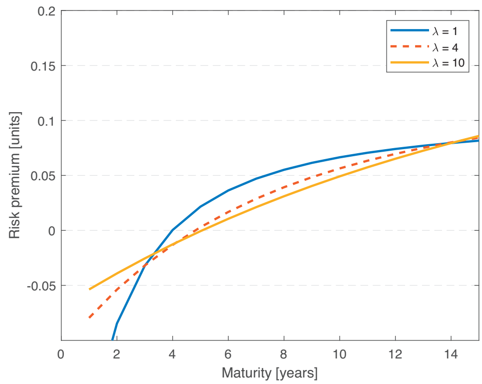
Note: The graph shows the term structure of firm's ROE and sales growth across firm age. Each solid line represents an average of the ROE (sales growth) quintile formed when the age equals 1. The age is proxied by the first occurrence of the firm in the Compustat database. Each dashed line represents the least squares fit of an AR(1) model, where AR(1) for ROE and sales growth portfolios equal to 0.89 and 0.40 respectively. Sample range: 1964-2021.

Figure C.II: Duration distribution

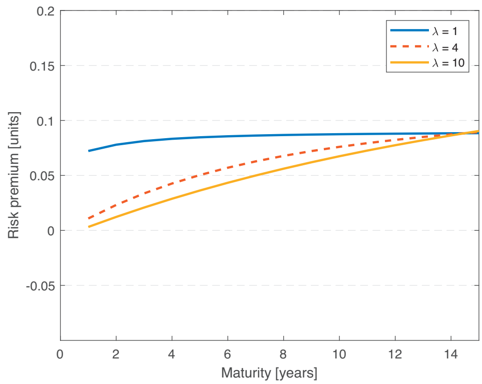


Note: The figure illustrates the year-firm cash flow duration distribution in the merged Compustat-CRSP-I/B/E/S sample (1980-2021). In order to calculate the cash flow duration, we use the model value of a firm, as in equation (3.2), and the equity term structure estimates from Figure I. For illustrative purposes, the duration is capped at 60-years.

Figure C.III: Sensitivity to λ



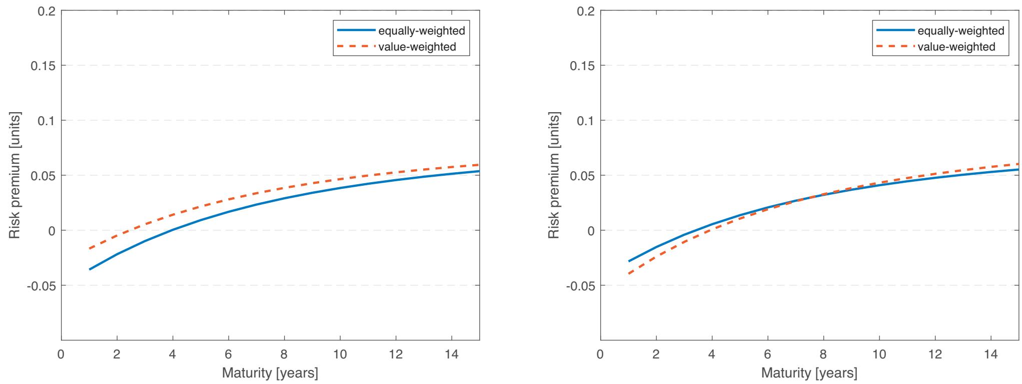
(a) Value weighting



(b) Equal weighting

Note: The figures illustrates the sensitivity of baseline estimates to changes in the decay parameter λ. Compustat-CRSP-I/B/E/S subsample: 1980-2021.

