Raffaele Corvino

Doctoral Thesis

Department of Finance, Cass Business School
City University of London

Essays in Asset Pricing and Corporate Finance

Supervisor:
Gianluca Fusai

Supervisor:
Enrique Schroth

PhD in Finance - March, 2019

Contents

List of Figures

List of Tables

*I am incredibly indebted to Giovanna Nicodano, who guided me throughout all my doctoral education with her extraordinary professional experience, her inestimable human qualities, and her timely and precious advises. I am grateful to Gianluca Fusai, who has believed in me since I was a Master student, has supervised my works from the very beginning of my PhD, and has imparted me new invaluable skills. I am also grateful to Enrique Schroth, who has offered to me a new perspective on academic research and has triggered in me great enthusiasm for working hard during my years at Cass. I want to thank intensely Dino Palazzo, Lucio Sarno, and Stefano Sacchetto, who has provided me extremely useful and constant feedbacks, much helpful comments and suggestions to my works, and have been always present and available. The advises of Giovanna, Gianluca, Enrique, Dino, Lucio, and Stefano have been crucial to achieve this target, and their support on the job market has been outstanding. I want also to thank Giovanni Cespa, my current PhD director, Ian Marsh, my former director, and all the faculty of finance at Cass, Fabio Bagliano and Carolina Fugazza at the University of Torino, my dearest friend and co-author Francesco Ruggiero, the Risk Management Institute at the National University of Singapore, where I have spent few but wonderful months during my visiting, and Claudio Zucca at the European Investment Bank. Finally, I want to thank Josef Zechner, who has discussed my thesis and has provided me very insightful comments to improve my papers towards the next steps of my academic career.

Abstract

The thesis investigates the asset pricing implications of different issues arising in financial markets: the heterogeneity across types of investors and their ownership policy over time, the premium required by the investors for bearing the default risk of a company, and the proper compensation owed to the investors for lending to riskier countries.

In the first chapter, I quantify private benefits of control, and their impact on stock prices, by estimating a structural model of optimal shareholding using data on the ownership dynamics of Italian public companies. The results show that controlling shareholders (i) extract private benefits on average around 2% of equity value, and (ii) generally have positive and persistent impact on stock prices. The results imply that controlling shareholders extract private benefits without cost for the rest of the company shareholders. I also provide evidence of a synergistic effect when the largest shareholder is a corporation.

In the second chapter, which is coauthored with Gianluca Fusai, we estimate a credit risk structural model using information from both credit and equity markets. We infer the dynamics of asset and debt, and the default boundary, for worldwide non-financial firms. Using our estimation results, we compute the premium that compensates the investors for bearing the default risk of the firm. Exploiting the insights from the structural model, we address the relation between the premium and both credit default swaps (CDS) spreads and equity prices. We provide evidence that the dynamics of the premium may match with opposite dynamics of credit and equity securities across firms, depending on whether the firm generates positive or negative excess return on the asset.

In the last chapter, which is coauthored with Francesco Ruggiero, we analyse the relative pricing between sovereign credit default swap (CDS) spreads and sovereign bond yields for European countries during and after the sovereign debt crisis of 2010-2012. We investigate whether riskier countries compensate their debtholders properly by paying out sufficiently higher bond yields compared to those of safer countries. We test whether the differences across countries in terms of the default risk priced in the CDS spreads are consistently priced in the cross section of the bond yields, and we show that an inconsistent cross-sectional relationship between CDS spreads and bond yields emerges during the crisis period for all European countries. However,

after the announcement of the Outright Monetary Transaction (OMT) program by the European Central Bank, the consistent cross-sectional relationship between default risk and bond yields is restored for the Eurozone countries only.

1. Dynamic Ownership, Private Benefits, and Stock Prices

1.1. Introduction

Controlling shareholders have additional motives to hold shares in the company compared to minority shareholders, as they are able to extract private benefits such as social status, public prestige, or discretionary power to divert cash flows or to pay excessive compensation to blockholders or their relatives. Private benefits of control, then, can be an important driver in the controlling shareholders' choice about the size of the initial ownership share, and about their subsequent trading decisions. Since the controlling shareholders' trading decisions affect the formation of investors' beliefs and the amount of shares floating on the market, private benefits may have significant impact on stock prices.

In this paper, I present and estimate a dynamic model of optimal shareholding to quantify the private benefits of controlling shareholders and to measure the impact of private benefits of control on stock prices over time.

In the model, a large shareholder and a mass of marginal investors hold shares in a company. The marginal investors are uninformed about the fundamental value of the firm, and they trade on their heterogenous expectations, which they revise over time using two pieces of information: the shocks to the fundamental value of the firm and the trading decision of the large shareholder, who has perfect information over the true value of the firm. However, the large shareholder also extracts private benefits from the stake, so that the information released by his trading is noisy.

The large shareholder trades off stock mispricing and private benefits against risk diversification and price impact of the trade, where the amount of private benefits extracted from the stake depends on the attainment of given thresholds of stake (for instance, 50% for the control of the firm). This assumption is in line with the institutional framework, according to which shareholders' rights and obligations arise as soon as shareholders get hold of a given percentage of the outstanding shares.

In the absence of private benefits, the large shareholder always trades on the mispricing of the marginal investors. However, the mispricing reduces over time as the marginal investors learn from the large shareholder's trades. Thus, it becomes less profitable for the latter to exploit the mispricing. With additional private benefits, instead, the large shareholder may not trade at all. The reason is that the large shareholder sells a block of shares only when the share is largely overvalued, so that the gains offset both the loss in private benefits and shares depreciation, given the negative price impact of the trade. Specular reasoning applies to the purchase of an additional block of shares.

The price impact of private benefits is two-fold. First, private benefits affect the decision of the large shareholder in terms of size of the stake, and so the number of shares tradable on the market. Second, when the controlling shareholder extracts private benefits from the ownership share, his trading decisions are less affected by the fundamental value of the firm, and so they are less informative about the true value of the firm. Therefore, the presence of private benefits makes it more noisy for the rest of the investors to extract information from the large shareholder's trade.

To estimate the model, I use data on Italian public companies for which large shareholders are required to disclose their stakes every six months, thus allowing for data with higher frequency with respect to previous studies on the ownership dynamics (e.g., Donelli et al. (2013)).[1] The data show that large shareholders trade infrequently, despite substantial variation in the economic fundamentals of the firm, stock prices, and trading volumes. Moreover, when they trade, large shareholders buy or sell large blocks of shares, while they trade small stakes very rarely.

The main estimation results are the following. First, I estimate private benefits of control of around 2% of equity value. Moreover, I show that the distribution of private benefits is highly positively skewed: 50% of the controlling shareholders extract private benefits of less than 1% of equity value, and the maximum rate is 20%.

Second, I show that private benefits of control generally have positive impact on stock prices. As for the private benefits, the price impact is very heterogenous across firms. For 15% of firms, in fact, private benefits have a negative impact greater than 1%, and for the same proportion of firms private benefits have a positive impact larger than 5%. Moreover, I find evidence of a synergistic effect. When the controlling shareholder is a corporation, the positive price impact

[1] In comparison, US disclosure rules allow to have information only on purchases of blocks above 5% of the outstanding shares (file 13D or 13G), and on the portfolio of big institutional investors with more than 100 millions of dollars of equity assets under management (file 13F).

of the large shareholder's stake is even larger compared to the case of individual controlling shareholders.

Third, I document that the presence of controlling shareholders has, overall, a substantial positive impact on stock prices. This positive price impact is larger during the European sovereign debt crisis of 2011-2012, so that controlling shareholders may be particularly beneficial to the rest of the investors during negative economic cycles.

Finally, I estimate the certainty equivalent payoff of the large shareholders' stake over time, which is unobservable when the large shareholder does not trade. The certainty equivalent payoff is the valuation of the share by an investor, and so the maximum price the investor would be willing to pay to buy the share.

My paper makes contributions to both theoretical and empirical studies on private benefits of control and controlling shareholders' ownership policy. To the best of my knowledge, this is the first paper to provide a measure of the impact of large shareholders on stock prices over time, only predicted in theory by the dynamic models of Collin-Dufresne and Fos (2015) and DeMarzo and Urosevic (2006), and estimated at the time of the block trade by Albuquerque and Schroth (2010).

Theoretical papers predict that the heterogenous valuation between large and small shareholders should always trigger trading by the large shareholder (Collin-Dufresne and Fos (2015), Hilli et al. (2013), DeMarzo and Urosevic (2006), and Gomes (2000)). The implication of persistent trading by the large shareholder is contradicted by empirical evidence showing that the frequency of trading by large shareholders is much less than expected from the predictions of current models (e.g., Donelli et al. (2013)). The infrequent trading by the large shareholder, then, emerges as implication of my model with extraction of private benefits.

Further, I propose a new approach to quantify private benefits of control. Previous studies focus on the acquisition of the controlling stake to measure private benefits of control (e.g., Barclay and Holderness (1989), Nenova (2003), Nicodano and Sembenelli (2004), Dyck and Zingales (2004), and Albuquerque and Schroth (2010)). However, if private benefits may impact on the ownership policy of the controlling shareholder over time, then the ownership policy of the controlling shareholder may contain crucial information to estimate such benefits. The intuition is that private benefits make the controlling shareholder's trading less sensitive to changes in the economic conditions of the firm, since the incentive to hold the stake may be mostly due to the opportunity to enjoy benefits that are unrelated to the economic fundamentals of the firm. For this reason, I use data on the ownership dynamics of large shareholders to assess the magnitude of private benefits of control and their impact on stock prices.

My structural estimates of private benefits of control are in line with that of Albuquerque and Schroth (2010). Similarly to Albuquerque and Schroth (2010) and Nicodano and Sembenelli (2004), I show that the distribution of private benefits is highly positively skewed. My results show that 50% of the large shareholders (40% in Albuquerque and Schroth (2010)) extract private benefits less than 1% of equity value, and the maximum rate of private benefits is 20% (15%). Moreover, my estimates are generally in line with the evidence on private benefits of control in countries with large minority investors protection, such as anglo-saxon and north European countries. This results is consistent with Dyck and Zingales (2004), who show that private benefits of control in Italy dropped dramatically after the passage of a corporate governance reform (known as Draghi reform) that substantially increased minority shareholders protection.

By estimating a structural model, with endogenous asset pricing, I can quantify the price impact of the private benefits of control, and of the presence of a controlling shareholder. Measuring the price impact requires a counterfactual analysis: what should be the stock price in the absence of private benefits? And, what should be the stock price in the absence of a controlling shareholder (i.e., completely dispersed ownership among atomistic investors)? I use the model pricing equations to quantify the difference between the actual stock price and the unobservable counterfactual stock price. Albuquerque and Schroth (2015) adopt the same methodology to quantify the price impact of illiquidity in the market for trading control blocks, and Lippi and Schivardi (2014) perform counterfactual analysis to show that private benefits of control have significant and negative impact on firm's profitability. Moreover, by using the optimality conditions of the controlling shareholder, I quantify the valuation of the stake by the controlling shareholder, thus addressing the question whether private benefits motivate controlling shareholders to hold large stakes even at the cost of excessive risk exposure.

Even though structural estimation relies on a specific theoretical model, I show that the estimated model performs well in replicating empirical facts on the dynamics of large shareholders' stakes that have not been explicitly targeted.

I review the relevant literature in the next section. In Section 1.3, I describe the theoretical framework and characterize the model equilibrium. Section 1.4 describes the estimation methodology, followed by the estimation results in Section 1.5. Section 1.6 concludes the chapter.

1.2. Literature Review

The theoretical framework of the paper builds on the literature that rationalises the ownership policy of a large shareholder (Collin-Dufresne and Fos (2015), Hilli et al. (2013), DeMarzo and Urosevic (2006), and Gomes (2000)). In Collin-Dufresne and Fos (2015), an activist shareholder accumulates shares and improves the firm's value as long as her stake is not fully revealed to the market. The large shareholder of DeMarzo and Urosevic (2006) is prevented to trade to her optimal portfolio allocation from moral hazard, since marginal investors revise the stock price on the base of the large shareholder's stake. As a result, the large shareholder adjusts gradually the stake towards the optimal risk-sharing allocation. Gomes (2000) derives similar trading pattern for the owner-manager of a firm in the presence of asymmetric information, when the small investors do not perfectly observe the managerial ability of the large shareholder. Hilli et al. (2013) show that, even with separation between management and control, the large shareholder trades gradually to her optimal portfolio allocation if there is divergence of interests between manager and owner. The common point among these papers is that the friction triggers persistent trading by the large shareholder, while in the absence of the friction the large shareholder trades immediately to the optimal allocation.

My model assumes an exogenous ownership structure, with one large shareholder only and a mass of small dispersed investors, as in Collin-Dufresne and Fos (2015), Hilli et al. (2013), DeMarzo and Urosevic (2006), and Gomes (2000). A few papers endogenise the ownership structure with different motivations. In the seminal work of Bolton and von Thadden (1998), inside investors trade-off the need of liquidity provided by potential external investors with the cost of being monitored by additional shareholders. In Zwiebel (1995), wealth constraints motivate shareholders to form a controlling coalition to undertake and finance investments, at the cost of sharing fixed private benefits of control. Several blockholders arise in the framework of Dhillon and Rossetto (2014) to mitigate the conflict of interests between the largest shareholder and the marginal investors. Yet, Dhillon and Rossetto (2014) show that one only big shareholder is optimal when such conflict is mild.

Many papers have documented ownership concentration across countries. Faccio and Lang (2002) show that families play a prominent role as controlling shareholder in Western European companies, and La Porta et al. (1999) report similar pattern across worldwide companies. Both papers find an inverse relationship between legal investors' protection and ownership concentration, consistently with the prediction of Dhillon and Rossetto (2014). In particular, Laeven and Levine (2008) show that half of their sample firms have one large shareholder only and residual ownership widely held across dispersed shareholders.

Albeit quantitative measures of the controlling shareholder's impact on stock prices over time are still missing, several studies have highlighted the negative impact of the controlling shareholder on the risk taking of the company, in line with the theoretical results of Admati et al. (1994). Rossetto and Staglian (2018) find supporting evidence of this theoretical prediction only when the controlling shareholder is the only large shareholder of the company. John et al. (2008) argue that controlling shareholders undertake sub-optimally conservative investment decisions to preserve their private benefits of control, while Faccio et al. (2011) motivate the inverse relationship between control and risk with the large risk exposure of the controlling shareholder due to the large stake held in the firm.

Recently, Roger and Schatt (2016) have rationalised private benefits of control as compensation to controlling shareholders for excessive risk exposure in the controlled firm. Roger and Schatt (2016) show that the opportunity to enjoy private benefits motivate shareholders to hold the controlling stake, otherwise inefficiently large. Dyck and Zingales (2004) point out that control confers also costs in terms of low diversification of the controlling investor. However, Odegaard (2016) find little evidence of this argument. With a unique dataset of Norwegian equity-holders, Odegaard (2016) concludes that the magnitude of private benefits enjoyed by large investors does not compensate properly their diversification loss.

Empirical literature has followed two prominent approaches to estimate private benefits of control. The first approach is based on the price difference between shares with voting rights and shares with cash-flows rights. Zingales (1995) and Nenova (2003), among others, claim that the price premium of superior vote shares is justified by the perspective of future benefits of control. The price premium, however, can only be observed in firms that issue dual-class shares (i.e., voting rights and cash-flows rights). For this reason, Benos and Weisbach (2004) argue that this approach overstates private benefits of control, since firms with dual-class shares are likely to confer much larger amount of private control benefits.

The second approach is based on the difference between the price of the controlling stake and the market price of the shares at the day of the block negotiation. This difference is defined as block premium. Barclay and Holderness (1989), Nicodano and Sembenelli (2004), and Dyck and Zingales (2004) use the block premium as an empirical proxy to quantify the private benefits of control. First, they show that controlling shareholders are willing to pay more than the market price to buy the controlling stake of a company. Then, they argue that the premium is justified by the opportunity for the controlling shareholder to extract private benefits from the controlling stake, and try to disentangle from the premium the amount of (expected) private control benefits embedded in the price difference. Albuquerque and Schroth (2010) adopt a structural approach, by using the block pricing model of Burkart et al. (2000), to identify the private benefits of control transferred in the trade of the controlling block. Structural estimation is motivated by the great challenge to disentangle in the block premium the private benefits from the change in the share value due to the takeover of the incoming controlling shareholder.

A few papers have quantified private benefits of control in Italy. The first attempt of Zingales (1995), by using the voting premium approach, delivers very large estimates for Italian listed companies, on average around 80% of the equity value. The results of Nenova (2003) yield much lower numbers, claiming that Zingales (1995) computes private benefits for each single share in the block instead of taking the block as a whole. Nicodano and Sembenelli (2004), using a different approach which adjusts the methodology of Barclay and Holderness (1989) for the degree of dispersion of the ownership structure, provide evidence in line with Nenova (2003). Zingales (1995), Nenova (2003), and Nicodano and Sembenelli (2004), estimate private benefits of control before the Draghi reform of 1998, that increased the minority shareholders protection. Dyck and Zingales (2004) show that private benefits of control in Italy dropped on average from 47% to 6% after the Draghi reform.

1.3. The Model

In this section, I describe the model. First, I state the model assumptions, then I derive the optimality conditions for the marginal investors and the large shareholder. Moreover, I derive the equilibrium stock price and large shareholder's stake. Finally, I summarize the main results.

1.3.1. The Setup

The economy consists of investors with two investment opportunities: the shares of a company and a riskless asset. The following assumptions describe this economy.

- **Assumption 1. Investment Opportunities**

  The firm is in unit supply and generates cumulative free cash flows described by the following diffusion

$$dD_t = \mu_t dt + \sigma_D dZ_t, \tag{1}$$

$$d\mu_t = \sigma dX_t, \tag{2}$$

  where μ_t is a time-varying drift, σ and σ_D are constant, and dZ_t and dX_t are two independent standard Brownian motions. The firm pays out all cash flows as dividends. The riskless investment pays a continuously compounded rate of return r, with perfect elastic supply. Without loss of generality, I assume that all cash flows are paid to all shareholders each period, in forms of dividends, share repurchases or issuance. The assumption of normality on the cash flows implies that they can be negative. Moreover, they are independent across periods. Both independence and normality ensure tractability of the model.

  I use the assumption of a time-varying drift to introduce asymmetric information between marginal investors and large shareholder, and the sequential learning of the marginal investors in bayesian fashion, using the observed dividends flow as noisy signal on μ_t (see Assumption 3).

- **Assumption 2. Investors Population**

  The economy is populated by a continuum of competitive investors, with measure M. All the investors are risk-averse, with standard CARA utility function, defined on the continuous flow of consumption. The agents have equal risk and intertemporal preferences. The investors live infinitely, and trade continuously the riskless asset and the company shares, with price P to be determined in equilibrium. Let $\alpha_{i,t}$ denote the number of shares owned at time t by the investor i, with i going from 1 to M. Each investor, then, chooses the amount of consumption and the number of shares to maximise

$$E_t \int_t^{\infty} e^{-R(s-t)} u(c_s) ds,$$

where $u(c) = -e^{-ac}$, a is the absolute risk-aversion coefficient, and R is the rate of intertemporal preferences, that is equal to r when $a = 0$. The wealth of each investor i is given by the riskless asset and the company shares, that is $W_i = B_i + \alpha_i P$. I assume that small shareholders cannot form coalitions and do not behave strategically.

There exists a large shareholder, who is risk-averse, with equal risk and intertemporal preferences as the marginal investors, lives infinitely, and has identical objective function as the marginal investors. However, the large shareholder differs from the rest of the investors in two directions. First, the large shareholder sets the optimal number of shares at discrete dates τ.[2] Moreover, the large shareholder extracts additional benefits from investing in the firm, that accrue to the total wealth of the shareholder given by $W_L = B_L + \alpha_L P + \Phi(\alpha_L)$. The private benefits from shareholding, $\Phi(\alpha_L)$, generate continuously an instantaneous inflow of additional wealth for the large shareholder, denoted by $\phi(\alpha_L)$, such that $dW_L = d(B_L + \alpha_L P) + \phi(\alpha_L)$, according to a discrete step function:

$$\phi(\alpha_L) = b * \alpha_j,$$

if $\alpha_j \leq \alpha_L < \alpha_{j+1}$, where α_j and α_{j+1} are given thresholds of stakes (for instance, 0.2 and 0.3, respectively), with $j = \{0, 1, 2, ..., J-1\}$, $\alpha_0 = 0$, and $\alpha_J = 1$.

I make this assumption for consistency with the actual institutional framework, in which the shareholders' rights and obligations are triggered as soon as the shareholders come into possession of a given percentage of the outstanding shares. It follows that the shareholders acquire additional rights (or have the duty to comply with additional obligations) only if they reach the next higher threshold, and they lose rights (or are free from complying with a given obligation) as soon as they hold one stock less than a given percentage of

[2] This assumption improves the tractability of the model and follows DeMarzo and Urosevic (2006).

shares. The private benefits take here the form of monetary incomes, or non-pecuniarity amenities that can be converted in additional wealth. b may take any real value, either positive or negative. Negative values imply that the large shareholder is bearing private costs from holding the stake in the company, in forms of outflow of money deducted from his total wealth.

- **Assumption 3. Heterogeneous Information and Beliefs Update**

  The time-varying drift of the dividends pay out, denoted by μ_t, is unobservable to the marginal investors and observable only to the large shareholder. The marginal investors have heterogenous prior on μ_t, conditioning on the information set at time t, denoted by $\mu_{i,t}$, and they have heterogenous prior variances, that are different conjectures on σ^2, denoted by σ_i^2. The marginal investors receive two types of signal to update their prior on μ_t, in a bayesian fashion.