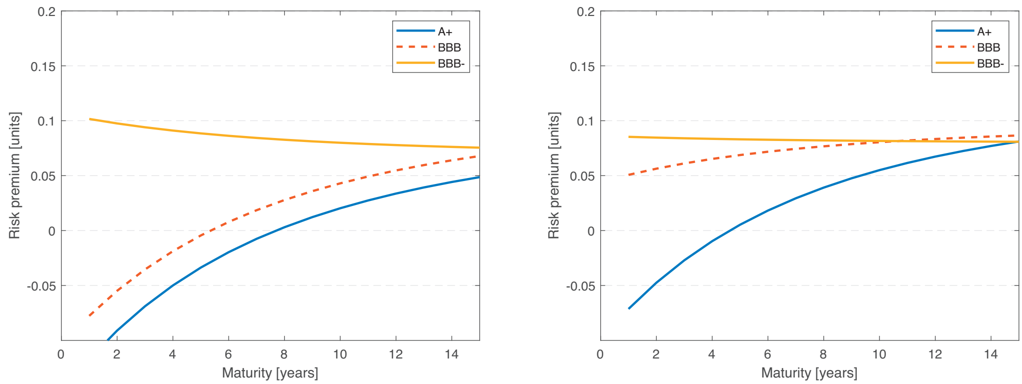
Figure C.IV: van Binsbergen, Brandt & Koijen (2012) sample



(a) Top size quintile **(b) S&P500 firms**

Note: The figures depict the estimated equity risk premia term structure in the van Binsbergen et al. (2012) sample proxied by the top size quintile firms (left graph) or the $S\&P500$ index firms in our sample (right graph). The blue solid line represents the benchmark value-weighted results, whereas the red dashed line presents results where we equally weight firms in the estimation. Compustat-CRSP-I/B/E/S subsample: 1996-2008.

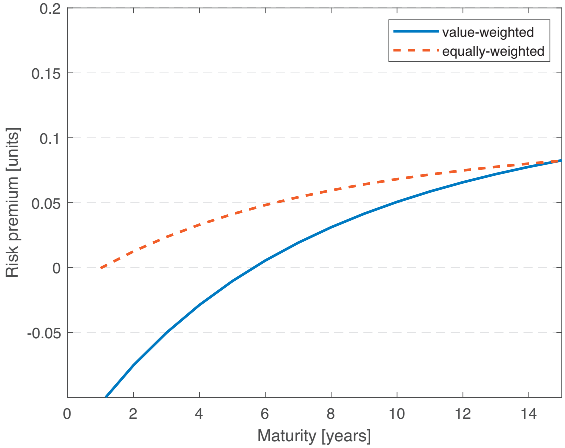
Figure C.V: Double sorts on credit rating and size



(a) small firms **(b) large firms**

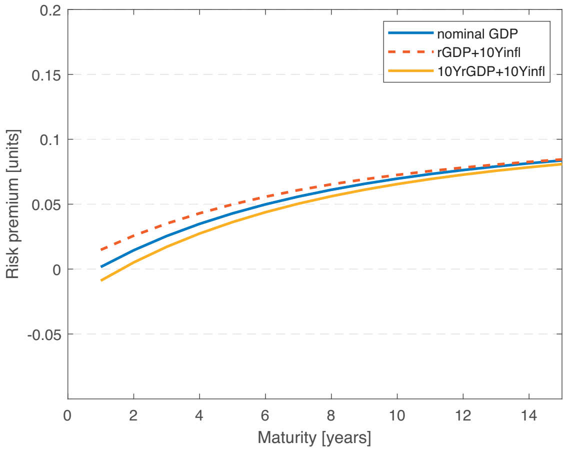
Note: The figures present the estimated equity risk premia term structure in different credit rating and size portfolios. The observations are first sorted based on the S&P credit rating and then split into two median size portfolios. A+ group contains all credit ratings above and including an $A-$ rating. BBB group includes all $BBB+$, BBB and $BBB-$ rated firms. Finally, BBB- includes all non-investment grade ratings. Compustat-CRSP-I/B/E/S subsample: 1985-2018.

Figure C.VI: Cash flows with corrected analyst bias



Note: The figure presents the estimated equity risk premium term structure with corrected analyst forecasts. The correction is made by using estimates from van Binsbergen et al. (2022) (the procedure is described more thoroughly in Section 3.6.1). Compustat-CRSP-I/B/E/S sample: 1980-2021.