  At each time t, the investors observe the continuous cash flow dD_t, that is a noisy signal on μ_t. The signal is noisy as the investors are not able to disentangle between the pure shock on the dividend payout, given by σdZ_t, and the true dividend drift μ_t. Therefore, each investor updates continuously her prior on μ_t according to the conjecture on σ. In particular, the larger is the investor's prior variance σ_i^2, the lower is the level of confidence of the investor about her prior, and the larger is the weight assigned to the noisy signal for revising the prior on μ_t. The investor's prior on μ_t, then, evolves according to the following equation

$$d\mu_{i,t} = k_i \eta_{i,t},$$

  where, following standard bayesian filtering results,

$$k_i = \frac{\sigma_i^2}{\sigma_D^2 + \sigma_i^2}, \eta_{i,t} = dD_t - E_{i,t}[dD_t],$$

  and $E_{i,t}[dD_t] = \mu_{i,t}$.

Therefore, the heterogeneity across investors is fully described by the distribution of the coefficient k_i, that is the weight assigned to the signal for updating the beliefs on the expected dividend payout. [3]

At the discrete dates τ, the investors observe the ownership share of the large shareholder $\alpha_{L,\tau}$, that is a noisy signal on μ_t, since investors know that the large shareholder has additional motives to invest in the firm, given by the private benefits, that are unobservable to the marginal investors. The update at τ of each marginal investor about μ_t is then described by the following equation:

$$\mu_{i,\tau} = \mu_{i,t<\tau} + g_i(\tau^-)(\alpha_{L,\tau} - \alpha_{L,\tau}(\mu_{i,t<\tau})),$$

where $\mu_{i,t<\tau}$ is the prior of the investor i before observing the choice of the large shareholder, $\alpha_{L,\tau}(\mu_{i,t<\tau})$ is the belief of the investor i about the optimal choice of the large shareholder, and $g_i(\tau^-)$ is the weight assigned to the observation of the large shareholder's choice for updating the prior on μ_t. Therefore, each investor i revises upwards (downwards) the conjecture on the true dividends drift of the firm when the large shareholder sets an ownership share above (below) the expected choice, that is interpreted as a positive (negative) signal on the true state of the firm. Following again standard bayesian filtering results, the weight $g_i(\tau^-)$ is equal to

$$g_i(\tau^-) = \frac{\alpha^{\mu}_{L,\tau^-}\sigma_i^2}{(\alpha^{\mu}_{L,\tau^-})^2\sigma_i^2 + \sigma_\epsilon^2},$$

where $\alpha^{\mu}_{L,\tau} = \frac{\partial \alpha_{L,\tau}}{\partial \mu_t}$ is the derivative of the optimal choice of the large shareholder with respect to the true dividends drift, and σ_ϵ^2 is the variance of the observation error. So, σ_ϵ^2 is a measure of the noise contained in the information on μ_t released by the large shareholder with his optimal choice.

[3] The reader can refer to a set of investors with different level of confidence or information on the future cash flows generated by a company. An investor with superior level of information, or high degree of confidence, has a low value of k, and relies little on the signals provided by the actual dividend payments. Instead, a poorly informed investor is easily conditioned by the fresher information provided by the new dividend payment, and has a high value of k.

In case of no private benefits, the marginal investors know that the large shareholder sets the demand for shares only on the base of the superior information on the dividends drift. Therefore, the observation α_{L,τ^-} is clean, that is $\sigma_\epsilon^2 = 0$, so $g_i(\tau^-) = \frac{1}{(\alpha^\mu_{L,\tau^-})}$, where α^μ_{L,τ^-} is observable, albeit with one period lag due to asymmetric information. With private benefits, the observation α_{L,τ^-} is not longer clean, that is $\sigma_\epsilon^2 > 0$.

Given assumptions 1, 2, and 3, the following sections characterize the model solution. First, I summarize the main results in figure 1, with a graphical simulation. Details on the simulation study are provided in the Appendix. Then, I describe analytically the investors' optimality and the equilibrium stock price, the large shareholder's optimality, finally deriving the model equilibrium. I leave the proofs for the Appendix to save in space and notation. Next, I depict the timing of the model.

Model Timing

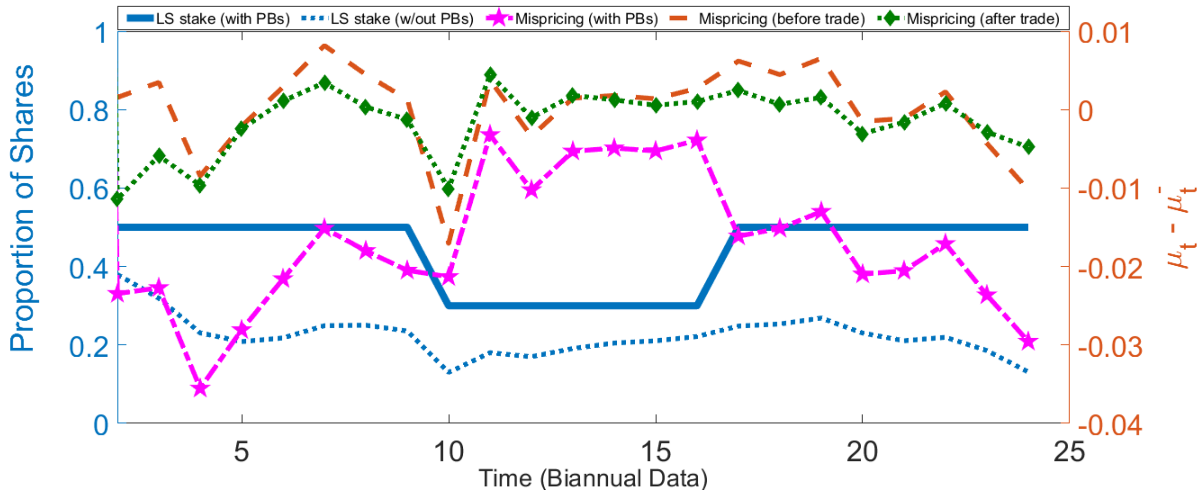
-∞ $\rightarrow$ .....

- Large shareholder chooses the stake α_{τ^-}
- Small investors update their beliefs based on α_{τ^-}
- Cash flow dD_{τ^-} is realized
- Small investors update their beliefs based on $dD_{\tau^-} \Rightarrow \{\bar{\mu}_{\tau^-}; P_{\tau^-}\}$
- ...
- Small investors update their beliefs based on $dD_{t<\tau} \Rightarrow \{\bar{\mu}_{t<\tau}; P_{t<\tau}\}$
- Large shareholder chooses the stake α_τ

$\rightarrow$ ..... $+\infty$

In Figure 1, the dotted line shows that in the absence of private benefits the large shareholder always trades, to exploit the mispricing of the marginal investors (dashed line). The mispricing is given by the difference between the true dividends drift and the average belief

Figure 1. Ownership Dynamics, Private Benefits and Mispricing



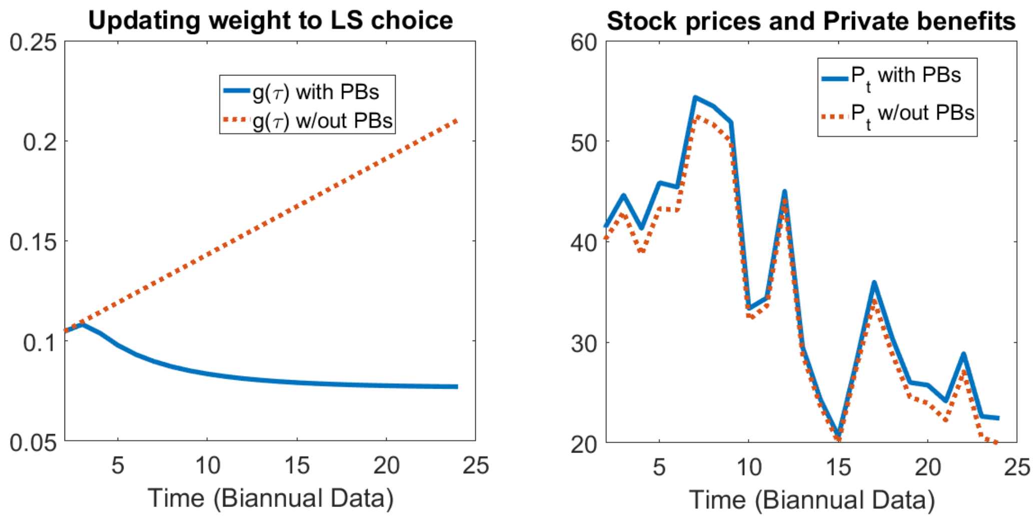
The figure shows the optimal demand for shares of the large shareholder in presence of private benefits (blue line), and in absence of private benefits (dotted line), against the difference between the true dividends drift (μ_t) and the average belief on the dividends drift ($\bar{\mu}_t$) by the marginal investors before the disclosure of the large shareholder's stake (dashed line), after the disclosure of the large shareholder's stake in absence of private benefits (diamond line), and after the disclosure of the large shareholder's stake in presence of private benefits (stars line). The parameters used for the numerical example are the same as for the simulation study described in the Appendix.

on the dividends drift by the marginal investors. However, the marginal investors learn from the large shareholder's trade revise their belief after observing the large shareholder's trade (diamond line). As a result, the large shareholder trades gradually to the optimal risk-sharing solution.

With private benefits, the large shareholder trades only at two points in time (blue line). The large shareholder sells (buys) a block of shares when the overvaluation (undervaluation) by the marginal investors makes the sale (purchase) convenient to the large shareholder: the trading gains (costs) offset both the loss (gain) in private benefits and the share depreciation (appreciation), due to the negative (positive) price impact of his trade.

Yet, in the presence of private benefits, the signal released by the large shareholder with his demand for shares is noisy. The noise in the information released by the large shareholder's demand generates a distortion in the update of the marginal investors' belief (stars line), and the marginal investors consider less reliable this information in the update of their belief compared to the case of no private benefits. The left panel of Figure 2 shows that the weight assigned by the marginal investors to the information released by the large shareholder's demand ($g(\tau)$) in the presence of private benefits is gradually lower compared to the case of no private benefits.

Figure 2. The impact of LS ownership policy



The left panel shows the average updating weight assigned to the observation of the large shareholder's stake by the marginal investors to update their belief on the dividends drift, with (blue line) and without (red dotted line) private benefits. The right panel shows the equilibrium stock price at the disclosure dates τ with (blue line) and without (red dotted line) private benefits. The parameters used for the numerical example are the same as for the simulation study described in the Appendix.

1.3.2. Equilibrium Share Price and Investor's Optimality

I start characterizing the model solution describing the marginal investors' optimality conditions.

At each time t, the optimal choice in terms of number of shares of the investor i is given by:

$$\alpha_{i,t} = \frac{\mu_{i,t} - \bar{\mu}_t + \rho}{ar\sigma^2}, \tag{3}$$

where $\bar{\mu}_t$ is the average expected dividend by the marginal investors, and the risk premium ρ, determined by market clearing condition, that is $\int_i \alpha_{i,t} di = 1 - \alpha_{L,t}$ for each t, is

$$\rho_t = (1 - \alpha_{L,t}) a^I r \sigma^2,$$

where $\alpha_{L,t}$ is the stake held by the large shareholder at time t, and a^I is the aggregate risk aversion coefficient:

$$\frac{1}{a^I} = \int_i \frac{1}{a^i} di.$$

Proof. Appendix 1.7.1

Equation (3) has a natural interpretation. While the demand for shares decreases with the dividends process variance and the risk aversion coefficient, the demand increases with the difference between the individual and the average belief about the expected dividend payout of the firm. Equation (3) reminds the familiar optimal risky asset allocation for a mean-variance investor, and it is equivalent to the optimal solution of the small price-taker investor of DeMarzo and Urosevic (2006). However, in DeMarzo and Urosevic (2006) the risk premium only depends on the large shareholder's trading, as the expected dividend payout is observable.

Next, I derive the equilibrium share price. At each time t, the stock price is the following:

$$P_t = \int_t^{\infty} e^{-r(s-t)}(\bar{\mu}_s - \rho_s)ds. \tag{4}$$

Proof. Appendix 1.7.2

Hence, the share price is the present value of the expected dividends by the marginal investors, minus the risk premium component. The former evolves continuously according to the following equation:

$$d\bar{\mu}_t = \bar{k}_t(dD_t - \bar{\mu}_t). \tag{5}$$

Proof. Appendix 1.7.3

Equation (5) shows that the average belief increases (decreases) when the actual dividend payout is greater (lower) than the expected dividend payout by the marginal investors, and the rate of growth is proportional to the average reaction of the marginal investors to the new signal. Then, share prices fluctuate also independently from the economic fundamentals of the company, due to over or under reaction of the investors to news and shocks. Chan (2003) shows that investors react slowly to valid information, while they overreact to price shocks, causing huge trading volume and price volatility.

1.3.3. Large Shareholder's Optimality

Given the average prior on the expected dividend payout by the marginal investors $\bar{\mu}_t$, reflected in the share price $P_t = P(\bar{\mu}_t)$, the large shareholder chooses at discrete dates τ the optimal number of shares, $\alpha_{L,\tau}$, to maximize the certainty equivalent payoff

$$V(\alpha_{L,\tau}) - (\alpha_{L,\tau} - \alpha_{L,\tau^-})P_\tau,$$

where

$$V(\alpha_{L,\tau}) = \int_\tau^\infty e^{-r(s-\tau)} v(\alpha_{L,s}) ds,$$

and

$$v(\alpha_{L,t}) = \alpha_{L,t}\mu_t - \frac{1}{2}\alpha_{L,t}^2 a^L \sigma_D^2 r + \phi(b, \alpha_{L,t}).$$

$v(\alpha_{L,t})$ is the net benefits flow to the large shareholder at each time t, given by the risk-adjusted instantaneous dividend process accrued to the large shareholder, plus the additional inflow of instantaneous private benefits generated by the stake.

So, the certainty equivalent payoff is given by the present value of the net benefits flow less (plus) the trading costs (gains), where α_{L,t^-} stands for the number of shares at the previous point in time. By taking the first-order derivative with respect to $\alpha_{L,t}$, I obtain the large shareholder's optimality condition:

$$V' = P_\tau + (\alpha_{L,\tau} - \alpha_{L,\tau^-})P', \tag{6}$$

where the prime index stands for the derivative with respect to the control variable. Equation (6) is the usual equilibrium condition that equates benefits and costs, where the marginal benefits are given by the risk-adjusted cumulative expected dividends plus the private benefits generated by an additional share, and the marginal costs are given by the share price plus the implicit cost of trading, due to the price impact of the large shareholder's stake.[4]

1.3.4. Equilibrium

I now characterize the equilibrium stock price and the optimal demand for shares of the large shareholder, taking into account his impact on the stock price. The equilibrium risk premium is derived by the market clearing condition in the static context (sum of investors' demand for shares equal to total number of shares minus the large shareholder's stake), while the average expected dividend payout is derived by the market clearing condition in the dynamic context

[4]Given the step function characterizing the private benefits of the large shareholder, the objective function displays local maxima at the thresholds, thus implying that the large shareholder in equilibrium holds a stake equal to either one of the thresholds.

(sum of shares purchased equal to sum of shares sold).

Proposition 1. At each discrete date τ, the equilibrium share price is given by

$$P_\tau = \int_\tau^\infty e^{-r(s-\tau)}(\bar{\mu_\tau} - \rho_\tau)ds,$$

where the equilibrium risk premium is $\rho_\tau = (1 - \alpha_{L,\tau}) * a^I \sigma_D^2 r$, and

$$\bar{\mu}_\tau = \bar{\mu}_{t<\tau} + \bar{g}(\tau^-)(\alpha_{L,\tau} - \alpha_{L,\tau}(\bar{\mu}_{t<\tau})).$$

Proof. Appendix 1.7.4

$\bar{\mu}_\tau$ is the average posterior belief on the dividend payout of the firm by the marginal investors, after the large shareholder's choice of ownership share, where $\bar{g}(\tau^-)$ is the average reaction by the marginal investors to the large shareholder's choice of ownership share, and $\alpha_{L,\tau}(\bar{\mu}_{t<\tau})$ is the average prior by the marginal investors about the optimal choice of the large shareholder:

$$\alpha_{L,\tau}(\bar{\mu}_{t<\tau}) = \frac{(1 + \alpha_{L,\tau^-})a^I \sigma_D^2 r}{2a^I \sigma_D^2 r + a^L \sigma_D^2 r},$$

where $\bar{\mu}_{t<\tau} = \bar{\mu}(dDt)$ is the average expected dividend payout by the marginal investors before the large shareholder's choice of ownership share, which is function of the continuous signals given by the company's dividends.

Proposition 2. Taking into account his price impact, the large shareholder's optimal demand for shares is the following

$$\alpha_{L,\tau} = \frac{\mu_\tau - \bar{\mu}_{t<\tau} + \phi'(b, \alpha_{L,\tau})}{2a^I \sigma_D^2 r + a^L \sigma_D^2 r + \bar{g}(\tau^-)} + \alpha_{L,\tau}(\bar{\mu}_{t<\tau}).$$

Proof. Appendix 1.7.5

The large shareholder's optimal demand for shares is the sum of three components: the speculation on the investors' mispricing, the expected optimal choice of shares conjectured by the marginal investors, and the gain in the private benefits generated by an additional share, which is positive only when the additional share allows to reach a higher threshold. The next section provides closed-form solution in two benchmark cases, and provides further details on the solution in the general case.

1.3.5. Main Results

- **Perfect Information and No Private Benefits**

In case of perfect information and no private benefits, the marginal investors set the weight to assign to the observation of the large shareholder's choice $\alpha_{L,\tau}$ at time τ, then $\bar{g}(\tau)$ is simply equal to $1/(2a^I\sigma_D^2 r + a^L\sigma_D^2 r)$, so that $\bar{\mu}_t = \mu_t$, since

$$\alpha_{L,\tau} - \alpha_{L,\tau}(\bar{\mu}_{t<\tau}) = \frac{\mu_\tau - \bar{\mu}_{t<\tau}}{2a^I\sigma_D^2 r + a^L\sigma_D^2 r}.$$

When all the investors have the same set of information, at each point in time, then the equilibrium stock price is simply given by the present value of the expected dividends payout of the firm, given the true dividends drift μ_t, minus the equilibrium risk premium $\rho = (1 - \alpha_{L,t})a^I r\sigma^2$, where

$$\alpha_{L,t} = \frac{a^I}{a^L + a^I},$$

that is the large shareholder *immediately* trades to the risk-sharing allocation. This result is equivalent to DeMarzo and Urosevic (2006) in the absence of moral hazard.

- **Asymmetric Information and No Private Benefits**

With asymmetric information and no private benefits, the large shareholder's demand for shares is informative on the true drift of the dividends process. However, the marginal investors set the weight to assign to the observation of the large shareholder's choice $\alpha_{L,\tau}$ at time τ^-, so that

$$\bar{g}(\tau) = \frac{1}{\alpha^\mu_{L,\tau}} = \frac{1}{2a^I\sigma_D^2 r + a^L\sigma_D^2 r + \bar{g}(\tau^-)}.$$

In turn, at τ, the large shareholder knows and anticipates the response of the investors, thus taking into account the impact of his demand for shares. Given his superior information, the large shareholder can exploit the mispricing of the marginal investors, thus trading on the difference between his stock valuation and the average belief by the marginal investors. Therefore, the large shareholder has always the temptation to trade.

However, as the marginal investors learn from the large shareholder's trade, the stock price increases when the large shareholder buys, and falls when the large shareholder sells. This mechanism makes convex (concave) the trading costs (profits) on the purchase (sale) of the shares to the large shareholder, and prevents the large shareholder from trading to his first-best solution, that is the trading policy in the absence of the investors' learning.

As a result, the large shareholder trades *gradually*, yet not monotonically, at the discrete dates τ, towards the optimal risk-sharing allocation, while the marginal investors update their belief on the expected dividends payout. Indeed, the magnitude of the large shareholder's trade, and so the speed of the adjustment towards the optimal risk-sharing solution, is inversely proportional to the average reaction of the marginal investors ($g(\tau^-)$) to his trade. This result is equivalent to DeMarzo and Urosevic (2006) in the presence of moral hazard.

- **Asymmetric Information and Private Benefits**

  Finally, with private benefits, the information released by the large shareholder with his optimal demand for shares is noisy with respect to the true drift of the dividends process, then

$$\bar{g}(\tau) = \frac{\alpha^{\mu}_{L,\tau}\sigma_i^2}{(\alpha^{\mu}_{L,\tau})^2\sigma_i^2 + \sigma_\epsilon^2},$$

  where the actual α^{μ}_{L,τ^-} is not observable, and therefore proxied by $1/(2a^I\sigma_D^2 r + a^L\sigma_D^2 r)$.

  While in the absence of private benefits the large shareholder always trades at the discrete dates τ, in the presence of private benefits the large shareholder may not trade at all. The no-trade occurs when the large shareholder's stake is at a given threshold ($\alpha_{L,\tau} = \alpha_j$). The purchase of a share would not produce any additional private benefit while making even more undiversified and suboptimal the investment portfolio of the large shareholder. On the other side, selling a share produces a loss in private benefits ($\phi(\alpha_{L,\tau} < \alpha_j) = b * \alpha_{j-1}$) which may not be compensated by the gain in risk diversification.

  The large shareholder, instead, sells a block of shares when the difference in the valuation, against the marginal investors, is negative (the share is overvalued) and large enough to make convenient the trade, even if this happens at the cost of losing a given amount of private benefits, and the shares depreciate because of the sale, given the price impact of his sale. On the other hand, the large shareholder is willing to buy an additional block of shares, in order to reach the higher threshold that generates additional private benefits, only when the difference in the valuation becomes positive (the share is undervalued) and large enough, taking into account the implicit cost of trading due to the share appreciation.

1.3.6. Price Impact of Private Benefits

Private benefits affect the equilibrium stock price in two directions. As shown above, the equilibrium stock price depends on both the average expected dividend by the marginal investors following the large shareholder's demand for shares, and the equilibrium risk premium ρ derived by the market clearing condition

$$P_\tau = P\left(\bar{\mu}(\alpha_{L,\tau}), \rho(\alpha_{L,\tau})\right).$$

Then, it is straightforward to note that private benefits increase the equilibrium stock price by reducing the equilibrium risk premium. In fact, private benefits increase the optimal demand for shares of the large shareholder, thus reducing the equilibrium risk premium sought by the marginal shareholders to invest in the firm. In practice, a larger demand for shares of the large shareholder reduces the number of shares tradable on the market, under the assumption of fixed shares supply, thus raising the stock price.