Figure C.VII: Time-varying g



Note: The figures illustrate equity term structure estimates using alternative proxies for the long-term cash flow growth rate g. The blue line represents estimates where g equals the average nominal GDP growth rate up to year t, as in Li et al. (2013) (Compustat-CRSP-I/B/E/S sample: 1980-2021). The red dashed line approximates g with the average real GDP growth rate up to year t plus the long-term (10Y) inflation forecasts (Compustat-CRSP-I/B/E/S subsample: 1982-2018). The yellow line measures g as a sum of the long-term (10Y) real GDP and long-term (10Y) inflation forecasts (Compustat-CRSP-I/B/E/S subsample: 1992-2021).

Table C.I: Analyst forecast bias in EPS

The table presents the average bias, its standard error, and the number of observations of analyst EPS forecast biases obtained from van Binsbergen et al. (2022). The standard errors are clustered by time dimension. Compustat-CRSP-I/B/E/S sample: 1980-2021.

	Analyst bias					
	1-year			2-year		
	mean	SE	N	mean	SE	N
Full sample	0.009	0.002	40,292	0.025	0.003	37,861
Normal times	0.006	0.001	34,683	0.021	0.003	32,486
NBER recessions	0.023	0.008	5,609	0.050	0.010	5,347

Table C.II: Additional filters and model averaging

The table presents the sensitivity of level and slope estimates to alternative filters and cash flow models. The first and the second row stand for scenarios where cash flow forecasts are averaged across two approaches (the mean-reversion model and the augmented version with I/B/E/S forecasts) using the equal-weighting scheme or weights based on the normal distribution, where the mean and the standard deviation are based on the empirical values. The third row represents the case where the estimation sample is trimmed at the higher level of 5th and 95th percentile of the sample IRR. The forth and the fifth rows describe scenarios where different procedures are combined. Finally, the last row summarizes the results when the estimation is performed excluding the most extreme book-to-market portfolios within each size quintile. Compustat-CRSP-I/B/E/S subsample: 1980-2021.

	level	slope
1) cash flows equally averaged	0.086	0.165
2) cash flows averaged using normal dist	0.089	0.189
3) trimmed at 5% and 95% sample IRR	0.087	0.097
combination of 1) & 3)	0.086	0.175
combination of 2) & 3)	0.089	0.200
excluding extreme BM portfolios	0.087	0.201

Table C.III: Industry analysis

The table presents the estimation results within Fama-French 15 industries. The technology industry classification is constructed separately following Barron et al. (2002). Weight denotes the market cap weight of the industry-specific observations in the total sample. All Fama-French industries are sorted based on the slope estimates, and negative slopes are highlighted in gray. Compustat-CRSP-I/B/E/S subsample: 1980-2021.

	level	slope	weight
Transportation	0.09	0.04	0.04
Oil and Petroleum Products	0.08	0.05	0.10
Steel Works Etc	0.08	0.06	0.01
Other	0.08	0.06	0.32
Automobiles	0.09	0.07	0.02
Mining and Minerals	0.08	0.07	0.01
Chemicals	0.08	0.10	0.02
Textiles, Apparel and Footware	0.07	0.11	0.01
Machinery and Business Equipment	0.09	0.12	0.13
Construction and Construction Materials	0.08	0.12	0.02
Fabricated Products	0.07	0.12	0.00
Food	0.08	0.14	0.05
Consumer Durables	0.07	0.16	0.01
Drugs, Soap, Prfums, Tobacco	0.10	0.17	0.12
Retail Stores	0.09	0.18	0.04
Technology	0.09	0.10	0.37

3.C Other Derivations

Notice that an m-maturity cash flow of firm i at time t:

$$\mathbb{E}_t[CF_{i,t+m}] = \mathbb{E}_t[CF_{i,t+m-1}(1 + growth_{i,t+m})]$$

where $growth_{i,t+m}$ is the growth rate of cash flows from period $t+m-1$ to $t+m$ ($m \geq 1$). We can rewrite this expression in terms of cash flow $CF_{i,t}$:

$$\begin{aligned}\mathbb{E}_t[CF_{i,t+m}] &= \mathbb{E}_t[\mathbb{E}_{t+m-1}[CF_{i,t+m-1}(1 + growth_{i,t+m})]] = \mathbb{E}_t[CF_{i,t+m-1}\,\mathbb{E}_{t+m-1}[1 + growth_{i,t+m}]] \\ &= \mathbb{E}_t[CF_{i,t+m-1}(1 + g_{i,t+m})] = CF_{i,t}\prod_{s=1}^{m}(1 + g_{i,t+s})\end{aligned}$$

where $g_{i,t+s} \equiv \mathbb{E}_{t+s-1}[growth_{i,t+s}]$ is the conditional expected growth rate of cash flows from period $t+s-1$ to $t+s$.