Moreover, private benefits affect the update, at τ, of the belief of the marginal investors on the true expected dividends payout, following the observation of the large shareholder's decision in terms of ownership share. First, since private benefits drive the optimal demand for shares of the large shareholder, then private benefits impact on the distance between the actual and the expected choice, by the marginal investors, of the large shareholder's ownership share:

$$(\alpha_{L,\tau} - \alpha_{L,\tau}(\bar{\mu}_{t<\tau})).$$

Further, private benefits affect the weight assigned by the marginal investors to the observation of the large shareholder's choice. Since this choice is driven by additional motives with respect to the dividend payout of the company, then the choice is a noisy signal on the dividend payout of the company, and so the marginal investors assign a lower weight to that observation in the update of their belief.

1.4. Structural Estimation

In this section, I report the actual data used for the estimation and I state the estimation problem, by describing the quantities to estimate, the parameters that are calibrated, and the observable variables involved in the estimation. Then, I describe the identification process, which links the observable variables to the unobservable quantities of the model.

1.4.1. Data

The universe of firms consists of the public companies listed in the Italian Stock Exchange, in which they are classified by market capitalization. I consider all the non-financial firms listed in the Large, Medium, and Small Capitalization indexes. The source of data for the ownership share of the large shareholders is Thomson Reuters Eikon, which combines public and private information on the ownership structure of public companies. However, I double-check manually the data by using the website of the Consob, the Italian security exchange commission, that releases information on the ownership structure of the public companies every six months, on the base of the company disclosure. So, I use biannual data on the largest shareholder's stake between March 2004 and September 2016 (24 observations). Thomson Reuters Datastream, on the other hand, provides also data on stock prices and earnings per share.

My final sample is obtained by selecting only the firms reporting the same largest shareholder for at least 75% of the observations. This filter allows to identify correctly the controlling shareholder of the firm. Further, I delete the firms with missing data over the time series. The final sample consists of 78 firms, 936 firm-year data on the earnings-per-share, 1,872 firm-biannual data on shareholdings and stock prices, and 280,800 firm-day observations on daily stock prices.

The large shareholder's average and median stake is around 50%, and it is quite stable over time (Figure 3, left panel). In fact, the trading activity of the large shareholder is very low. I observe a trade in only 18% of the total observations, where I refer to trade as a non-zero difference between two consecutive stake observations, and the average number of trades across

Table 1. Descriptive Statistics

	Variable	Mean	Median	Std	10th pct	90th pct
Company Data	Total Assets	1.24	0.15	3.42	0.01	2.79
	Debt-To-Equity	1.72	0.87	14.77	0.11	2.45
	Earning-Per-Share	0.30	0.14	1.02	-0.25	1.17
	FCF-Per-Share	0.94	0.48	1.49	-0.01	2.39
Stock Data	Daily Returns (%)	1.48	-0.38	20.92	-19.32	22.83
	Daily Turnover	0.38	0.23	0.60	0.02	5.26
LS Data	Stake (%)	48.63	53.29	17.53	18.88	66.96
	Trade (%)	4.27	0.41	1.17	0.00	4.35
	N of Trades	3.33	3	2.34	1	6

The table reports the descriptive statistics at company, stock, and largest shareholder levels. The statistics are the mean, the median, the standard deviation, the 10th and the 90th percentiles, computed over the 78 final sample, between March 2004 and Septmber 2016. Company data are on annual basis, and are the total value of assets (in millions of euro), the debt-to-equity ratio, the earnings-per-share, and the free cash flow-per-share. Stock data are on daily basis, and are the returns and the number of shares traded divided by the number of outstanding shares. Largest shareholder data are on biannual basis, and are the percentage of shares held by the largest shareholder, the variation over time of the largest shareholder's stake, and the number of times the stake of the largest shareholder changes.

firms is slightly above 3 (out of 24 observations for each firm). The right panel of Figure 3 shows the distribution of the number of trades across the largest shareholders in the sample.

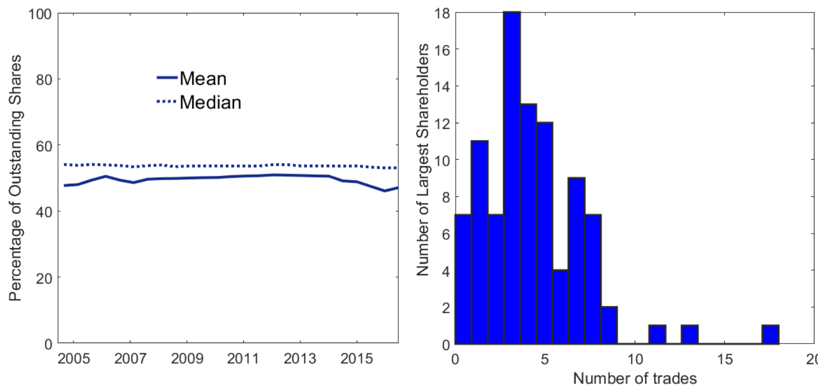
Moreover, when they do, large shareholders usually trade big blocks of shares: the mean (median) trade is 4.27% (1.17%) of the outstanding shares. Finally, they rarely trade small blocks of shares. I observe a trade that involves less than 1% of the outstanding shares in only 7.79% of total observations.

In summary, large shareholders show an ownership dynamics quite stable over time, and they trade very infrequently, usually buying or selling large blocks of shares. By contrast, the sample firms are characterised by a large trading volume over time, and stock prices fluctuate significantly, thus showing a substantial trading activity of the mass of shareholders and investors operating in the market. On average, 0.38% of the outstanding shares are traded every day.

1.4.2. Parameters Calibration

The steps of the private benefits function (α_j), and the aggregate risk aversion coefficient (a^I), are calibrated to the Italian data, for consistency with the actual dataset used for the estimation. I set a^I equal to 2, following Guiso et al. (2018) who measure the aggregate risk

Figure 3. Largest Shareholders: Stake and Trading



The left panel shows the mean (blue line) and the median (dotted line) ownership share, as percentage of outstanding shares, across the largest shareholders of the final 78 sample firms, between March 2004 and September 2016, at biannual frequency. The right panel shows the distribution of the number of trades of the largest shareholders of the final 78 sample firms, where a trade is defined as a non-zero difference between two consecutive stake observations.

Table 2. Ownership Share Thresholds

Ownership Share	Right/Commitment
3%	Obligation to stake disclosure
10%	Right to call shareholders meeting
30%	Obligation to launch takeover
50%	Company control
66%	Right to call extraordinary meeting
100%	Full ownership

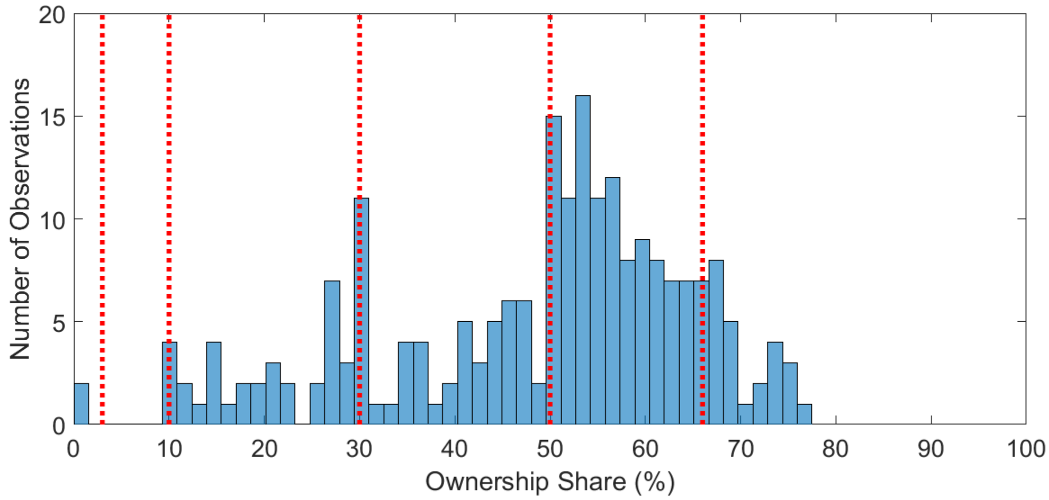
The table describes the ownership share thresholds that a shareholder has to attain in order to acquire a given right, or that triggers a given commitment, according to the Italian commercial law. The thresholds are expressed as percentage of the outstanding shares.

aversion on a large set of clients of an Italian bank. The private benefits thresholds follow the Italian law on the ownership structure of public listed companies. The thresholds are listed in table II.

The rights (and the obligations) linked to each stake level are triggered *as soon as* the shareholder reaches a given threshold. This rule motivates the assumption that the stake generates a given amount of private benefits for a given threshold only if the stake is greater or equal than that threshold, while holding a stake even one share lower than a threshold generates private benefits according to the lower threshold (if $\alpha_L = 29.99\%$, then $\phi(\alpha_L) = b * (10\%)$).

The following figure shows the empirical distribution of the no-trade thresholds. I define no-trade threshold the percentage of shares at which the largest shareholder of the company does

Figure 4. Empirical No-Trade Thresholds



The figure shows the distribution of the no-trade ownership shares observed on the 78 final sample firms. The no-trade share is defined as the ownership share observed at least for two consecutive observations for a given largest shareholder, that is the ownership share at which the largest shareholder does not trade at least across two periods. The red vertical dotted lines are the stake thresholds described in Table II.

not trade, that is the largest shareholder holds that ownership share for at least two consecutive observations. Note that the largest number of no-trade thresholds are observed around 30%, and between 50% and 70%.

1.4.3. Structural Parameters and Latent Variables

I estimate the model firm-by-firm. For each firm, I estimate the following set of parameters:

$$\theta = \{a^L, \sigma_D, \sigma, b, \bar{g}(0), \sigma_\epsilon^2\},$$

that is the large shareholder's absolute risk aversion coefficient, the volatility of shocks to dividends, the volatility of shocks to dividends drift, the parameter that quantifies the private benefits extracted by the large shareholder, for a given discrete step function, the initial weight assigned to the large shareholder's demand for shares by the marginal investors to update their belief on μ_t, and the noise contained in the information released by the large shareholder's demand.

Moreover, for each firm, I infer the dynamics of the following set of latent variables:

$$X_t = \{\mu_t, \bar{\mu}_t\},$$

that includes the true dividends drift and the average expected dividends drift by the marginal investors.

I estimate the model by using stock prices, ownership share of large shareholders, and earnings-per-share.

1.4.4. First Step

In the first step, I estimate $\{\sigma_D, \sigma\}$ by using daily stock prices, that proxy the continuous evolution of the equilibrium share price in the model. To estimate, I discretize the equations (5) and (2), respectively, take the conditional expectation at time t, and derive the following set of diffusion equations that link the two latent variables:

$$E_t[\bar{\mu}_{t+1}] = (1 - \bar{k}_t)\bar{\mu}_t + \bar{k}\mu_t, \tag{7}$$

where k_t measures the average reaction to the new observation of the dividends payout by the marginal investors, and

$$E_t[\mu_{t+1}] = \mu_t, \tag{8}$$

as $E_t[dD_t] = \mu_t$, that is the true drift of the dividends process, and $E_t[d\mu_t] = 0$. The conditional covariance matrix of the two latent variables is diagonal[5]. The diagonal matrix depends on both σ_D and σ:

$$\Sigma_{t+1|t}(X_t) = Diag(\sigma^2, \bar{k}_t\sigma_D^2).$$

[5]In the next formula, k_t is allowed to vary over time to proxy the evolution of the average reaction of the marginal investors to the new signal provided by the firm's payout:

$$k_t = \frac{\nu_t}{\nu_t + \sigma_D^2},$$

where $\nu_t = w_t + \sigma$, and at each time step w_t is updated by using $(1 - k_{t-1})\nu_{t-1}$, and initializing the recursion with a large value of ν_0. This procedure allows to proxy the prior update on the dividends drift across the marginal investors.

On the other hand, the stock price is linked to the two state variables according to the equation (4), that can be written as follows

$$P_t = \frac{1}{r}[(\sigma_D^2 a^I(1-\alpha_{L,t})) + \bar{\mu}_t]. \tag{9}$$

Therefore, given a prior on μ_t and $\bar{\mu}_t$, I can compute a predicted stock price at each time t, by using the above equation, then obtaining a prediction error

$$e_t = \tilde{P}_t - \hat{P}_t,$$

where $\tilde{P}_t$ is the actual stock price, and $\hat{P}_t$ is the predicted stock price. The errors are function of the structural parameters and the latent variables, i.e. $e_t = e(a^I, \sigma_D^2, X_t)$, and the covariance matrix of the prediction errors depends on the derivative of the stock price with respect to the state variables and the conditional covariance matrix of the state variables:

$$\Sigma(e) = f\left(\frac{\partial P_t}{\partial X_t}, \Sigma_{t+1|t}(X_t)\right),$$

where

$$\frac{\partial P_t}{\partial X_t} = \left[\frac{1}{r}\bar{k}_t; \frac{1}{r}\right].$$

So, I construct a likelihood function on the prediction errors, under the assumption of normality, that I maximize with respect to σ:

$$\hat{\sigma} = argmax_\sigma \ln \ell(e_t;\sigma) = -\frac{1}{2}\sum_{t=0}^{T}\ln|\Sigma(e)| - \frac{1}{2}\sum_{t=0}^{T} e_t{}'\Sigma(e)^{-1}e_t,$$

under the restriction that

$$\hat{\sigma}_D = argmin_{\sigma_D}\left[\sigma_D - \sqrt{\frac{var(\delta(D_t)) - \hat{\sigma}^2}{2}}\right]^2,$$

where $\delta(D_t)$ stands for the innovations in the earnings-per-share, and the above condition is derived from equation (1), noting that

$$var(\delta(dD_t)) = var(\delta(d\mu_t)) + 2var(dD_t - \mu_t) = \sigma^2 + 2\sigma_D^2.$$

The maximization of the likelihood function, combined with the condition on the variance of the innovations in the earning-per-share, allows to simultaneously estimate σ_D and σ, and to infer the dynamics of the two state variables. The second result is achieved by iterating the updating and the predicting equations of the linear Kalman filter.[6] In particular, for each time t, the prior estimate on X_t is updated on the base of the prediction error, thus obtaining a posterior estimate of X_t in a bayesian fashion, that is used as prior estimate for the next point in time.

1.4.5. Second Step

In the second step, I estimate $\{a_L, \bar{g}(0), \sigma_\epsilon^2\}$ by using biannual stock prices and contemporaneous ownership share of large shareholders. By using the model equilibrium conditions, I arrive at one equation that describes the stock price at the discrete dates τ, when the large shareholder discloses his ownership share, as function only of the exogenous variables and the ownership share of the large shareholder, by eliminating all the remaining endogenous quantities determined in the model.

Given $\bar{\mu}_{t<\tau}$ that I have estimated in the previous step with daily stock prices, the evolution of the biannual stock prices is endogenously determined using $\alpha_{L,\tau}$, $\bar{g}(\tau)$, and $\alpha_{L,\tau}(\bar{\mu}_{t<\tau})$, where $\bar{g}(\tau)$ is function of the exogenous parameters and $\bar{g}(\tau^-)$, and $\alpha_{L,\tau}(\bar{\mu}_{t<\tau})$ is function of the exogenous parameters and α_{L,τ^-}.

Hence, using again the equation (9), and similar approach to the previous step, I can compute a predicted stock price at each date τ, where I use $\bar{\mu}_\tau$, as determined in the model, to form my prediction on the stock price, thus obtaining again a prediction error at each date τ. Once more, the errors are function of the structural parameters and the latent variable $\bar{\mu}_\tau$. The conditional variance of the state variable is now simply $\bar{k}_t\sigma_D^2$, while the variance of the prediction errors depends again on the derivative of the stock price with respect to the state variable, now simply given by $1/r$, and the conditional variance of the state variable. Then,

[6] Details on the Kalman filter implementation, and details on the identification of k_t are provided in the appendix.

I construct a likelihood function on the prediction errors, under the assumption of normality, that I maximize with respect to $\{a_L, \bar{g}(0), \sigma_\epsilon^2\}$.

1.4.6. Identifying Private Benefits

Now, I describe the identification process of the private benefits parameter b. In particular, I derive upper and lower bounds for b, for each firm-large shareholder. First, let $J_\tau^L = J(\alpha_{L,\tau}, b)$ denote the maximal certainty equivalent payoff for the large shareholder at each time τ, given the optimal demand for shares $\alpha_{L,\tau}$

$$J_\tau^L = \int_\tau^\infty e^{-r(s-\tau)} v_s^*(\alpha_{L,\tau}) ds - (\alpha_{L,\tau} - \alpha_{L,\tau^-}) P(\alpha_{L,\tau}),$$

where $v^*(\alpha_{L,\tau})$ is the maximal net benefits flow to the large shareholder, given the optimal demand for shares $\alpha_{L,\tau}$.

J_τ^L can be written as $J_\tau^L = J_\tau^C + J_\tau^B$, where

$$J_\tau^C = J(\alpha_{L,\tau}, 0),$$

and

$$J_\tau^B = \int_\tau^\infty e^{-r(s-\tau)} \phi(b, \alpha_{L,s}) ds.$$

So, J_τ^C is the present value of the risk-adjusted instantaneous dividend process accrued to the large shareholder less (plus) the trading costs (gains), and J_τ^B is the present value of the private benefits flow. I define J_τ^C as the marginal utility of the large shareholder. Then, note that without private benefits the actual choice $\alpha_{L,\tau}$ is not optimal, that is

$$J_\tau^C < J(\alpha_{m,\tau}, 0),$$

where $\alpha_{m,\tau}$ denotes the number of shares that the large shareholder would choose as optimal solution in the absence of private benefits, thus behaving as a marginal investor with perfect information. $\alpha_{m,\tau}$ solves the utility maximization problem of the large shareholder when $b = 0$, for a given dynamics of μ_t and $\bar{\mu}_t$, risk aversion coefficient a^L, dividends shock volatility σ_D, initial updating weight $\bar{g}(0)$, and using the fact that in the absence of private benefits $\sigma_\epsilon^2 = 0$. Therefore, it is possible to compute both J_τ^C and $J(\alpha_{m,\tau}, 0)$, by using the implied dynamics of

μ_t, and the parameters estimates.

The gain in terms of private benefits must be at least equal to the loss in terms of marginal utility, choosing $\alpha_{L,\tau}$ rather than $\alpha_{m,\tau}$:

$$\phi(\alpha(L,t)) - \phi(\alpha(m,t)) > J\left(\alpha_{m,t}, 0\right) - J_{\tau}^{C},$$

from which I derive the lower bound for b, where $\phi(\alpha(L,t)) = b * \alpha_j(L,t)$ and $\phi(\alpha(m,t)) = b * \alpha_j(m,t)$, and $\alpha_j(L,t)$ and $\alpha_j(m,t)$ are the thresholds associated to $\alpha_{L,t}$ and $\alpha_{m,t}$, respectively. So,

$$b > \frac{J\left(\alpha_{m,t}, 0\right) - J_{\tau}^{C}}{\left(\alpha_j(L,t) - \alpha_j(m,t)\right)} = b^l.$$

On the other hand, the private benefits that the large shareholder can extract from the stake are not large enough to make convenient for the large shareholder to increase his stake up to a higher threshold. In other words, jumping to a higher threshold would produce gains in terms of private benefits not sufficient to cover the loss in terms of marginal utility:

$$\phi(\alpha_{j+1}(L,t)) - \phi(\alpha_j(L,t)) < J\left(\alpha_{j+1,t}, 0\right) - J_{\tau}^{C},$$

from which I derive the upper bound for b, where $\alpha_{j+1}(L,t)$ is the threshold above the one associated to $\alpha_{L,t}$ (e.g., if $\alpha_{L,t}$ is 53%, then $\alpha_j(L,t)$ is 50% and $\alpha_{j+1}(L,t)$ is 66%). Therefore,

$$b < \frac{J\left(\alpha_{j+1,t}, 0\right) - J_{\tau}^{C}}{\left(\alpha_{j+1}(L,t) - \alpha_j(L,t)\right)} = b^u.$$

In table of results in the following section, I report the mid point between lower and upper bounds, that is $(b^l + b^u)/2$.

1.4.7. Estimating the Price Impact

The estimation of the price impact of the private benefits involves a simple counterfactual analysis. First, note that the observed stock price is the stock price that reflects the private benefits. In fact, the observed stock price, at the discrete dates at which the large shareholder discloses the stake, depends on the large shareholder's stake.

Then, in the counterfactual analysis, I compare the observed stock price, at the discrete disclosure dates, with two unobservable stock prices: (i) the stock price in the absence of private benefits, that is the equilibrium stock price when there is one large shareholder that is fully informed about the true expected dividends payout of the firm, and (ii) the stock price in the absence of a large shareholder, that is the equilibrium stock price when the marginal investors do not receive any additional signal to update their beliefs at the discrete disclosure dates. Let P^m denote the equilibrium stock price in the absence of private benefits, and P^n the equilibrium stock price in the absence of a large shareholder, then

$$P_\tau^m = P\left(\bar{\mu}(\alpha_{m,\tau}), \rho(\alpha_{m,\tau})\right),$$

$$P_\tau^n = P\left(\bar{\mu}_\tau = \bar{\mu}_{t<\tau}, \rho(0)\right),$$

and

$$\psi^m = \frac{\tilde{P}_\tau - P_\tau^m}{P_\tau^m},$$

$$\psi^n = \frac{\tilde{P}_\tau - P_\tau^n}{P_\tau^n},$$

where ψ^m and ψ^n denote the (percentage) price impact in the two different cases, respectively.

I estimate the price impact at the disclosure discrete dates τ, that is at biannual frequency. I compute both P^m and P^n by using equation (9). As for P_τ^m, I first determine the optimal stake of the large shareholder in the absence of private benefits, $\alpha_{m,\tau}$, for a given dynamics of μ_t and $\bar{\mu}_t$, risk aversion coefficient a^L, dividends shock volatility σ_D, initial updating weight $\bar{g}(0)$, and using the fact that in the absence of private benefits $\sigma_\epsilon^2 = 0$. Then, I derive the equilibrium risk premium and the updated belief on the dividends drift by the marginal investors ($\bar{\mu}_\tau$), thus obtaining the equilibrium stock price. Computing P_τ^n, instead, requires only the dynamics of $\bar{\mu}_t$, and the equilibrium risk premium is simply given by $a^I \sigma_D^2 r$.

1.5. Results

1.5.1. Model Fit

Table III reports parameter estimates (top panel), and goodness of fit in terms of empirical moments on stock prices and large shareholders' stake and trading (bottom panel).

Table 3. Parameters Estimates and Goodness of Fit

	Mean	Median	Std	10th pct	90th pct
σ	0.024	0.016	0.028	0.002	0.051
σ_D	0.279	0.210	0.294	0.038	0.639
a_L	2.664	2.137	1.452	1.131	4.999
b	0.001	0.000	0.008	-0.003	0.009
$\bar{g}(0)$	0.137	0.041	0.172	0.000	0.493
σ_ϵ^2	0.584	0.764	0.405	0.000	0.998
	Mean		Median		
	Actual	Fitted	Actual	Fitted	
Stock Price volatility	1.98	2.05	1.22	1.19	
LS Trading volatility	0.029	0.031	0.016	0.014	
Trade (% of shares)	7.73	11.73	3.64	5.46	
	Actual	Fitted			
N of Trades $\geq 1\%$	10.76%	6.63%			

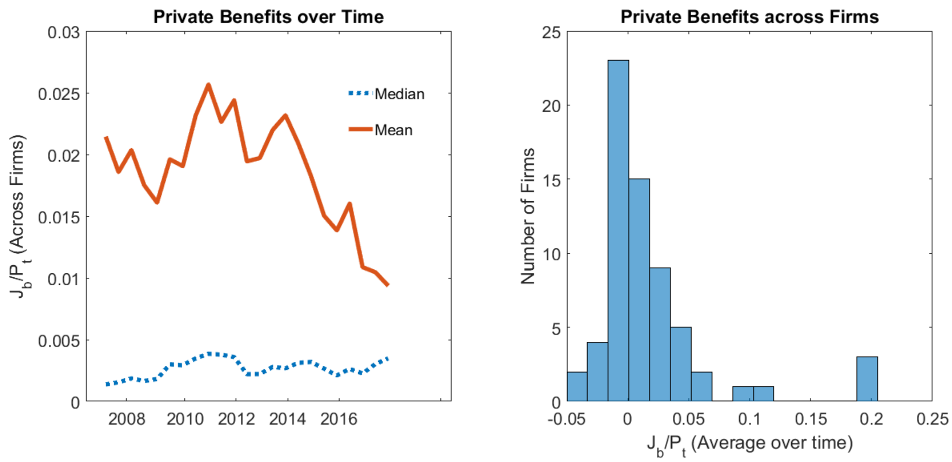
The table reports statistics on the parameters estimates for the final sample of 78 firms. The parameters are: the volatility of shocks to drift (σ), the volatility of shocks to dividends (σ_D), the large shareholder's risk aversion (a^L), the private benefits b, the initial updating weight of the marginal investors' belief $\bar{g}(0)$, and the noise in the large shareholder's trade observation σ_ϵ^2. The parameters are estimated firm-by-firm, by using maximum likelihood. σ and σ_D are estimated with daily stock prices and annual earnings-per-share, a^L, $\bar{g}(0)$ and σ_ϵ^2 are estimated with biannual stock prices and ownership shares, b is the average between the lower bound b^l and the upper bound b^u. The goodness of fit shows the comparison between actual empirical moments on stock price, large shareholder's stake and trading (averaged across firms), and estimation-implied moments on stock price, large shareholder's stake and trading. Moreover, the bottom line compares the observed number of trades with the estimated-implied number of trades, above 1% of the outstanding shares, as percentage of the total number of observations.

The volatility of shocks to dividends (σ_D) is widely larger than the volatility of shocks to the fundamental value of the firm (σ), proxied by the time-varying drift of the dividends process. We can interpret in the real data the continuous dynamics of the dividends process of the model as the daily arrival of news and information about the state of the firm.

The estimate of the large shareholder's risk aversion coefficient is close to the value calibrated for the aggregate risk aversion. Further, estimates on $\bar{g}(0)$ document heterogeneity across firms in terms of initial weight assigned to the large shareholder's stake by the marginal investors, and also in terms of noise of the information released by the large shareholders with their demand for shares.

The bottom panel compares the model in-sample predictions on the dynamics of stock prices and large shareholders' stake to their corresponding actual values in the data. The estimated model performs well in replicating these features of the data, even though the stylized facts on

Figure 5. Private Benefits over Stock Price



The left panel shows the mean (red line) and the median (blue dotted line) across firms, for each biannual date between March 2004 and September 2016, of the present value of the private benefits flow (J_b), divided by the stock price at the same date. The right panel shows the distribution of the present value of the private benefits flow divided by the stock price over the 78 sample firms, where for each firm I compute the average ratio over time.

the dynamics of large shareholders' stake and trading have not been explicitly targeted. Mean and median across firms of the large shareholders' trading volatility are almost equal between real data (1.98%, and 1.22%, respectively) and model-implied estimates (2.05%, and 1.19%), and close in terms of size of trade (7.73%, and 3.64% in real data, and 11.73% and 5.46% in the model). The differences between real data and model predictions, in terms of size of trade, and also in terms of number of trades (10.76% and 6.63%, respectively), are likely due to the exogenous thresholds imposed in model, so that the model tends to predict trades of larger blocks compared to the actual trades, but with a lower frequency.

1.5.2. Private Benefits

I use the estimated parameter b, and the actual stake of the large shareholder, to compute the present value per share of the private benefits flow, defined in the model as J_b, for each firm, and each disclosure date τ. I report results on J_b in terms of stock price, then $J_{b,\tau}/P_\tau$ is the measure of the present value of the private benefits flow extracted by large shareholders with their stake, in terms of equity value of the firm. The left panel of Figure 5 reports the time series of mean and median across firms at each biannual date, and the right panel of Figure 5 reports the distribution of the average over time for each sample firm. Statistics on the distribution are shown in Table 4.

Private benefits amount to approximatively 2% of equity value on average. This number is slightly lower than the estimate of Albuquerque and Schroth (2010). Similarly to Albuquerque and Schroth (2010), I also document a pronounced positive skewness in the distribution across firms, where the mean is much higher than the median, and does not provide an accurate picture of the results. Half of the large shareholders (40% in Albuquerque and Schroth (2010)) extract private benefits less than 1% of the total equity value, and the maximum rate of private benefits is 20% (15%).

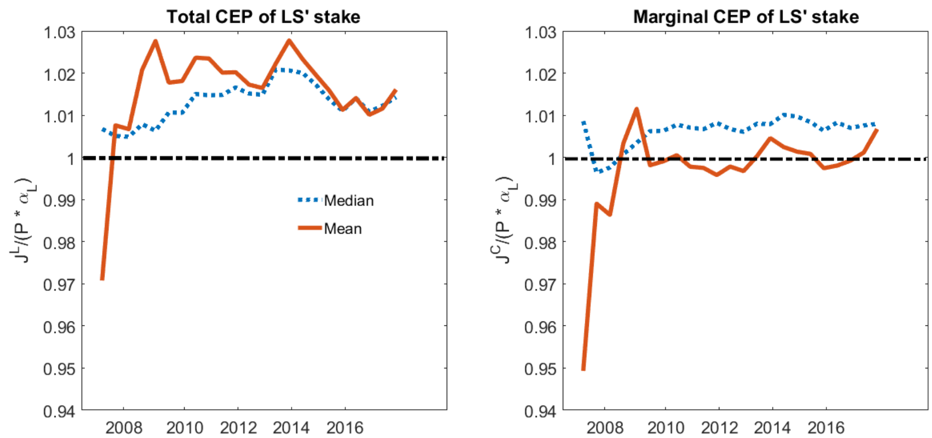
I compute the certainty equivalent payoff of the large shareholders' stake, defined in the model as J^L, that is the sum of the present value of the private benefits flow plus the marginal utility, defined as the present value of the risk adjusted dividends flow plus (minus) the trading costs (gains). Remind that the certainty equivalent payoff is simply the valuation of the stake by an investor, so the maximum price at which the investor is willing to buy that block of shares. I report results in terms of stock price, dividing by the large shareholders' stake, so that $J^L_\tau/(P_\tau * \alpha_{L,\tau})$ is equal to 1 when the large shareholder valuates one share exactly as the market does. The left panel of Figure 6 reports the time series of mean and median across firms at each biannual date for $J^L_\tau/(P_\tau * \alpha_{L,\tau})$, and the right panel of Figure 6 reports the time series of mean and median across firms at each biannual date for $J^C_\tau/(P_\tau * \alpha_{L,\tau})$, that is the marginal component of the certainty equivalent payoff of the large shareholders' stake. Statistics on the distributions of the average over time for each firm are shown in Table 4.

Intuitively, large shareholders value their stakes more than the market price. With risk aversion, an investor is willing to buy a stock only if the certainty equivalent is above the market price of the stock. The difference between the total and the marginal large shareholders' certainty equivalent is due to the private benefits. So, private benefits contribute to compensate large shareholders for holding very undiversified portfolios, thus keeping the stake valuation of the large shareholders always above the market price of the stock.

1.5.3. Price Impact

Finally, I use the model estimates to compute the (percentage) price impact of the private benefits of control, and of the overall large shareholder's stake, defined as ψ^m and ψ^n, respectively. Figure 7 reports the time series of mean and median across all sample firms at each biannual date, and Figure 8 reports the time series of the average across two different types

Figure 6. Certainty Equivalent of Large Shareholders



The left panel shows the mean (red line) and the median (blue dotted line) across firms, for each biannual date between March 2004 and September 2016, of the certainty equivalent payoff of the largest shareholder (J^L), divided by the stock price at the same date, and the largest shareholder's stake. The right panel shows the mean (red line) and the median (blue dotted line) across firms, for each biannual date between March 2004 and September 2016, of the marginal component of the certainty equivalent payoff of the largest shareholder (J^C), divided by the stock price at the same date, and the largest shareholder's stake.

of large shareholders at each biannual date: corporations and individuals. Statistics on the distributions of the average over time for each firm are shown in Table 4.

In general, private benefits of control have positive impact on stock prices, so they are not extracted with cost for the rest of the company shareholders. The average price impact over time fluctuates between 1% and 3%, and the mean is always significantly larger than the median, signalling again highly positive skewness. The price impact of private benefits, in fact, is quite heterogenous across firms. For 15% of the firms private benefits have a negative impact greater than 1%, and for the same proportion of firms private benefits have a positive impact larger than 5%.

The impact of the large shareholder's stake, overall, is even more beneficial to the rest of the investors. For 15% of the firms, in fact, the presence of the large shareholder increases the stock price by more than 10%, and for only 6% of the firms the negative impact is greater than 1%. In general, the positive price impact of both private benefits and large shareholder's stake is substantially greater during the crisis of 2011-2012. This suggests that large shareholders support stock prices and are significantly beneficial to the rest of the investors over negative economic cycles. In untabulated results, I find that the average price impact of private benefits

Figure 7. Price impact over time. All Firms

The left panel shows the mean (red line) and the median (blue dotted line) across firms, for each biannual date between March 2004 and September 2016, of the price impact of the private benefits extracted by the largest shareholder (ψ^n). The right panel shows the mean (red line) and the median (blue dotted line) across firms, for each biannual date between March 2004 and September 2016, of the price impact of the overall large shareholder's stake (ψ^m).

is 3.17% at the beginning of 2012 (at the boom of the European sovereign debt crisis) and only 0.96% at the beginning of 2007, before the start of the great financial crisis. Moreover, the average price impact of large shareholders, overall, is 7.87% at the beginning of 2012, and only 2.27% at the beginning of 2007.

Table 4 shows that corporate large shareholders exert more beneficial effect to the stock price than individuals, both in terms of private benefits and presence of large shareholder, and both over time and across firms. The combination between positive price impact of corporate large shareholders and positive estimate of private benefits extracted by corporate large shareholders offers the evidence of a synergistic effect between owner and owned firm. In this case, in fact, I document a reciprocal beneficial effect so that private benefits are not extracted at cost for the rest of shareholders, and they can be properly defined as synergies.

1.6. Conclusion

The paper estimates private benefits of control using the restrictions provided by a dynamic model of optimal shareholding, with asymmetric information and heterogenous shareholders. The model equilibrium conditions allows to quantify the price impact of large shareholders and private benefits of control over time. The estimation results provide evidence that large shareholders have positive impact on stock prices that does not vanish over time, so that they

Table 4. Private Benefits and Stock Price

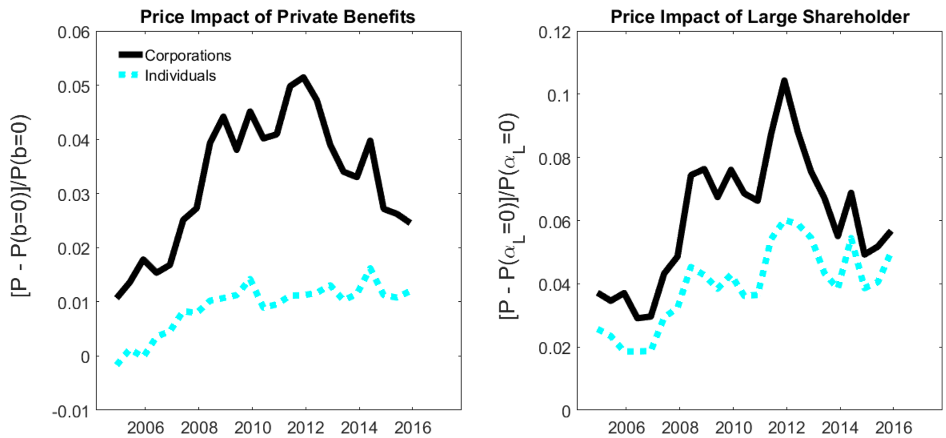
	Panel A: All Sample			
	Mean	Median	10th pct	90th pct
J^L/P	1.015	1.014	0.984	1.053
J_b/P	0.018	0.003	-0.015	0.054
J_C/P	0.997	1.006	0.945	1.054
ψ^m	0.020	0.002	-0.021	0.096
ψ^n	0.048	0.028	-0.001	0.128
	Panel B: Corporations			
	Mean	Median	10th pct	90th pct
J^L/P	1.018	1.022	0.982	1.063
J_b/P	0.029	0.009	-0.016	0.105
J_C/P	0.989	1.002	0.865	1.052
ψ^m	0.032	0.003	-0.018	0.127
ψ^n	0.060	0.046	0.001	0.149
	Panel C: Individuals			
	Mean	Median	10th pct	90th pct
J^L/P	1.013	1.014	1.001	1.023
J_b/P	0.011	0.001	-0.010	0.011
J_C/P	1.002	1.007	0.967	1.059
ψ^m	0.009	0.002	-0.026	0.029
ψ^n	0.039	0.021	-0.002	0.132

The table reports statistics on the total certainty equivalent payoff of the large shareholder (J^L), the marginal certainty equivalent payoff of the large shareholder (J^C), and the present value of the private benefits flow (J^b), divided by the stock price at the disclosure date τ, and on the price impact of private benefits (ψ^m) and the total price impact of the large shareholder's stake (ψ^n). First, I compute the average over time for each firm, then I report in the table mean, median, and 80% confidence interval across firms, for all sample firms, and by type of large shareholder.

extract private benefits without cost for the rest of the company shareholders. On the other side, private benefits contribute to compensate large shareholders for holding very undiversified portfolios. When the large shareholder is a corporation, for instance, this reciprocal beneficial effect sheds light on a synergy between owned and owner firm.

While this paper proposes the first structural approach to quantify the price impact of large shareholders over time, both the theoretical and the empirical analysis is far from being exhausted. The main challenge for future research is to disentangle the two drivers of the heterogeneity between large shareholder and marginal investors, namely the opportunity to extract private benefits and the superior information.

Figure 8. Price impact over time. Types of shareholders



The left panel shows the mean, across the firms where the largest shareholder is a corporation (black line) and the firms where the largest shareholder is an individual (sky-blue line), for each biannual date between March 2004 and September 2016, of the price impact of the private benefits extracted by the largest shareholder (ψ^n). The right panel shows the mean, across the firms where the largest shareholder is a corporation (black line) and the firms where the largest shareholder is an individual (dotted line), for each biannual date between March 2004 and September 2016, of the price impact of the overall large shareholder's stake (ψ^m).

1.7. Appendix

1.7.1. Proofs: Investors' Optimality Conditions

To derive the optimality conditions of the marginal investor, I follow the same approach of DeMarzo and Urosevic (2006). First, I formulate the investor's conjecture about the price process, and I conjecture a value function for the investor in the customary form as in DeMarzo and Urosevic (2006). Then, I derive the Bellman equation, to be maximized with respect to the control variables α and c. Finally, I derive the equilibrium risk premium, and the equilibrium share price in Proposition 1.

In a CARA-utility framework, with normality assumption on the dividend payout, the conjecture of the price process is the following

$$dP_t = (rP_t + \rho_t - \bar{\mu}_t)dt,$$

that is the share price grows at the riskless rate, plus a risk premium component ρ to compensate the investor's risk aversion, and to be determined in equilibrium, minus the biased expected dividend payout.

Then, for a given price process, the investor's optimality conditions are the followings:

$$u_c = J_W,$$

where u_c denotes the marginal utility from consumption, and J_W is the partial derivative of the value function of the investor, that is his expected payoff on each point in time, with respect to the state variable wealth, and

$$\alpha_{i,t} = \frac{\mu_{i,t} - \bar{\mu}_t + \rho}{ar\sigma^2},$$

where $\bar{\mu}_t$ is the average expected dividend across the marginal investors, and the risk premium ρ compensates the investor's risk aversion. By market clearing, that is $\int_i \alpha_{i,t} di = 1 - \alpha_{L,t}$ for each t, the equilibrium risk premium is given by

$$\rho_t = (1 - \alpha_{L,t}) a^I r\sigma^2,$$

where $\alpha_{L,t}$ is the stake held by the large shareholder at time t, and a^I is the aggregate risk aversion coefficient

$$\frac{1}{a^I} = \int_i \frac{1}{a^i} di$$

In fact, let the riskless holdings of the investor to evolve as follows, for given consumption level c_t, shareholding α_t, and share price P_t

$$dB_t = (rB_t - c_t)dt + \alpha_t dD_t - P_t d\alpha_t,$$

The wealth of the investor is defined as $W = \alpha P + B$. Then, the expected value of the wealth accumulation over time, $dW_t = dB_t + \alpha_t dP_t + P_t d\alpha_t$, is

$$\frac{1}{dt} E_t[dW] = rW_t + \alpha_t(\mu_{i,t} + \rho_t - \bar{\mu}_t) - c_t = rW_t + \alpha_t(\rho_t + e_{i,t} - \bar{e}_t) - c_t$$

Consider the value function

$$J(W,t) = \frac{1}{r} u\left(r[W + y_t] + \frac{(R - r)}{ar} \right),$$

where y_t is the certainty equivalent of the investor at time t

$$y_t = \int_t^\infty e^{-r(s-t)} \left(\frac{1}{2} \alpha_t^2 r \sigma^2 a \right) ds$$

The value function $J(W,t)$ satisfies the following HJB equation

$$\max_{\alpha,c} J_t + J_W dW + \frac{1}{2} J_{WW} dW^2 + u(c) = RJ,$$

Substituting dW, taking the expectation at t, and maximizing over c and α, I obtain the equations of optimality in the desired form, where $J_W = arJ$, and $J_{WW} = -aJ_W$.[7]$\square$

1.7.2. Proofs: Equilibrium Stock Price

Consider the i-th investor's certainty equivalent at time t for the stake $\alpha_{i,t}$,

$$\int_t^\infty e^{-r(s-t)} \left(\alpha_{i,t}\mu_{i,t} - \frac{1}{2} \alpha_{i,t}^2 a^i \sigma_D^2 r \right) ds,$$

that, with constant interest rate r, can be written as

$$\frac{\alpha_{i,t}\mu_{i,t} - \frac{1}{2}\alpha_{i,t}^2 a^i \sigma_D^2 r}{r},$$

that in equilibrium must be equal to the cost of the stake, that is

$$\frac{\alpha_{i,t}\mu_{i,t} - \frac{1}{2}\alpha_{i,t}^2 a^i \sigma_D^2 r}{r} = \alpha_{i,t} P_t$$

Then, take the derivative of both sides with respect to α, and solve for α

$$\alpha_{i,t} = \frac{\mu_{i,t} - rP_t}{ar\sigma^2}.$$

Then, impose the market clearing condition, so that $\int_i \alpha_{i,t} = 1(1 - \alpha_{L,t})$, substitute $\alpha_{i,t}$ and solve for P_t,

$$P_t = \frac{\bar{\mu}_t - (1 - \alpha_{L,t}) * a^I \sigma_D^2 r}{r},$$

[7] For the proof that the Bellman equation holds, and the conditions to avoid doubling strategies and Ponzi scheme, the reader can refer to DeMarzo and Urosevic (2006).

where $\bar{\mu}_t = \int_i \mu_{i,t} di$, and $a^I = \int_i (1/a^i) di$. So, P_t can be written as

$$P_t = \int_t^{\infty} e^{-r(s-t)} (\bar{\mu}_s - \rho_s) ds$$

where $\rho_s = (1 - \alpha_{L,t}) * a^I \sigma_D^2 r$

1.7.3. Proofs: Evolution of average expected dividend

First, I derive the evolution over time of the demand for shares of the investors, by setting $\alpha_{i,t}$ as function of $\mu_{i,t}$ and $\bar{\mu}_t$, then using Ito's lemma. The investor, then, buys (sells) company shares when the change in her expectation about the future dividend is greater (lower) than the change in the average expectation about the future dividend. In fact, setting $\alpha_{i,t} = f(\mu_{i,t}, \bar{\mu}_t)$, then

$$d\alpha_{i,t} = \frac{1}{ar\sigma^2}(d\mu_{i,t} - d\bar{\mu}_t),$$

where $d\mu_{i,t} = k_i(dDt - E_{i,t}(dD_t))$.