Since we make an assumption that all firms after T years converge to the same equilibrium growth rate g_t, for $m > T$, the expression simplifies to:

$$\mathbb{E}_t[CF_{i,t+m}] = CF_{i,t}\prod_{s=1}^{T}(1 + g_{i,t+s})(1 + g_t)^{m-T} = \mathbb{E}_t[CF_{i,t+T}](1 + g_t)^{m-T}$$

This implies that the present value of cash flows after year T is just:

$$\sum_{m=T+1}^{\infty}\frac{E_t[CF_{i,t+m}]}{(1 + r_{i,t,T})^m} = \frac{E_t[CF_{i,t+T+1}]}{(1 + r_{i,t,T})^T(r_{i,t,T} - g_t)}$$

and the equation (3.2) follows.

CENTER DISSERTATION SERIES

CentER for Economic Research, Tilburg University, the Netherlands

No.	Author	Title	ISBN	Published
672	Joobin Ordoobody	The Interplay of Structural and Individual Characteristics	978 90 5668 674 1	February 2022
673	Lucas Avezum	Essays on Bank Regulation and Supervision	978 90 5668 675 8	March 2022
674	Oliver Wichert	Unit-Root Tests in High-Dimensional Panels	978 90 5668 676 5	April 2022
675	Martijn de Vries	Theoretical Asset Pricing under Behavioral Decision Making	978 90 5668 677 2	June 2022
676	Hanan Ahmed	Extreme Value Statistics using Related Variables	978 90 5668 678 9	June 2022
677	Jan Paulick	Financial Market Information Infrastructures: Essays on Liquidity, Participant Behavior, and Information Extraction	978 90 5668 679 6	June 2022
678	Freek van Gils	Essays on Social Media and Democracy	978 90 5668 680 2	June 2022
679	Suzanne Bies	Examining the Effectiveness of Activation Techniques on Consumer Behavior in Temporary Loyalty Programs	978 90 5668 681 9	July 2022
680	Qinnan Ruan	Management Control Systems and Ethical Decision Making	978 90 5668 682 6	June 2022
681	Lingbo Shen	Essays on Behavioral Finance and Corporate Finance	978 90 5668 683 3	August 2022
682	Joshua Eckblad	Mind the Gales: An Attention-Based View of Startup Investment Arms	978 90 5668 684 0	August 2022
683	Rafael Greminger	Essays on Consumer Search	978 90 5668 685 7	August 2022
684	Suraj Upadhyay	Essay on policies to curb rising healthcare expenditures	978 90 5668 686 4	September 2022

No.	Author	Title	ISBN	Published
685	Bert-Jan Butijn	From Legal Contracts to Smart Contracts and Back Again: An Automated Approach	978 90 5668 687 1	September 2022
686	Sytse Duiverman	Four essays on the quality of auditing: Causes and consequences	978 90 5668 688 8	October 2022
687	Lucas Slot	Asymptotic Analysis of Semidefinite Bounds for Polynomial Optimization and Independent Sets in Geometric Hypergraphs	978 90 5668 689 5	September 2022
688	Daniel Brosch	Symmetry reduction in convex optimization with applications in combinatorics	978 90 5668 690 1	October 2022
689	Emil Uduwalage	Essays on Corporate Governance in Sri Lanka	978 90 5668 691 8	October 2022
690	Mingjia Xie	Essays on Education and Health Economics	978 90 5668 692 5	October 2022
691	Peerawat Samranchit	Competition in Digital Markets	978 90 5668 693 2	October 2022
692	Jop Schouten	Cooperation, allocation and strategy in interactive decision-making	978 90 5668 694 9	December 2022
693	Pepijn Wissing	Spectral Characterizations of Complex Unit Gain Graphs	978 90 5668 695 6	November 2022
694	Joris Berns	CEO attention, emotion, and communication in corporate financial distress	978 90 5668 696 3	November 2022
695	Tom Aben	The (long) road towards smart management and maintenance: Organising the digital transformation of critical infrastructures	978 90 5668 697 0	December 2022
696	Gülbike Mirzaoğlu	Essays in Economics of Crime Prevention and Behavior Under Uncertainty	978 90 5668 698 7	February 2023
697	Suwei An	Essays on incentive contracts, M&As, and firm risk	978 90 5668 699 4	February 2023
698	Jorgo Goossens	Non-standard Preferences in Asset Pricing and Household Finance	978 90 5668 700 7	February 2023