Note then that the market clearing condition holds also in a dynamic fashion: $\int_i d\alpha_{i,t} = 0$, so that

$$\int_i d\mu_{i,t} di = M d\mu_t$$

where M is the measure of the marginal investors. It follows that

$$d\mu_t = \frac{1}{M} \int_i (k_i(dDt - E_{i,t}(dD_t))) = \bar{k}(dD_t - \mu_t),$$

since k_i and $E_{i,t}(dD_t)$ are independent.

1.7.4. Proofs: Proposition 1

The proof of the proposition 1 is simply the combination between the two previous proofs, at the disclosure dates τ. When the large shareholder discloses his stake, then each investor i revises the prior on μ_t, and so the optimal demand for shares, so that, applying again Ito's lemma,

$$d\alpha_{i,\tau} = \frac{1}{ar\sigma^2}(d\mu_{i,\tau} - d\bar{\mu}_\tau),$$

where $d\mu_{i,\tau} = g_i(\tau^-)(\alpha_{L,\tau} - E_{t<\tau}(\alpha_{L,\tau}))$, with $E_{t<\tau}(\alpha_{L,\tau})$ equal across the investors as it is common knowledge at τ, and

$$d\bar{\mu}_\tau = \bar{\mu}_\tau - \bar{\mu}_{t<\tau}$$

Again, the market clearing condition holds in a dynamic fashion: $\int_i d\alpha_{i,\tau} = 0$, so that

$$\int_i d\mu_{i,\tau} di = M d\bar{\mu}_\tau$$

It follows that

$$d\bar{\mu}_\tau = \frac{1}{M}\int_i (g_i(\alpha_{L,\tau} - E_{t<\tau}(\alpha_{L,\tau})))di = \bar{g}(\tau^-)(\alpha_{L,\tau} - E_{t<\tau}(\alpha_{L,\tau}))$$

Then, consider the i-th investor's certainty equivalent at time τ for the stake $\alpha_{i,\tau}$,

$$\frac{\alpha_{i,\tau}\mu_{i,\tau} - \frac{1}{2}\alpha_{i,\tau}^2 a^i \sigma_D^2 r}{r},$$

that in equilibrium must be equal to the cost of the stake, that is

$$\frac{\alpha_{i,\tau}\mu_{i,\tau} - \frac{1}{2}\alpha_{i,\tau}^2 a^i \sigma_D^2 r}{r} = \alpha_{i,\tau} P_\tau$$

Then, take the derivative of both sides with respect to α, and solve for α

$$\alpha_{i,\tau} = \frac{\mu_{i,\tau} - rP_\tau}{ar\sigma^2}.$$

Then, impose the market clearing condition, so that $\int_i \alpha_{i,\tau} = (1 - \alpha_{L,\tau})$, substitute $\alpha_{i,\tau}$ and solve for P_τ,

$$P_\tau = \frac{\bar{\mu}_\tau - (1 - \alpha_{L,\tau}) * a^I \sigma_D^2 r}{r},$$

where $\bar{\mu}_\tau = \int_i \mu_{i,\tau} di$, and $a^I = \int_i (1/a^i) di$. So, P_τ can be written as

$$P_\tau = \int_\tau^\infty e^{-r(s-\tau)}(\bar{\mu}_s - \rho_s)ds$$

where the equilibrium risk premium is $\rho_\tau = (1 - \alpha_{L,\tau}) * a^I \sigma_D^2 r$

1.7.5. Proofs: Proposition 2

Next, consider the large shareholder's certainty equivalent payoff,

$$V(\alpha_{L,\tau}) - (\alpha_{L,\tau} - \alpha_{L,\tau^-})P_\tau,$$

write P_τ and V_τ in the explicit form, take the derivative with respect to α, and solve explicitly for $\alpha_{L,\tau}$,

$$\alpha_{L,\tau} = \frac{\mu_t - \bar{\mu}_\tau + (1 + \alpha_{L,\tau^-})a^I \sigma_D^2 r + \phi(\alpha_{L,\tau})}{\Delta},$$

where $\Delta = 2a^I\sigma_D^2 r + a^L\sigma_D^2 r$.

Now, let define

$$\alpha_{L,\tau}(\bar{\mu}_{t<\tau}) = \frac{(1 + \alpha_{L,\tau^-})a^I \sigma_D^2 r}{\Delta},$$

then

$$\alpha_{L,\tau} = \frac{\mu_t - \bar{\mu}_\tau + \phi(\alpha_{L,\tau})}{\Delta} + \alpha_{L,\tau}(\bar{\mu}_{t<\tau})$$

then substitute $\bar{\mu}_\tau = \bar{\mu}_{t<\tau} + \bar{g}(\tau^-)(\alpha_{L,\tau} - \alpha_{L,\tau}(\bar{\mu}_{t<\tau}))$, collect common terms over $\alpha_{L,\tau}$ and $\alpha_{L,\tau}(\bar{\mu}_{t<\tau})$, and solve again explicitly for $\alpha_{L,\tau}$, thus obtaining the desired form.

1.7.6. Kalman Filter

The Kalman filter allows to reconstruct the dynamics of a latent variable, by using observable variables, and a given known relationship between latent and observable variables. The relationship between observed and unobserved variables forms the *measurement equation*, while the evolution over time of the latent variable is called *transition equation*.

In few words, the filter starts from a prior on the latent variable, and forms a prediction of the next step value of the latent following the diffusion described in the transition equation. Then, the filter makes a forecast of the observable variable based on the prediction of the latent, by using the relationship described in the measurement equation. At each time step, the filter generates an error, given by the distance between the actual value of the observable and the forecast. The error is then used to update the prior on the latent, for a given weight assigned to the error, which is called *Kalman Gain*. Moreover, the errors depend on the model parameters, so under the assumption of normality the errors are used to construct a likelihood function that is maximized with respect to the model parameters.

In my set up, the transition equations describe the evolution of the average expected dividend across the marginal investors, and the true time-varying dividends drift, that are defined in the equation (7) and (8), respectively. The parameter $\bar{k}$ in equation (7) is allowed to vary over time

$$k_t = \frac{\nu_t}{\nu_t + \sigma_D^2},$$

where $\nu_t = w_t + \sigma$, and at each time step w_t is updated by using $(1 - k_{t-1})\nu_{t-1}$, and initializing the recursion with a large value of ν_0. This procedure allows to proxy the prior update on the dividends drift across the marginal investors. Moreover, the state prediction on $\bar{\mu}_t$ has variance equal to $\omega_t + k_t^2\sigma_D^2$, where the second term is the variance of the state, conditional at time t, derived by the market clearing equation. The state prediction on μ_t has variance equal to $\omega_t + \sigma^2$, ω_t is updated at each time step according to the Kalman Gain, and the recursion is initialized with a large value of ω_0.

The measurement processes, instead, come from the pricing equation defined in (9), and I assume that the actual prices are observed with noise, for instance due to microstructure issues, so that

$$\tilde{P}_t = P_t + \epsilon_{1,t},$$

where the noises are gaussian, with zero mean and variance σ_P^2.

Hence, starting from an initial guess $\{\mu_0,\bar{\mu}_0\}$, I define the prediction for $\{\mu_1,\bar{\mu}_1\}$ by using the transition equations. Then, given the prediction on the state, I compute the forecast for the share price P_1, thus obtaining a prediction error, by using the actual observations $\tilde{P}_1$. Combining the errors with the Kalman Gain, I update the prior for the state, and iterate recursively the filter up to the end of the time series. The Kalman Gain is the optimal weight to assign to prediction error in order to revise the prior on the state variable. It is derived by minimizing the conditional variance (covariance matrix) of the state variable(s).

Table 5. Parameters Estimates: Simulation Study

	True	Mean	Median	10th pct	90th pct
σ	0.1	0.093	0.093	0.087	0.099
σ_D	0.2	0.202	0.199	0.132	0.274
a_L	8	8.71	7.97	7.10	10.36
b	0.02	0.019	0.021	0.007	0.027
$\bar{g}(0)$	0.1	0.097	0.099	0.082	0.105
σ_ϵ^2	0.2	0.189	0.188	0.138	0.245

1.7.7. Simulation Study

To test the accuracy of the estimation methodology, I perform a numerical analysis over 1000 simulations. I simulate the dynamics of the dividends time varying drift, and the dividends process of the firm, according to the equations (1) and (2), with daily frequency ($\delta t = 1/250$), for a set of arbitrary values of σ and σ_D. I also simulate a mass of marginal CARA-maximizer investors, who observe dD_t and update their prior on μ_t according to their heterogenous prior variances, thus obtaining an average expected dividend across the marginal investors,$\bar{\mu}_t$, for each t.

Next, for given parameter b and thresholds of the private benefits function, aggregate risk aversion coefficient a^I, large shareholder's risk aversion coefficient a^L, initial average weight $\bar{g}(0)$ and noise σ_ϵ assigned to the large shareholder's choice by the marginal investors for updating their beliefs at the discrete disclosure dates, I obtain the time series of the large shareholder's stake and the equilibrium stock price with biannual frequency ($\delta\tau = 0.5$).

Then, using daily stock prices, I estimate σ and σ_D, and I infer the dynamics of μ_t and $\bar{\mu}_t$, which I use in the second step to estimate the remaining parameters by using biannual stock prices and large shareholder's stakes. Finally, I identify the parameter b following the identification approach described above. Table 5 reports mean, median, and 80% confidence interval, over 1000 simulations, for the six parameters against the true arbitrary value of the parameters.

2. The Relative Pricing of Sovereign Credit Risk After the Eurozone Crisis

2.1. Introduction

Classical asset pricing theory suggests that investors should be compensated with higher return when they hold a security issued by a company subject to default risk. In other words, investors seek a premium for the risk they are bearing by holding a risky security, and the default risk should be priced in the expected return, thus implying a positive and monotonic relationship between risk and expected return over a cross-section of assets. The empirical validation of this positive relationship has attracted much attention in the finance literature, and still receives great consideration in the academic debate. Researchers and practitioners have intensively investigated whether stocks of riskier firms generate higher expected returns. In this paper, we move the focus of the analysis to the sovereign level, by testing whether bonds of riskier countries pay out yields sufficiently higher compared to those of safer countries, to compensate properly their debtholders.

The idea is that the positive and monotonic relationship between default risk and expected return in the general framework, and between default risk and bond yields in our sovereign analysis, is only a necessary but not sufficient condition to guarantee that investors are compensated for the risk they are bearing. Investors, in fact, are willing to switch the safer and riskier assets only if they expect to receive proper additional return. It turns out that the difference in terms of default risk should be reflected on the difference in terms of expected returns. Therefore, we investigate whether the differences in terms of default risk across countries are reflected on the differences in terms of bond yields, thus testing whether the sovereign default risk is priced in the cross-section of the bond yields. Our main research question, then, is whether it exists a consistent, rather than only positive and monotonic, cross-sectional relationship between CDS spreads and bond yields across countries.

Following the results of the literature on credit risk, we measure default risk using Credit Default Swaps (CDS) spreads.[8] By using the CDS spreads, we can also exploit the well-known no-arbitrage relationship between CDS spreads and bond yields. CDS and bonds issued by

[8]Several papers have documented that the CDS market provides a reliable proxy of default risk (Blanco et al. (2005), Almeida and Philippon (2007), and Longstaff et al. (2005), among others)

the CDS reference entity, in fact, produce similar exposure to the investor in terms of risk and return: the CDS provides protection to the acquirer in case of default of the reference entity, while the bond pays out yields to the bondholder as long as the reference entity is able to comply with its obligations. Hull et al. (2004) point out that, under a set of assumptions that ensure the absence of frictions in the market, a portfolio including a bond and the protection on the bond provided by a CDS generates cash flows equal to a riskless bond in all states of the world. Consequently, the CDS spread should be equal to the excess bond yield over the risk-free rate to prevent arbitrage. This equilibrium condition is called the zero-basis condition, where the basis is the difference between the CDS spread and the asset swap spread of the bond.

Therefore, our paper builds a bridge across two fields of the literature: the analysis on the empirical validation of the risk-return relationship, and the study on the empirical validation of the CDS-Bond relationship. By uncovering this strand of research, we make our main contribution: we show that the deviation from the CDS-Bond no-arbitrage condition for single countries, while does not affect the positive and monotonic relationship between CDS spreads and bond yields over the cross section of countries, prevents the consistent cross-sectional relationship between CDS spreads and bond yields.

We study the relationship between sovereign CDS and sovereign bonds for European countries during and after the sovereign debt crisis of 2010-2012. Our main finding is the following: an inconsistent cross-sectional relationship between CDS spreads and bond yields emerges during the crisis period for all European countries. However, after the announcement of the Outright Monetary Transaction (OMT) program by the European Central Bank on July 26, 2012, the consistent cross-sectional relationship between default risk and bond yields is restored for the Eurozone countries only.

Mispricing has been documented over different markets. Dammon et al. (1993) present evidence on the violation of the arbitrage pricing restrictions across high-yield bonds issued by a representative company, while Lamont and Thaler (2003) show that inconsistency in stock prices emerges during equity carve-outs due to equity short-selling constraints and very costly trading strategies with options. Fleckenstein (2013), and Fleckenstein et al. (2014), shed light on the mispricing puzzle across inflation-linked and nominal bonds in the US. Corradin and Rodriguez-Moreno (2014) report mispricing across USD and Euro denominated sovereign bonds, widened by massive purchases programs of central banks, and Jarrow et al. (2018) offer puzzling results on the expected relationship across spreads of CDS traded with different maturities.

Recently, the literature on mispricing has extensively tested the violation of the theoretical no-arbitrage condition between CDS spreads and bond yields. Fontana (2010) shows that during the great recession the CDS-Bond basis tends to be negative, and does not converge to zero because of funding costs that prevent the purchase of the bonds to exploit the arbitrage opportunity. Bai and Collin-Dufresne (2013) confirm this result, however they report a large cross-sectional dispersion of the basis across firms, that they relate to several potential limit to arbitrage explanations. They also confirm that the basis is slightly positive during normal times, as reported also by Longstaff et al. (2005), and Blanco et al. (2005). These papers argue that CDS spreads are faster in price discovery, thus reacting more quickly to changes in the credit conditions. A few papers focus on the CDS-Bond basis at the sovereign level. Foley-Fisher (2010) strengths the result of Fontana (2010) for sovereign CDS and bonds, and Fontana and Scheicher (2016) stress that the basis is negative during the crisis periods in particular for risky countries, such as the peripheral countries of the Eurozone. Arce et al. (2013) argue that the results of Longstaff et al. (2005), and Blanco et al. (2005) are only partially true at the sovereign level, and price discovery is largely state dependent.

We start our analysis by documenting that the equilibrium condition between CDS spreads and bond yields is violated before the announcement of the OMT program for all European countries and is restored afterwards for the Eurozone countries only, and in particular for the peripheral countries of the Eurozone. Instead, the deviation from the equilibrium condition persists even after the OMT announcement for the European countries out of the Eurozone.

Since the violation of the equilibrium condition generates arbitrage opportunities, we corroborate the result with a portfolio analysis based on the deviation from the zero-basis condition. We show that arbitrage opportunities are large and persistent before the OMT announcement across all European countries and then quickly disappear after the OMT announcement for Eurozone countries only. Instead, arbitrage opportunities persist even after the OMT announcement for countries outside the Eurozone.

We proceed with our analysis by showing that the deviation from the zero-basis condition does not imply the violation of the expected positive and monotonic relationship across countries between CDS spreads and bond yields. In fact, we show that the cross-sectional rank correlation between CDS spreads and bond yields is always close to 1 for both Eurozone and non-Eurozone countries. This result provides evidence that riskier countries issue debt securities that pay out higher yields.

The empirical contradiction of the positive relationship between the risk and expected return is known in the financial literature as a *distress puzzle*. The distress puzzle has been widely investigated in the context of corporate securities by studying the relationship between the default risk and expected stock return. The empirical evidence is far from univocal. Vassalou and Xing (2004) find that the default risk is consistently priced in the cross-section of the expected equity returns, while Campbell et al. (2008) report that riskier firms tend to generate lower expected returns. Chava and Purnanandam (2010) confirm the positive relation between default risk and stock returns using the cost of capital as proxy for the expected returns rather than the realized returns, while Garlappi and Yan (2011) document an hump-shaped relation between default risk and stock returns due to shareholders' recovery upon default. Friewald et al. (2014) argue that both positive and negative relations between default risk and expected returns are consistent with structural models. Ferreira Filipe et al. (2016) and Anginer and Yildizhan (2017) show that the default risk is consistently priced in the cross-section of the expected equity returns when taking into account the systematic component of the default probability, that is the exposure to aggregate default risk. Gao et al. (2017) and Eisedorfer et al. (2017) document that distress anomalies prevail in developed countries, while Aretz et al. (2018) offer evidence of a positive risk-return relationship at the international level, by combining the approach of Campbell et al. (2008) with the distance-to-default based on the approach of Merton (1974). Schneider et al. (2019) argue that distress anomalies can be rationalised considering a skewness premium on stock returns. To the best of our knowledge, however, an analysis of the puzzle at the sovereign level is still missing. As countries do not issue equity, we focus on debt securities.

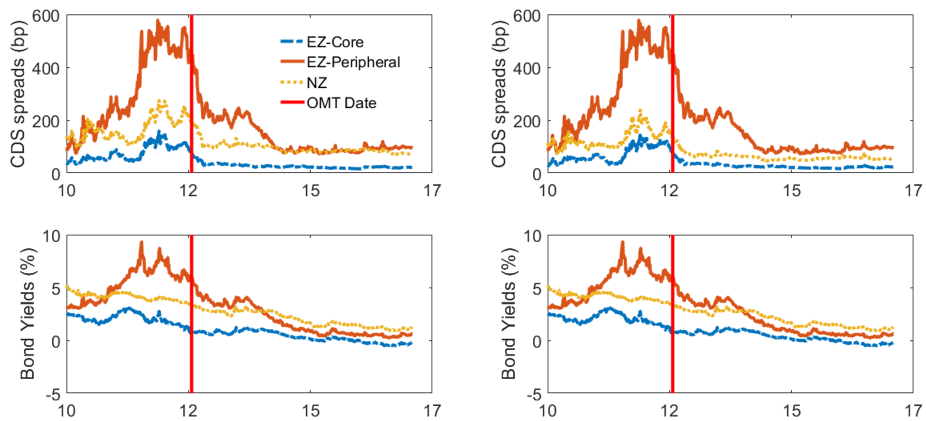
Then, we make one step ahead, and we study whether riskier countries pay out sufficiently higher bond yields compared to those of the safer countries, by testing whether the differences in terms of bond yields are consistent with the differences in terms of the default risk priced in the corresponding CDS spreads. We show that an inconsistent cross-sectional relationship between CDS spreads and bond yields emerges during the crisis period for Eurozone countries and is restored after the announcement of the OMT program. Therefore, while the deviation from the zero-basis equilibrium condition does not affect the monotonicity in the cross-sectional relationship between CDS spreads and bond yields, it generates inconsistency in the cross section of the bond yields across countries: the differences across countries in terms of the default risk, priced in the CDS spreads, are not consistently priced in the cross section of the bond yields.

To determine the proper distance between bond yields across countries, we adopt a contingent claim model. In the model, bond and CDS are implicitly related at each point in time, as both the securities are derivative contracts on the same underlying quantity, which are the assets and liabilities of the reference entity. In particular, we adopt a first-passage time model, where the issuer defaults as soon as the value of the assets crosses from above a default boundary, which is assumed to be deterministic and constant. This framework is an extension of the seminal model of Merton (1974), where the issuer may default only at the maturity of the liability. Gapen et al. (2011) introduce a contingent claim analysis to study sovereign credit risk using a Merton model. We estimate the model with a nonlinear Kalman filter using daily data on CDS spreads, and we compute the bond yields implied by the model estimation using Monte Carlo (MC) simulations. These yields are consistent with the default risk of the country priced in the CDS spreads.

We corroborate our results with a portfolio analysis based on the difference between observed and implied bond yields. We show that arbitrage opportunities are large and persistent before the OMT announcement across all European countries, then converge to zero after the OMT announcement for the Eurozone countries only. Arbitrage opportunities do not disappear even after the OMT announcement for the European countries outside of the Eurozone.

Finally, we conjecture that the arbitrage opportunities before the OMT announcement were created by high transaction costs. Therefore, for each country, we estimate the threshold below which arbitrage profits are insufficient to cover the costs to implement the strategy. The idea is that arbitrageurs enter the market only if the arbitrage strategy generates profits above such costs. We show that, before the OMT announcement, the arbitrage opportunities are not cleared because of high transaction costs. Then, we estimate a strong reduction in the transaction costs for the Eurozone countries only, following the ECB intervention. Consequently, the arbitrage opportunities are cleared, and the equilibrium condition in the Eurozone sovereign debt market is restored. However, we do not estimate a similar reduction in the transaction costs for the

Figure 9. CDS Spreads and Bond Yields.



The figure reports the mean and median across countries for the sovereign CDS spreads and bond yields between the 1^{st} of January 2010 and the 1^{st} of February 2017, at 5-year maturity for three different groups of countries: Eurozone-core (blue line), Eurozone-peripheral (red line), and non-Eurozone (yellow line). The CDS spreads are expressed in basis points, and the bond yields are expressed in percentage terms. The red line is the OMT announcement date.

non-Eurozone countries. Therefore, for those countries, we observe a persistent mispricing even after the OMT announcement.

Our paper is organized as follows. We describe the data in the next section. Then, we provide empirical evidence on the relationship between CDS spreads and bond yields during and after the OMT announcement in Section 2.3. In Section 2.4, we focus on the cross-sectional analysis of CDS spreads and bond yields. We detail the underlying credit risk model and our estimation methodology to compute the implied bond yields. In addition, we compare observed and implied yields, and we perform the cross-sectional correlation analysis between CDS spreads and bond yields. Finally, we estimate the transaction costs before and after the OMT announcement and compare such costs with the arbitrage profits in Section 2.5. Section 2.6 concludes the paper.

2.2. Data

Our main source of data is Thomson Reuter's Datastream. We download daily data for sovereign CDS spreads and sovereign bond yields for several European countries from January 2010 to February 2017. We collect 1850 daily observations for each country, for both CDS spreads and bond yields, and for three time maturity levels: 1, 5, and 10 years. Datastream provides reference par yields for sovereign bonds at different maturities. The par yield is the

Table 6. Descriptive Statistics by Country.

Statistics:	CDS Spreads			Bond Yields		
Mean	B/OMT	A/OMT	Diff	B/OMT	A/OMT	Diff
Eurozone						
Core:						
Austria	78.19	20.22	-57.97*	3.14	1.21	-1.93*
Belgium	143.11	33.93	-109.19*	3.77	1.44	-2.33*
Finland	46.50	24.74	-21.76*	2.79	1.14	-1.65*
France	83.17	31.86	-51.31*	3.13	1.37	-1.76*
Germany	39.15	12.58	-26.57*	2.48	0.93	-1.55*
Netherlands	67.26	31.74	-35.53*	7.63	2.31	-5.31*
Peripheral:						
Ireland	485.07	80.00	-405.07*	4.94	2.81	-2.13*
Italy	229.15	138.40	-90.75*	2.80	1.15	-1.65*
Portugal	633.77	247.09	-386.67*	8.85	4.35	-4.49*
Slovakia	136.00	61.90	-74.10*	4.18	1.89	-2.29*
Slovenia	164.69	168.27	3.58	5.10	2.90	-2.20*
Spain	243.27	115.66	-127.62*	5.62	3.07	-2.55*
Non-Eurozone						
Bulgaria	258.99	130.61	-128.39*	6.53	4.28	-2.25*
Croatia	316.38	274.95	-41.43*	3.72	1.30	-2.41*
Czech Republic	98.95	49.67	-49.28*	2.56	1.08	-1.48*
Denmark	60.44	17.83	-42.61*	7.86	4.66	-3.20*
Hungary	353.35	191.21	-162.14*	6.75	6.36	-0.39*
Norway	26.58	16.23	-10.35*	5.17	3.85	-1.32*
Poland	160.49	71.75	-88.74*	5.51	2.47	-3.04*
Romania	301.57	145.09	-156.47*	3.01	1.98	-1.02*
Sweden	36.48	12.96	-23.52*	5.79	3.45	-2.34*
UK	65.54	27.81	-37.73*	7.51	4.49	-3.02*

The table reports the mean over time of the sovereign CDS spreads and bond yields for each country across the periods before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. The CDS spreads are expressed in basis points, and the bond yields are expressed in percentage terms. The third column is the difference between the two periods. The * indicates that the difference is significant at the 5% level.

internal rate of return (yield to maturity) of a bond traded at par, and it is expressed as an annualized figure. The CDS spread is expressed in basis points and represents the percentage of the CDS notional value that the protection buyer must pay, usually at quarterly frequencies, to the protection seller. Similarly, CDS spreads are expressed in annualized terms.

We use all the maturities of the CDS spreads to implement the estimation methodology.

Table 7. Descriptive Statistics by Asset.

	Average of Means			Average of Medians		
	B/OMT	A/OMT	Diff	B/OMT	A/OMT	Diff
All Samples:						
CDS	183.10	86.57	-96.53*	125.70	52.40	-73.30*
Yields	4.95	2.66	-2.29*	4.70	0.24	-4.45*
Country Groups:						
Eurozone-core						
CDS	73.23	25.84	-50.39*	70.57	26.24	-44.33*
Yields	3.82	1.40	-2.42*	3.13	0.13	-3.01*
Eurozone-periphery						
CDS	315.33	135.22	-180.11*	239.24	125.14	-114.10
Yields	5.25	2.69	-2.55*	4.84	0.27	-4.58*
Non-Eurozone						
CDS	167.88	93.81	-74.07*	129.99	60.74	-69.24*
Yields	5.44	3.39	-2.05*	5.84	0.36	-5.48*

The table reports statistics of the sovereign CDS spreads and bond yields before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date, and the relative difference, across all countries and the three groups of countries. The Average of Means is computed as the mean over time of the cross-sectional average CDS spreads and bond yields across countries. The Average of Medians is computed as the mean over time of the cross-sectional median CDS spreads and bond yields across countries. The CDS spreads are expressed in basis points, and the bond yields are expressed in percentage terms. The * indicates that the difference is significant at the 5% level.

However, throughout the paper, we focus on the 5-year maturity to show our results in the empirical analysis. We also collect data on the Euribor as a proxy of the European short-term risk-free interest rate. At longer maturities, we proxy the risk-free rate with the euro area yield curve computed exclusively on AAA-rated central government bonds. We use the Nelson-Siegel technique to bootstrap the maturities of the risk-free curve needed to obtain the present values of CDS that we use in the arbitrage strategies.

We apply a filter to the sample, excluding countries that report an excessive number of missing data points on bond yields or CDS spreads (more than 40% of the total observations for at least one maturity), thus dropping Cyprus, Luxembourg, and Malta from the sample. We also exclude Greece, which deserves a specific analysis, due to the dramatic turbulence experienced during the sample period. We drop Estonia, Latvia, and Lithuania from the sample because they changed their status from non-Eurozone to Eurozone over the sample period. We end up with a final sample of 22 countries: 12 countries belong to the Eurozone, and 10 countries are outside of the Eurozone. Throughout the analysis, we also divide the sample of the Eurozone countries in two subgroups: core and periphery. The list of countries is reported in Table 6.

Figure 9 shows that bond yields and CDS spreads significantly drop after the announcement of the OMT program for all groups of countries. In Table 6, we report data on CDS spreads and bond yields for each single country in the sample. Table 6 shows that the differences are significant at the 5% level, considering both the mean and median.

In Table 7, we report statistics on the time series of the mean and median across countries before and after July 2012. We also provide a breakdown of the mean and median by different groups of countries. We observe that bond yields and CDS spreads are generally lower for the core Eurozone countries compared to both the peripheral and non-Eurozone countries before and after the OMT announcement. Yet, the reduction in both CDS spreads and bond yields is significant at the 5% level even for the core Eurozone countries.

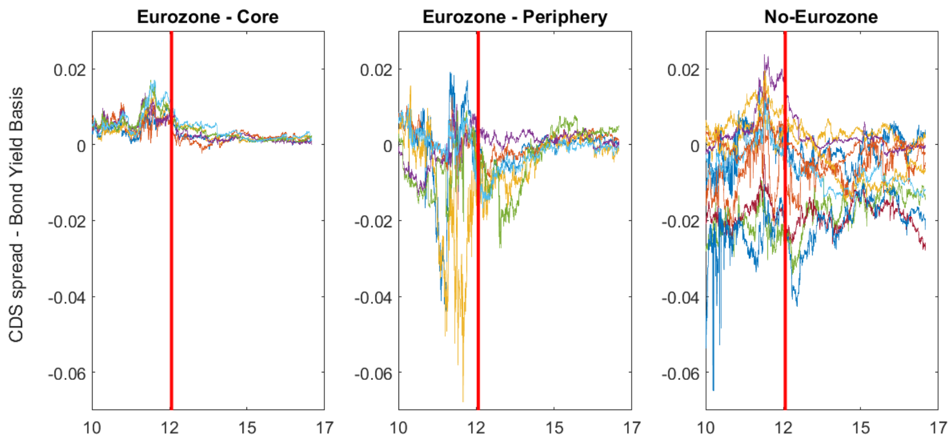
2.3. The CDS - Bond Basis

In this section, we analyze the theoretical equilibrium condition between CDS spreads and bond yields for each European country over the time series. The CDS spreads and yields on a risky bond issued by the reference entity of the CDS contract are strictly related. The CDS provides protection to the acquirer in case of default of the reference entity, while the bond pays out yields to the bondholder as long as the reference entity is able to comply with its obligations. Hull et al. (2004) have pointed out that, under a given set of assumptions, the T-year CDS spread should be equal to the T-year excess yield on a risky bond issued by the reference entity over the T-year riskless bond:

$$s = y - r, \tag{10}$$

where s is the T-year CDS spread, y is T-year yield on the risky bond, and r is the T-year yield on the riskless bond. The reason is simple: if the assumptions listed by Hull et al. (2004) hold, a portfolio including a T-year CDS and a T-year par yield bond issued by the reference entity generates cash flows equal to a T-year par yield riskless bond in all states of the world. The *basis* is the difference between the T-year CDS spread and the T-year excess yield on a

Figure 10. CDS spreads - Bond Yields Basis.



The figure reports the CDS spread - bond yield basis for each country between the 1st of January 2010 and the 1st of February 2017 at 5-year maturity for the three different groups of countries. The names of the countries belonging to each group are provided in Table 6. The basis is expressed in percentage terms. The red line is the OMT announcement date.

risky bond issued by the reference entity over the T-year riskless bond. In equilibrium, the basis must be equal to zero. Therefore, the basis is a straightforward signal to detect a relative mispricing between CDS spreads and bond yields for a given country that can be analyzed by simply using observed data.

We group our sample countries in three sub-samples: Eurozone-core (EC), Eurozone-peripheral (EP), and non-Eurozone (NZ). Figure 10 shows the dynamics of the basis for each country. The core countries have a substantially lower basis than both the peripheral and non-Eurozone countries. More importantly, the basis of both core and peripheral countries of the Eurozone converge to zero right after the OMT announcement. The non-Eurozone countries, instead, do not show the same pattern, with their basis spread around zero before and after the OMT announcement.

This result is also evident when examining the average of the absolute basis across groups of countries. Table 8 reports that the absolute basis has substantially reduced for the Eurozone countries in the second period of the time series (-65% for the Eurozone-core and -55% for the Eurozone-peripheral, respectively), while the decrease is much less pronounced for the non-Eurozone countries (-10%).

This empirical observation provides evidence on the disequilibrium between CDS spreads and bond yields for all European countries before the OMT announcement, that persists even

Table 8. Average Absolute Basis (CDS Spreads - Bond Yields).

	Euro-Core	Euro-Periphery	Non-Eurozone
Before OMT	0.63	0.78	1.05
After OMT	0.22	0.36	0.90

The table reports the average CDS spreads - bond yields basis across countries for the three groups of countries before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. The basis is expressed in basis points. Both CDS spreads and bond yields are at 5-year maturity.

after the OMT announcement for the non-Eurozone countries only. This deviation from the equilibrium condition should generate arbitrage opportunities in the market before the OMT announcement for all countries in the sample. Next, we compute the potential profits obtained by exploiting the violation of the no-arbitrage condition.

2.3.1. Arbitrage Strategy

If the basis is different from zero, an arbitrage opportunity arises in the market by trading CDS, risky bonds, and riskless assets, under the set of assumptions exhaustively explained in Hull et al. (2004). Here, we report only the most relevant assumptions that support the flow of our argument:

1. Market participants can short sovereign bonds;
2. Market participants can short the risk-free bonds (they can borrow money at the risk-free rate);
3. The cheapest-to-deliver bond option is ruled out, so that the profit is not affected by the ability of the protection seller to find a cheaper bond to deliver in case of default;
4. The recovery rate of the bond in case of default is equal to zero.

We express all the variables in monetary terms, thus computing the present value of the CDS, risk-free bond, and risky bond using continuous compounding, such that the no-arbitrage condition can be rewritten as follows:

$$P_{CDS} = P_{BY} - P_{RF},$$

where P_{CDS}, P_{BY}, P_{RF} denote the present value of the CDS, risky bond, and riskless bond, respectively. We omit the subscripts i and t to reduce notation.

The arbitrage strategy is based on the CDS spread - bond yield basis. When this relationship is not in equilibrium, there is a signal of an arbitrage opportunity arising on the market. Suppose that, for the i-th country, at time t,

$$P_{CDS} > P_{BY} - P_{RF},$$

then the arbitrageur sells the risk-free asset and purchases the CDS and risky bond issued by the CDS reference entity. The mispricing of the bond generates a positive difference that is exactly the risk-free arbitrage profit. Conversely, if

$$P_{CDS} < P_{BY} - P_{RF},$$

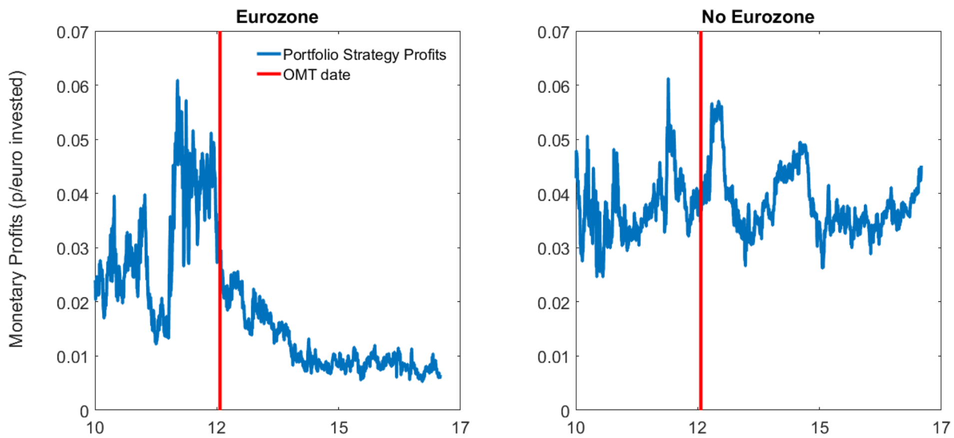
then the arbitrageur obtains the same arbitrage profit by reversing the strategy: the arbitrageur purchases the risk-free asset and sells the risky bond and CDS.

Figure 11 shows the arbitrage profits generated by a portfolio equally weighted in terms of countries. The left panel shows the profits that the arbitrageur can obtain by trading assets of the Eurozone countries, and the right panel shows the profits that the arbitrageur can obtain by trading assets of non-Eurozone countries. The profits are large and volatile before the OMT announcement in both Eurozone and non-Eurozone areas. After the announcement, however, the profits drop immediately and start to converge to zero for Eurozone countries. Instead, the riskless profits remain positive and volatile for the countries outside the Eurozone.

Table 9 reports the mean and standard deviation of the potential profits obtained with the arbitrage strategy before and after the OMT announcement and for Eurozone and non-Eurozone countries, respectively. In Table 9, we report a significant difference in the average profits between the two periods for Eurozone countries. Further, the standard deviation drops sensibly after the announcement. For the non-Eurozone area, Table 9 reports results on the means and standard deviations that are very similar across the two periods. The differences are not statistically different from zero.

Therefore, we observe that potential arbitrage profits are large and persistent for all coun-

Figure 11. Arbitrage Profits - Strategy 1.



The figure shows the arbitrage profits on an equally weighted across-country portfolio of sovereign CDSs and bonds using portfolio strategy 1 between the 1^{st} of January 2010 and the 1^{st} of February 2017. The strategy is implemented using either Eurozone sovereign CDSs and bonds only (left panel) or non-Eurozone sovereign CDSs and bonds only (right panel). The profits are expressed in monetary terms assuming a nominal value of 1 for the bonds, and the CDS price is computed as the present value of the CDS spreads expressed in percentage terms. The red line stands for the OMT announcement date.

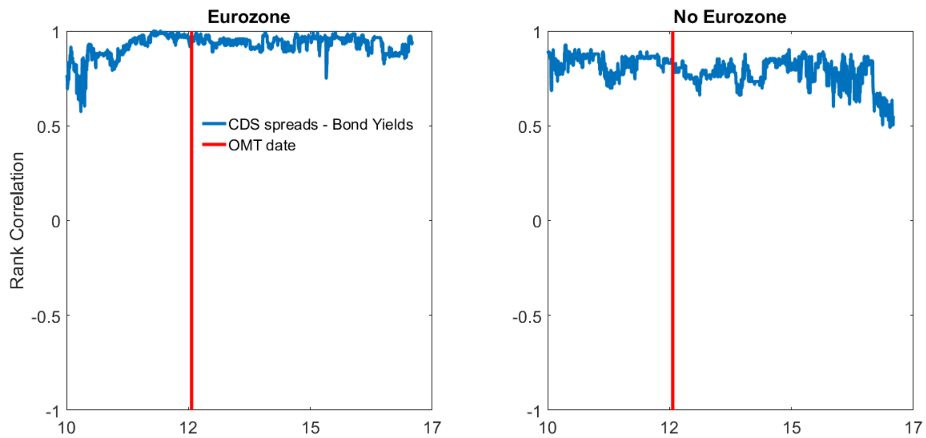
Table 9. Arbitrage Profits - Strategy 1.

Statistic:	Before OMT	After OMT	Difference
Eurozone			
Mean	0.034	0.014	-0.020*
Std. Dev.	0.012	0.005	
Non-Eurozone			
Mean	0.036	0.036	-0.000
Std. Dev.	0.006	0.006	

The table reports the mean and standard deviation of the profits on an equally weighted across-country portfolio of sovereign CDSs and bonds using portfolio strategy 1 before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. The strategy is implemented using either Eurozone sovereign CDSs and bonds only, or non-Eurozone sovereign CDSs and bonds only. The profits are expressed in monetary terms assuming a nominal value of 1 for the bonds, and the CDS price is computed as the present value of the CDS spreads expressed in percentage terms. In the last column, we report the difference across the two periods. The * indicates that the difference is significant at the 5% level.

tries before the OMT announcement and quickly converge to zero for the Eurozone countries only. This result is consistent with the evidence on the deviation from the zero-basis condition documented in the previous section.

Figure 12. CDS Spreads - Bond Yields: Cross-sectional Correlations.



The plots show the cross-sectional rank correlation between sovereign CDS spreads and bond yields at 5-year maturity between the 1^{st} of January 2010 and the 1^{st} of February 2017 across Eurozone (left panel) and non-Eurozone countries (right panel). The red line is the OMT announcement date.

2.4. Cross-Sectional Analysis

The previous section analyses the dynamics of the relationship between CDS spreads and bond yields over time for each country. However, our main target is to investigate the relationship between CDS spreads and bond yields over the cross-sectional dimension. The consistent relationship between CDS spreads and bond yields implies that a riskier country issues debt securities that pay out higher yields. Consequently, we should observe a monotonic and positive relationship between CDS spreads, the price of default risk, and bond yields. We use the Spearman correlation coefficient, which evaluates the rank correlation, to perform our analysis. If a positive and monotonic relationship between CDS spreads and bond yields exists, the rank correlation between CDS spreads and bond yields is equal to 1 over the cross section of countries for any point in time.

Figure 12 shows that the correlation between CDS spreads and observed yields is close to 1 over the time series for both groups of countries. This result implies that the relationship between CDS spreads and observed yields is monotonically positive: the riskier countries pay out higher bond yields compared to the safer countries. This result suggests that the deviation from the zero-basis condition documented in the previous section does not affect the monotonic relationship between CDS spreads and bond yields across countries. In other words, even when the CDS spreads - bond yields bases for single countries are far from zero, the cross-sectional monotonic relationship between CDS spreads and bond yields still holds.

However, the cross-sectional monotonicity between CDS spreads and bond yields is a necessary but not sufficient condition for a consistent relationship. A riskier country, in fact, should pay out bond yields that are not only higher than the bond yields paid out by a safer country, but are also sufficiently higher to compensate the bondholder properly for bearing that higher level of risk. The difference in terms of bond yields between riskier country and safer country should be large enough to be consistent with the difference in terms of the default risk priced in the corresponding CDS spreads. To perform this analysis, we examine the rank correlation over the cross section of countries between CDS spreads and *net yields*. We define net yield as the difference between the actual bond yield and the bond yield implied by the CDS spreads. The latter is the unobservable yield implied by a given level of default risk priced in the CDS spreads of the country.

The idea behind the analysis of the relationship between CDS spreads and net yields should be clear with a simple numerical example. We consider two countries, A and B, and suppose that the CDS spread of A is larger than the CDS spread of B, that is, A is riskier than B. We suppose also that the observed yields are $Y(A) = 0.1$ and $Y(B) = 0.05$, respectively. It turns out that country A is paying out a higher yield, and so the monotonicity condition between CDS spreads and bond yields is verified. However, is $Y(A)$ higher enough to compensate the bondholders for the higher risk associated with the country A? We suppose that the yields implied by the CDS spreads for the two countries are $I(A) = 0.15$ and $I(B) = 0.02$, respectively. Finally, we compute the net yields, which are $N(A) = -0.05$ and $N(B) = 0.03$.

The first consequence is that the bond of country A is overvalued; the actual yield is lower than the risk-implied yield, and the price of the bond is larger than the risk-implied price. On the other hand, the bond of country B is undervalued. Moreover, the monotonicity condition between CDS spreads and net yields is not verified. The difference between the observed bond yields across the two countries is inconsistent with the difference in terms of default risk. The riskier country A, in fact, is paying out an insufficiently higher bond yield compared to the safer country B to compensate the bondholder.

We compute the implied bond yields for each country by estimating a contingent claim model. We adopt a first-passage time model, where the issuer defaults as soon as the value of the assets crosses from above a default boundary, which is assumed to be deterministic and constant. Next, we detail the underlying model and the model estimation procedure and report our results. Then, we describe the MC simulation approach to obtain the implied bond yields, and we compare the implied bond yields with the observed bond yields. Finally, in this section, we implement the cross-sectional correlation analysis between CDS spreads and bond yields, using both the observed and implied bond yields. We also corroborate our findings with a portfolio analysis based on the arbitrage strategy described in the previous section.

2.4.1. Underlying Model

The asset value of the i-th country is described by a geometric Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, \mathcal{P})$:

$$dV_{i,t} = \mu_{V_i} V_i dt + \sigma_{V_i} V_i dW_{i,t},$$

where μ_{V_i} and σ_{V_i} are the $\mathcal{P}$-drift and diffusion constant coefficients, and $W_{i,t}$ is a standard Brownian motion under the physical probability measure $\mathcal{P}$.