No.	Author	Title	ISBN	Published
699	Santiago Bohorquez Correa	Risk and rewards of residential energy efficiency	978 90 5668 701 4	April 2023
700	Gleb Gertsman	Behavioral Preferences and Beliefs in Asset Pricing	978 90 5668 702 1	May 2023
701	Gabriella Massenz	On the Behavioral Effects of Tax Policy	978 90 5668 703 8	May 2023
702	Yeqiu Zheng	The Effect of Language and Temporal Focus on Cognition, Economic Behaviour, and Well-Being	978 90 5668 704 5	May 2023
703	Michela Bonani	Essays on Innovation, Cooperation, and Competition Under Standardization	978 90 5668 705 2	June 2023
704	Fabien Ize	The Role of Transparency in Fairness and Reciprocity Issues in Manager-Employee Relationships	978 90 5668 706 9	June 2023
705	Kristel de Nobrega	Cyber Defensive Capacity and Capability: A Perspective from the Financial Sector of a Small State	978 90 5668 707 6	July 2023
706	Christian Peters	The Microfoundations of Audit Quality	978 90 5668 708 3	June 2023
707	Felix Kirschner	Conic Optimization with Applications in Finance and Approximation Theory	978 90 5668 709 0	July 2023
708	Zili Su	Essays on Equity Incentive and Share Pledging in China	978 90 5668 710 6	September 2023
709	Rafael Escamilla	Managing the Nanostore Supply Chain: Base-of-the-Pyramid Retail in Emerging Markets	978 90 5668 711 3	September 2023
710	Tomas Jankauskas	Essays in Empirical Finance	978 90 5668 712 0	August 2023

Tomas Jankauskas (Vilnius, Lithuania, 1993) received his Bachelor's degree in Economics at Vilnius University in 2016. In 2018 he obtained his Research master's degree in Finance at Tilburg University. In the same year, he started writing his dissertation as a Ph.D. candidate under the supervision of Prof. Dr. Joost Driessen and Prof. Dr. Lieven Baele. In the spring of 2019, he visited the European Central Bank, and in the fall of 2022, he visited Yale University as a Visiting Researcher, sponsored by Prof. Dr. Stefano Giglio. In 2023 he will join the New York Federal Reserve Bank as a Financial Research Economist.

This dissertation is a collection of three chapters on the unconventional monetary policy effects and the term structure of corporate bond and equity risk premia. The first chapter sheds light on the impact of the recent European Central Bank lending policies on the financial sector stability. The key finding is that the central bank interventions reduced the fragility of the banking sector, predominantly by reducing the severity and prevalence of bank-run equilibria. The second chapter provides a new implied estimate of the term structure of corporate bond risk premia. The main result is that the maturity-matched risk premium is increasing with maturity, in line with most canonical asset pricing models; however, the cross-sectional sorts reveal a high degree of heterogeneity in the term structure slope. Finally, the third chapter proposes a novel methodology to estimate the term structure of equity term structure for any cross-sectional portfolio. The method documents similar qualitative patterns for equities as the second chapter presents for corporate bonds.

ISBN: 978 90 5668 712 0
DOI: 10.26116/qh2x-nq45

TILBURG SCHOOL OF ECONOMICS
AND MANAGEMENT

CentER Dissertation Series