We define the i-th market value of leverage as $L_{i,t} = \ln\left(\frac{F_i}{V_{i,t}}\right)$, following an arithmetic Brownian motion:

$$dL_{i,t} = \mu_{L_i} dt - \sigma_{L_i} dW_{i,t}, \tag{11}$$

where $\mu_{L_i} = -\left(\mu_{V_i} - \frac{1}{2}\sigma^2_{V_i}\right)$ is the $\mathcal{P}$-leverage drift coefficient, $\sigma_{L_i} = \sigma_{V_i}$ is the leverage diffusion component, and the minus before the diffusion component is the result of the perfect and negative correlation between the Brownian motions of asset and leverage.

In the first-passage time framework, the default occurs as soon as the asset value crosses from above a constant and deterministic barrier C_i that we assume to be below the face value of the debt at any time s with $t \leq s \leq T$, where T is the outstanding debt maturity. The default risk of the country is priced in the CDSs issued with different maturity τ_j, with j going from 1 to J, where the longest maturity τ_J matches the debt maturity T. In a CDS contract, the protection buyer pays a fixed premium at each period until either the default event occurs or the contract expires, and the protection seller is committed to buy back the defaulted bond from the buyer at its par value.

Therefore, the price of the CDS (i.e., the premium paid by the insurance buyer) is defined at the inception date of the contract to equate the expected value of the two contractual legs. By assuming the existence of a default-free money market account appreciating at a constant continuous interest rate r, and M periodical payments occurring over one year, the CDS spread γ with time to maturity τ_j priced at $t = 0$ solves the following equation:

$$\sum_{m=1}^{M} T \frac{\gamma}{M} \exp\left(-r\frac{m}{M}\right) \mathcal{E}_0^{\mathcal{Q}}[1_{t^* > \frac{m}{M}}] = \mathcal{E}_0^{\mathcal{Q}}[\exp(-rt^*)\alpha 1_{t^* < \tau_j}],$$

where t^* stands for the time of default, α is the amount paid by the protection seller to the protection buyer in case of default, and $\mathcal{E}_0^{\mathcal{Q}}$ indicates that the expectation is taken under the risk-neutral measure $\mathcal{Q}$. Therefore, $\mathcal{E}_0^{\mathcal{Q}}[1_{t^* < \tau_j}]$ is the probability that the country defaults at any time before τ_j, which is the probability that the asset value crosses from above the barrier C_i. At t, this probability is equal to the following:

$$\begin{aligned} PD_{i,t}^{\mathcal{Q}}(\tau_j) = \Phi &\left(\frac{K_i + L_{i,t} - \left(r - \frac{1}{2}\sigma_{L_i}^2\right)(\tau_j - t)}{\sigma_{L_i}\sqrt{(\tau_j - t)}} \right) \\ &+ \exp\left((K_i + L_{i,t}) \left(\frac{2r}{\sigma_{L_i}^2} - 1 \right) \right) \Phi \left(\frac{(K_i + L_{i,t}) + \left(r - \frac{1}{2}\sigma_{L_i}^2\right)(\tau_j - t)}{\sigma_{L_i}\sqrt{(\tau_j - t)}} \right), \end{aligned} \tag{12}$$

if $\tau_j < T$, otherwise

$$\begin{aligned} PD_{i,t}^{\mathcal{Q}}(\tau_J) = 1 - \Phi &\left(\frac{-L_{i,t} + \left(r - \frac{1}{2}\sigma_{L_i}^2\right)(\tau_J - t)}{\sigma_{L_i}\sqrt{(\tau_J - t)}} \right) \\ &+ \exp\left((K_i + L_{i,t}) \left(\frac{2r}{\sigma_{L_i}^2} - 1 \right) \right) \Phi \left(\frac{(2K_i + L_{i,t}) + \left(r - \frac{1}{2}\sigma_{L_i}^2\right)(\tau_J - t)}{\sigma_{L_i}\sqrt{(\tau_J - t)}} \right), \end{aligned} \tag{13}$$

as $\tau_J = T$. Equation 12 defines the early bankruptcy risk, and equation 13 defines the probability that the country is not able to pay back the outstanding debt F_i at time T, even though the asset value never crossed the default boundary. In the equations (12) and (13), Φ stands for the cumulative distribution function of a standard normal variable, and $K_i = \ln\left(\frac{C_i}{F_i}\right)$. Since the default barrier is below the face value of the debt, K_i assumes only negative values. The larger the absolute value of K_i is, the larger the distance is between the face value of the debt F_i and the default barrier C_i.

2.4.2. Estimation Methodology

We adopt the following procedure to estimate the model. First, we reconstruct the unobservable dynamics of the leverage, defined as the debt-to-asset ratio, for each country by performing a nonlinear Kalman filter using the CDS spreads as observable variables. The Kalman filter allows to retrieve the dynamics of a latent variable by exploiting the relationship between observable and unobservable variables. The relationship between the observed and unobserved variables forms the *measurement equation*, while the evolution over time of the latent variable is called the *transition equation*. We estimate the model parameters by adopting a quasi-maximum likelihood algorithm, in conjunction with the Kalman filter. Details of the estimation methodology are provided in Appendix **??**.

We formulate our problem with a state-space model, where the measurement equations are the equations (12) and (13). The noise terms associated with the CDS implied-default probability for different times to maturity τ_j are assumed to be uncorrelated and to have equal variance:

$$PD^{\mathbb{Q}}_{i,t}(\tau_j) = g\left(L_{i,t}; K_i, \sigma_{L_i}\right) + \epsilon_{i,t}(\tau_j), [j = 1, 5, 10],$$

where the time to maturity is expressed in years, and $j = 10$ stands for the maturity T of the outstanding debt F_i (i.e., 10 years). The function g defines the nonlinear relationships between the observable and latent variables, and $\epsilon_{i,t}(\tau_j)$ is the measurement noise associated with the time horizon j. The measurement noises, for each country i, are assumed to follow a multivariate normal distribution with zero mean and a diagonal covariance matrix R_i. We assume a homoscedastic covariance matrix, which varies by country.

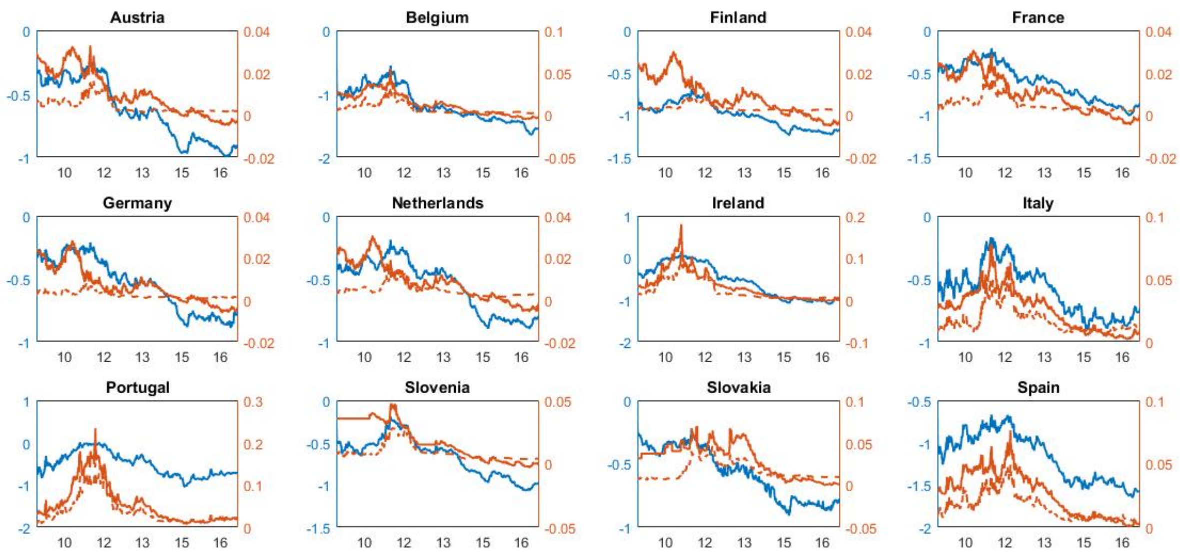
The transition equation describes the evolution of the leverage. It follows from the discretization of the stochastic process defined in (11):

$$L_{i,t+\delta t} = L_{i,t} + \mu_{L_i}\delta t + \eta_{i,t+\delta t},$$

where $\eta_{i,t+\delta t} = \sigma_{L_i}(W_{i,t} - W_{i,t+\delta t}) \backsim \mathcal{N}(0, Q_i)$ is the transition error, and $Q_i = \sigma^2_{L_i}\delta t$.

The dynamics of $L_{i,t}$ and the parameters of the model, such as μ_{L_i}, σ_{L_i}, and K_i, are esti-

Figure 13. Leverage, CDS Spreads, and Bond Yields: Eurozone Countries.



The plots show the leverage of Eurozone countries (blue line), as defined in equation (11), reconstructed with the nonlinear Kalman filter using the 5-year CDS spreads as the observable variable between the 1^{st} of January 2010 and the 1^{st} of February 2017. Moreover, we report the 5-year CDS spreads (dashed line) and the 5-year bond yields (red line) expressed in percentage terms.

mated by performing a nonlinear Kalman filter in conjunction with a quasi-maximum likelihood algorithm.

Figure 13 provides an idea of the estimation results. In Figure 13, we compare the reconstructed dynamics of the leverage and the observed dynamics of the 5-year CDS spreads and bond yields for our sample countries. The dynamics of both CDS spreads and bond yields are in line with the dynamics of the leverage. When CDS spreads and bond yields approach very low values, such as in the last part of the time series, we estimate a leverage moving far away from zero, toward negative values.

2.4.3. Implied and Observed Bond Yields

We obtain the implied bond yields for each point in time and for each country by performing an MC simulation analysis. In particular, for each point in time t, and each country i, we simulate the dynamics of the leverage for a time interval going from t to $t + M * 360$, where M is the maturity of the bond expressed in years. The leverage of a country is simulated using the equation (11), where dt is a one-day step. The parameters of the stochastic process are the estimates obtained in the previous step, and we use the estimated leverage at time t as the starting point of the simulated dynamics. We generate $M * 360$ normally distributed random numbers for each country to simulate the daily increment of the Brownian motion, thus finally obtaining the simulated dynamics of the leverage of length $M * 360$.

Then, we use the condition of default implied by the model. The country defaults if $V_{i,t} < C_i$, which corresponds to $L_{i,t} > (-K_i)$. Therefore, if the simulated leverage of the country is above $-K_i$, at least for one point in time over the time horizon, we impose that the bond defaults, and the t-value of the bond is zero. Otherwise, the t-value of the bond is equal to the risk-free discount factor using the risk-free rate at time t. We compute the bond price for each time t as an average over 10,000 simulations, and the corresponding yield by simple inversion. If we define B the price of the bond obtained with the MC simulations, then the implied yield Y is equal to the following:

$$Y = \log \left(\frac{1/B}{M * 360} \right).$$

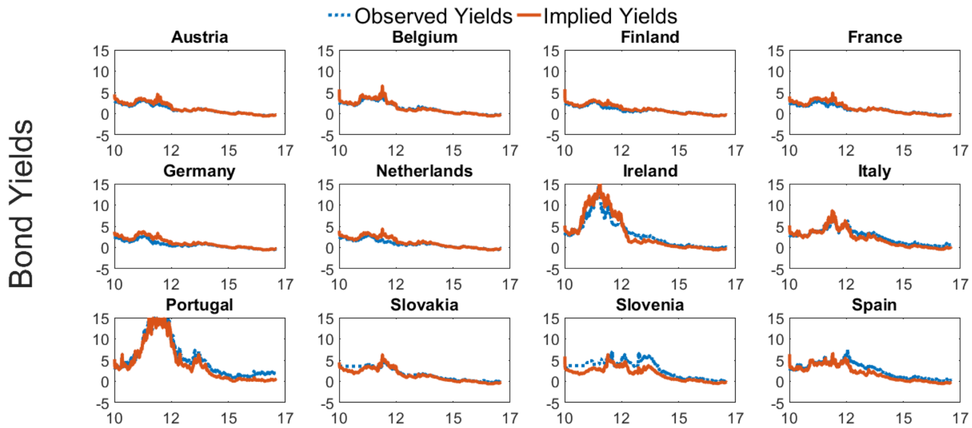
The difference between observed and implied yields should be zero for each country and each point in time, if the observed risky yields of a country are consistent with the default risk priced in the CDS spreads. Indeed, the maintained assumption behind this statement is that the model-implied yields are well estimated, and the model is able to fully capture whatever drives the relationship between default risk and bond prices. With these caveats in mind, we compare observed and implied yields for each country over the sample time-series.

Figures 14 and 15 show that the implied yields are generally closer to the observed yields for the Eurozone countries compared to those of the non-Eurozone countries. Within the Eurozone group, we obtain implied yields that are very close to the observed yields for the core countries in the second part of the time series. At the opposite, the non-Eurozone countries show a persistent distance between implied and observed yields over the entire time series.

2.4.4. Correlation Analysis

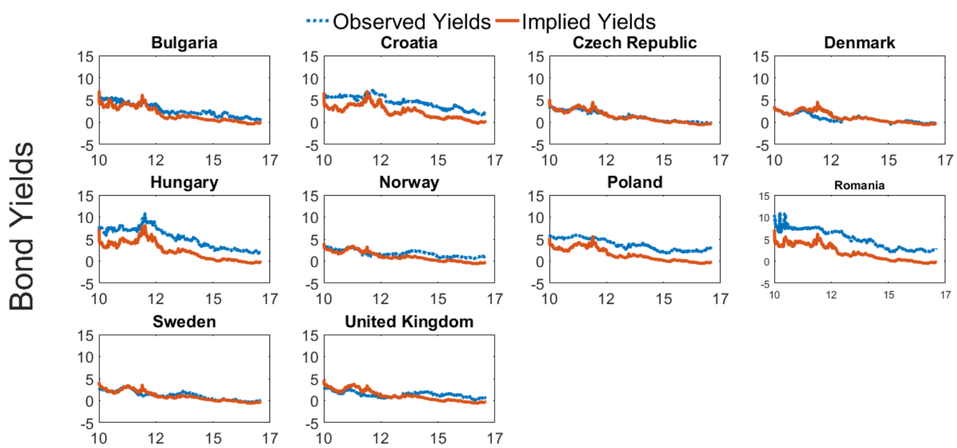
We now focus on the main result of the paper: the cross-sectional correlation for each point in time between CDS spreads and net bond yields across our sample countries. For each point in time, we compute the Spearman correlation coefficient between CDS spreads and bond yields

Figure 14. Implied Versus Observed Yields: Eurozone.



The plots show the observed (blue line) and the model-implied (red line) bond yields at 5-year maturity for Eurozone countries between the 1st of January 2010 and the 1st of February 2017. Bond yields are expressed in percentage terms. We compute the model-implied yields using the estimation methodology described in Section 4.

Figure 15. Implied Versus Observed Yields: Non-Eurozone.

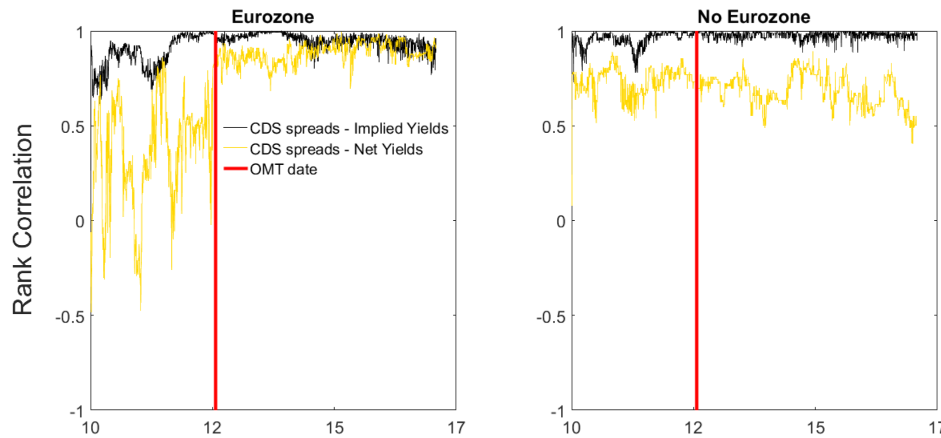


The plots show the observed (blue line) and model-implied (red line) bond yields at 5-year maturity for the non-Eurozone countries, between the 1st of January 2010 and the 1st of February 2017. Bond yields are expressed in percentage terms. We compute the model-implied yields using the estimation methodology described in Section 4.

using both implied and net yields across the sample countries. The next figure graphically represents the main result of the paper. Figure 16 shows the dynamics of the cross-sectional correlations between the 5-year CDS spreads and the implied bond yields, and between the 5-year CDS spreads and the net yields for the Eurozone and non-Eurozone countries, respectively.

The plots show that the correlation between CDS spreads and implied yields is close to 1 over the entire time series and for both groups of countries. This result is natural since the

Figure 16. CDS Spreads - Net Yields: Cross-sectional Correlations.



The top plots show the cross-sectional rank correlation between sovereign CDS spreads and model-implied yields (black line), and between sovereign CDS spreads and net yields (yellow line) at 5-year maturity across Eurozone (left panel) and non-Eurozone countries (right panel) between the 1^{st} of January 2010 and the 1^{st} of February 2017. We compute the model-implied yields using the estimation methodology described in Section 4, and we compute the net yields as the difference between observed and model-implied yields.

implied yields are estimated using the CDS spreads. Moreover, the correlation is not perfectly equal to 1, as the model is subject to error. The yields obtained by MC simulations are also subject to error.

More importantly, the yellow line in Figure 16 shows the dynamics of the cross-sectional correlation between CDS spreads and net yields. The correlation randomly moves around zero for the Eurozone countries before the OMT announcement, then approaches 1 right after the OMT announcement, and remains stable afterwards. It turns out that, before the OMT announcement, the differences between sovereign bond yields across the Eurozone countries are not consistent with the cross-sectional differences in terms of default risk, and this consistency is restored right after the announcement.

This result is even more interesting if we compare Eurozone and non-Eurozone countries. In fact, non-Eurozone countries do not show any change over time in the cross-sectional correlation between CDS spreads and net yields. The correlation is stable over the entire time series, never approaching 1.

Table 10 reports the average correlation between CDS spreads and the different measures of bond yields across countries in each group for the two periods (i.e., before and after the OMT announcement). The average correlation between CDS spreads and both actual and

Table 10. Correlation CDS spreads - Bond Yields.

	Eurozone			Non-Eurozone		
	Obs Yields	Imp Yields	Net Yields	Obs Yields	Imp Yields	Net Yields
Before OMT	0.883	0.938	0.367	0.956	0.895	0.737
	(0.003)	(0.000)	(0.275)	(0.000)	(0.002)	(0.026)
After OMT	0.951	0.927	0.885	0.978	0.818	0.683
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.044)

The table reports the cross-sectional rank correlation between sovereign CDS spreads and observed bond yields, between sovereign CDS spreads and model-implied bond yields, and between sovereign CDS spreads and net yields at 5-year maturity across Eurozone and non-Eurozone countries before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. We compute the model-implied yields using the estimation methodology described in Section 4, and we compute the net yields as the difference between observed and model-implied yields. We report p -values in parentheses.

implied yields is very close to 1 for both groups in each period. The average correlation across Eurozone countries between CDS spreads and net yields is more than double in the second period compared to the first period. This correlation, instead, is very similar across the two periods for non-Eurozone countries and is even lower after the OMT announcement.

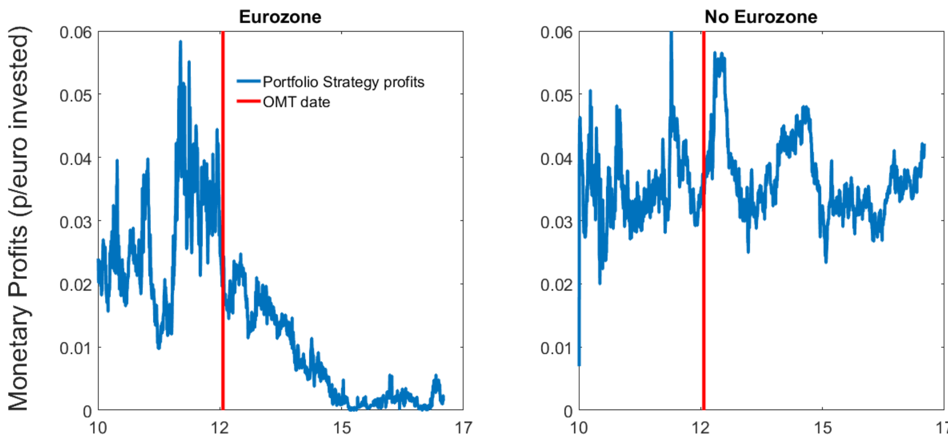
2.4.5. Net Yields and Arbitrage Strategy

To corroborate our findings, we construct long-short portfolio strategies based on the net yields to exploit the deviation of the observed yields from the model-implied yields that are consistent with the default risk priced in the CDS spreads. For each point in time, we classify the sample countries as *undervalued* when the net yield is positive and as *overvalued* when the net yield is negative.

If the i-th country is undervalued, the arbitrageur sells the risk-free asset and purchases the CDS and risky bond issued by the CDS reference entity. Otherwise, if the i-th country is overvalued, the arbitrageur purchases the risk-free asset and sells the risky bond and CDS to obtain the risk-free profit.

The implementation of this strategy works exactly as for the arbitrage strategy described in Section 2.3. The difference between the two strategies is only given by the signal of the riskless profit opportunity arising on the market. In the first strategy, the signal is the zero-basis condition. In this strategy, the signal is the distance between the observed and implied yields.

Figure 17. Arbitrage Profits - Strategy 2.



The figure shows the arbitrage profits on an equally weighted across-country portfolio of sovereign CDSs and bonds using portfolio strategy 2 between the 1^{st} of January 2010 and the 1^{st} of February 2017. The strategy is implemented using either Eurozone sovereign CDSs and bonds only (left panel), or non-Eurozone sovereign CDSs and bonds only (right panel). The profits are expressed in monetary terms assuming a nominal value of 1 for the bonds, and the CDS price is computed as the present value of the CDS spreads expressed in percentage terms. The red line stands for the OMT announcement date.

In Figure 17, we compare the profits obtained on an equally weighted portfolio across countries by trading assets of Eurozone and non-Eurozone countries, respectively, using the arbitrage strategy described in this section. The profits reported in Figure 17 are very similar to those presented in Figure 11. Arbitrage opportunities are persistent for both groups of countries before the OMT announcement; however, they quickly converge to zero for the Eurozone countries after the announcement of the OMT program.

Table 11 reports the mean and standard deviation of the arbitrage profits before and after the OMT announcement for the Eurozone and non-Eurozone countries, respectively. Table 11 shows a pronounced difference in the average profits between the two periods for the Eurozone countries. Further, the standard deviation drops sensibly after the OMT announcement. Such numbers indicate that, after the OMT announcement, the arbitrage opportunities are approximately zero. Instead, for the non-Eurozone area, Table 11 reports similar figures for the mean and standard deviation between the two periods. The differences between the two periods, in fact, are not statistically different from zero.

Table 11. Arbitrage Profits - Strategy 2.

Statistic:	Before OMT	After OMT	Difference
Eurozone			
Mean	0.029	0.003	-0.027*
Std. Dev.	0.012	0.005	
Non-Eurozone			
Mean	0.020	0.012	-0.008
Std. Dev.	0.013	0.017	

The table reports the mean and standard deviation of the profits on an equally weighted across-country portfolio of sovereign CDSs and bonds using portfolio strategy 2 before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. The strategy is implemented using either Eurozone sovereign CDSs and bonds only, or non-Eurozone sovereign CDSs and bonds only. The profits are expressed in monetary terms assuming a nominal value of 1 for the bonds, and the CDS price is computed as the present value of the CDS spreads expressed in percentage terms. In the last column, we report the difference across the two periods. The * indicates that the difference is significant at the 5% level.

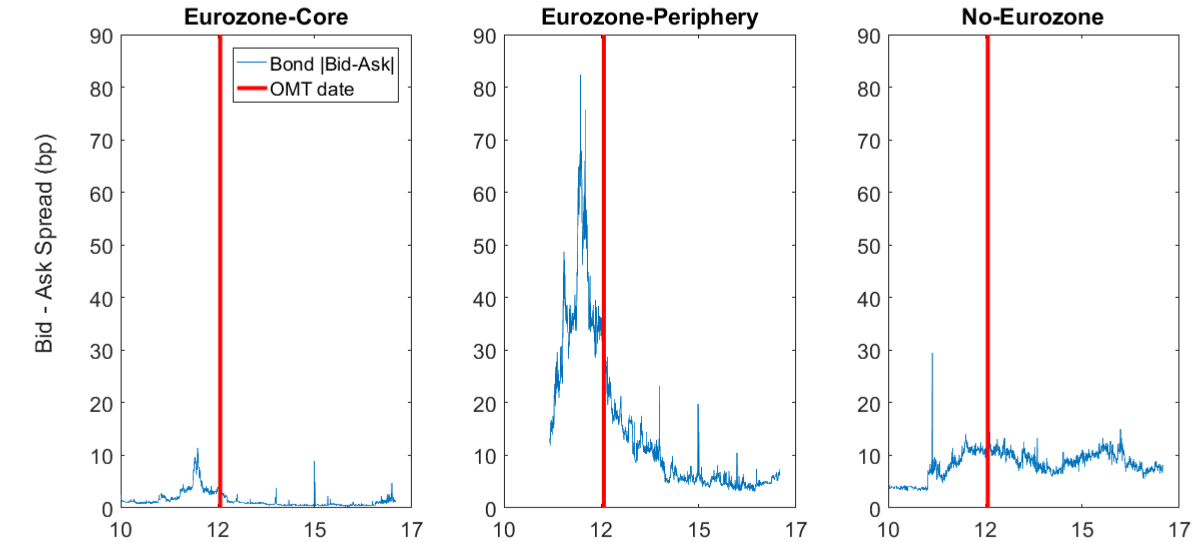
2.5. Arbitrage and Transaction Costs

Arbitrage opportunities can persist in the market if the riskless profits are insufficient to cover the costs to implement the arbitrage strategy. The idea is that arbitrageurs enter the market only if the arbitrage strategy still generates profits once the transaction costs have been paid. Therefore, we control for transaction costs in two ways. First, we use bid and ask prices of sovereign bonds and CDS for our sample countries to compute the performance of the two arbitrage strategies, and we check whether the strategies are still profitable. Second, we estimate the threshold beyond which the riskless trading gains become sufficiently profitable.

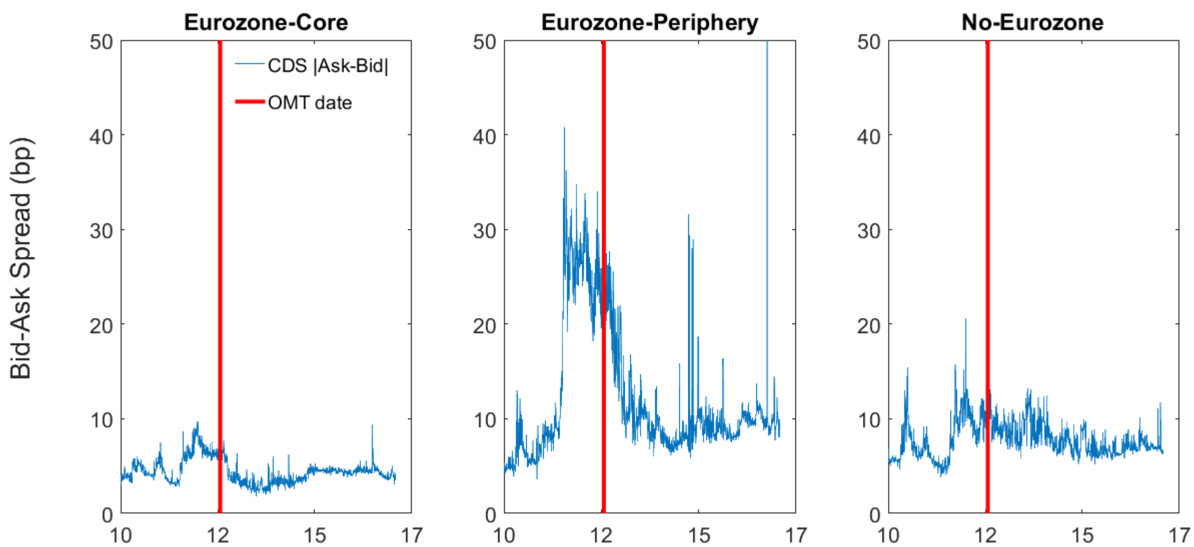
2.5.1. Bid-Ask Prices

In this section, we use bid and ask prices for CDS and bonds between January 2010 and February 2017. We compute the bid-ask spread for our sample countries at each time t for both assets. Figure 18 shows the average bid-ask spread across countries on the 5-year maturity bond (top panel) for Eurozone-core, Eurozone-periphery, and non-Eurozone countries, respectively. The plots show that, for both the Eurozone groups, the average spread has a spike during the sovereign debt crisis followed by a strong and persistent reduction over the subsequent years. Albeit the two Eurozone groups show similar dynamics, the average spread across the peripheral countries is higher in magnitude compared to the average spread across the core countries, in particular before the OMT announcement. The average bid-ask spread across the non-Eurozone

Figure 18. Bid-Ask Spreads.



The figure shows the average 5-year maturity bond bid-ask spread across countries for the three groups of countries between the 1^{st} of January 2010 and the 1^{st} of February 2017. The spreads are expressed in basis points. The red line stands for the OMT announcement date.



The figure shows the average 5-year maturity CDS bid-ask spread across countries for the three groups of countries between the 1^{st} of January 2010 and the 1^{st} of February 2017. The spreads are expressed in basis points. The red line stands for the OMT announcement date.

countries, instead, is substantially persistent across the two periods. Similar discussion applies to the CDS (bottom panel). However, the reduction in the average bid-ask spread for both the Eurozone groups occurs right after the OMT announcement. In Table 12, we report the

Table 12. Bid-Ask Spreads.

	Bond		
	Euro-Core	Euro-Periphery	Non-Eurozone
Before OMT	2.55	22.83	7.02
After OMT	0.87	8.46	9.10
	CDS		
	Euro-Core	Euro-Periphery	Non-Eurozone
Before OMT	5.20	14.95	7.82
After OMT	4.01	10.56	7.82

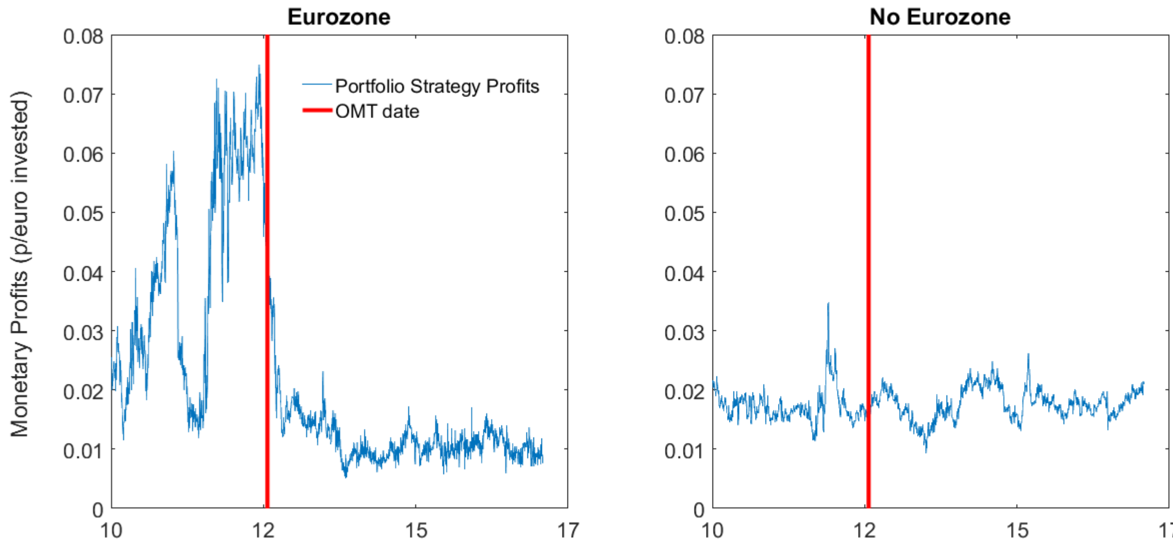
The table reports the average bid-ask spread for bonds and CDSs across countries for the three groups of countries before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. The spreads are expressed in basis points. Both CDSs and bonds are at 5-year maturity.

average bid-ask spread across countries for bonds (top panel) and CDS (bottom panel) for the three groups of countries before and after the OMT announcement date.

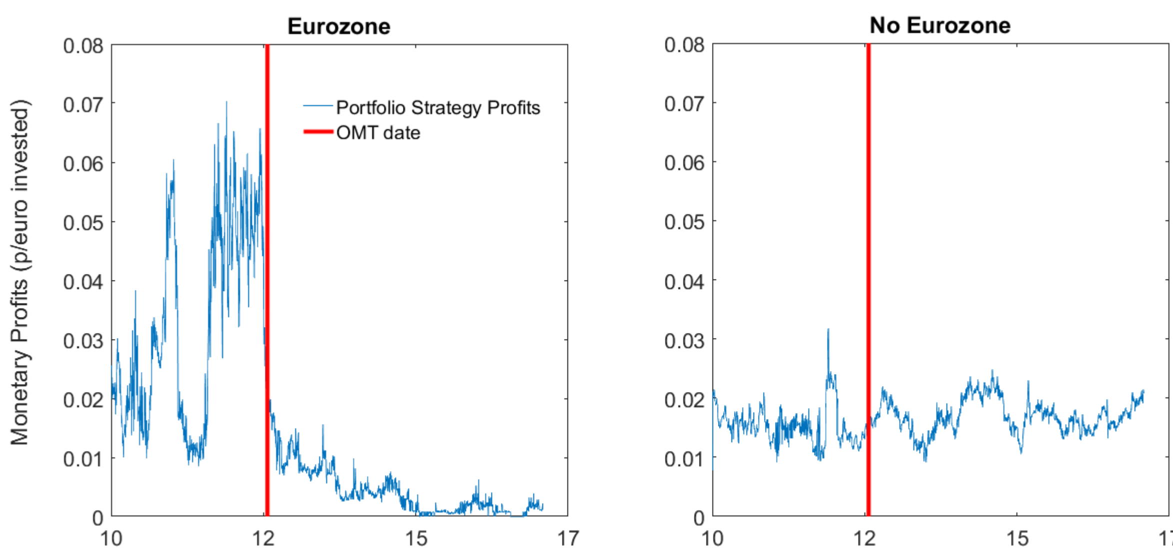
Using data on bid-ask spreads, we compute the riskless profits generated by the two arbitrage strategies described in the previous sections. By doing that, we use the actual price that an arbitrageur should pay to implement the strategy: the bid price, when the arbitrageur sells the asset, and the ask price, when the arbitrageur buys the asset. In Figure 19, we plot the profits generated by arbitrage strategy 1 (top panel) and arbitrage strategy 2 (bottom panel), described in Sections 2.3 and 2.4, respectively. The plots show that the arbitrage profits computed with bid and ask prices have very similar patterns to the arbitrage profits computed using only one price for each asset.[9]

[9] Differences in magnitude between the profits computed in this section compared to those calculated in Sections 2.3 and 2.4 are due to the different data sources of the CDS prices. The CDS market is an over-the-counter market; therefore, different data providers may report different prices. We use Datastream-Thomson Reuters in Sections 2.3 and 2.4, and we use Bloomberg in this section.

Figure 19. Arbitrage Profits.



The figure shows the arbitrage profits on an equally weighted across-country portfolio of sovereign CDSs and bonds between the 1^{st} of January 2010 and the 1^{st} of February 2017 using portfolio strategy 1, for which each transaction occurs at the quoted bid or ask price. The strategy is implemented using either Eurozone sovereign CDSs and bonds only (left panel), or non-Eurozone sovereign CDSs and bonds only (right panel). The profits are expressed in monetary terms assuming a nominal value of 1 for the bonds, and the CDS price is computed as the present value of the CDS spreads expressed in percentage terms. The red line stands for the OMT announcement date.



The figure shows the arbitrage profits on an equally weighted across-country portfolio of sovereign CDSs and bonds between the 1^{st} of January 2010 and the 1^{st} of February 2017 using portfolio strategy 2, for which each transaction occurs at the quoted bid or ask price. The strategy is implemented using either Eurozone sovereign CDSs and bonds only (left panel), or non-Eurozone sovereign CDSs and bonds only (right panel). The profits are expressed in monetary terms assuming a nominal value of 1 for the bonds, and the CDS price is computed as the present value of the CDS spreads expressed in percentage terms. The red line stands for the OMT announcement date.

The persistence of riskless profits across Eurozone countries before the OMT announcement date and across non-Eurozone countries before and after the OMT announcement date suggests that there are costs to implement the strategies not explained by the bid-ask spread. These costs, which prevent investors from exploiting the mispricing in the sovereign bond market, are not observable. Therefore, in the next section, we estimate such costs.

2.5.2. Estimating Transaction Costs

For each country, we estimate a vector error correction model (VECM) that includes CDS spreads and bond yields in excess of the risk-free rate adjusted for the nonlinearity due to the transaction cost threshold (TVECM). In a linear VECM, any deviation from the long-run equilibrium (zero-basis condition) would trigger trades leading the market back to the equilibrium. It turns out that, in absence of frictions, such as transaction costs, we should observe a basis moving around zero. Instead, when frictions arise in the market, we expect to observe a persistent deviation from the equilibrium. In particular, with non-zero transaction costs, the deviation should persist as long as the magnitude of the deviation is below a given threshold, which introduces nonlinearity in the error correction model.

Following Gyntelberg et al. (2017), we model CDS spreads and excess risky bond yields in vector form as follows:

$$\Delta y_t = [\lambda^L ec_{t-1} + \Gamma^L(\ell)\Delta y_t] d_{Lt}(\beta, \theta) + [\lambda^U ec_{t-1} + \Gamma^U(\ell)\Delta y_t] d_{Ut}(\beta, \theta) + \epsilon_t,$$

where $ec_{t-1} = CDS_{t-1} - \beta_0 - \beta_1 ER_{t-1}$ is the error correction term with ER standing for the excess risky bond yield, $\Gamma(\ell)\Delta y_t$ is the VAR term of order ℓ, and ϵ_t are white noise shocks. Moreover, d_{Lt} and d_{Ut} are defined as follows:

$$d_{Lt} = I(ec_{t-1} \leq \theta)$$

$$d_{Ut} = I(ec_{t-1} > \theta),$$

where I is an indicator function, and θ is the threshold to be estimated. We force β_1 to be equal to 1, and we estimate β_0. An estimate of β_0 different from zero stands for a persistent non-zero basis. Therefore, the average transactions costs faced by the arbitrageurs are given by $\theta + \beta_0$. We estimate the model following the approach of Hansen and Seo (2002), who estimated

Table 13. Average Transaction Costs.

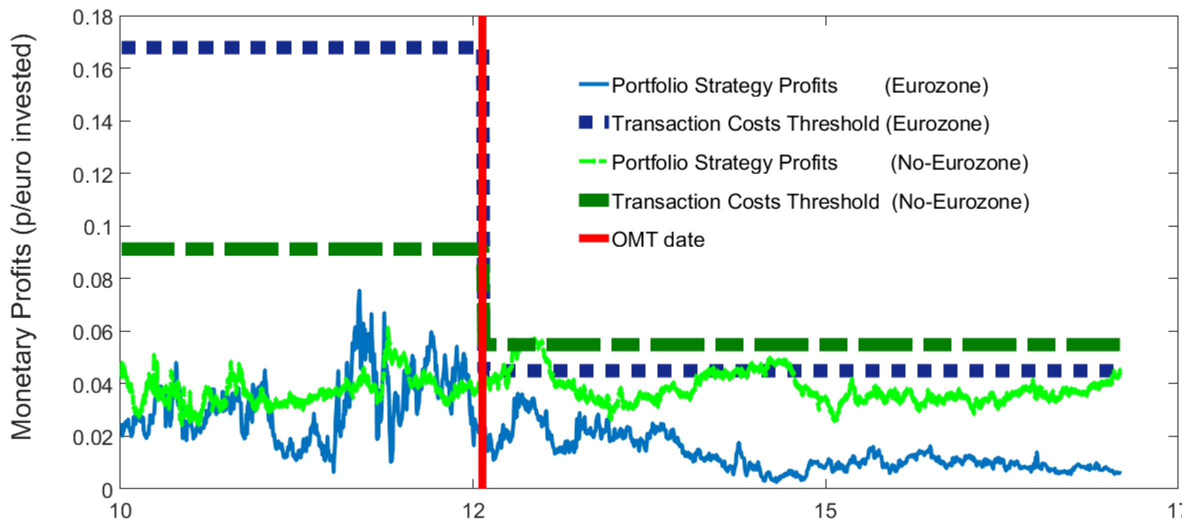
Countries	Before OMT	After OMT	% Diff
Austria	0.0152	0.0025	-83
Belgium	0.0468	0.0073	-84
Finland	0.0105	0.0036	-66
France	0.0131	0.0060	-54
Germany	0.0133	0.0013	-90
Netherlands	0.0098	0.0059	-40
Average Core	**0.0181**	**0.0044**	**-75**
Ireland	0.2466	0.0312	-87
Italy	0.1255	0.0508	-59
Portugal	0.4179	0.0897	-78
Slovakia	0.0723	0.0131	-81
Slovenia	0.0321	0.0368	+14
Spain	0.1123	0.0479	-57
Average Peripheral	**0.1678**	**0.0449**	**-73**
Average Eurozone	**0.0930**	**0.0247**	**-73**
Bulgaria	0.1132	0.0542	-52
Croatia	0.1958	0.1161	-40
Czech Republic	0.0303	0.0084	-72
Denmark	0.0152	0.0044	-71
Hungary	0.2156	0.0975	-54
Norway	0.0111	0.0407	+268
Poland	0.1170	0.0789	-32
Romania	0.2012	0.1034	-48
Sweden	0.0070	0.0109	+56
United Kingdom	0.0057	0.0340	+495
Average Non-Eurozone	**0.0912**	**0.0548**	**-39**
All Countries	**0.0922**	**0.0384**	**-58**

The table reports the average transaction costs ($\theta + \beta_0$) for each country before (January 1, 2010 - July 25, 2012) and after (July 26, 2012 - February 1, 2017) the OMT announcement date. The average transaction costs are expressed in percentage terms. The last column reports the variation in percentage terms across the two time periods for each country. We also report the mean across groups of countries (Eurozone core, Eurozone-peripheral, and non-Eurozone).

a two-regime TVECM using a maximum likelihood algorithm. We also estimate the model for two periods: before and after the OMT announcement. As result, we obtain an estimate of the average transaction costs for each country and for each period. Table 13 reports our results.[10]

[10]The statistical significance of the thresholds is evaluated following the approach of Hansen and Seo (2002), who calculate standard errors by means of both parametric and non-parametric bootstrap analysis. Gyntelberg et al. (2017) provide a short description of the two alternative bootstrap procedures and the decision criterion for the threshold statistical significance.

Figure 20. Arbitrage Profits and Transaction Costs.



The figure shows the profits generated by arbitrage strategy 1 for Eurozone countries (blue line) and non-Eurozone countries (green line) against the average transaction costs across Eurozone countries (blue dotted line) and non-Eurozone countries (green dotted line), respectively, between the 1^{st} of January 2010 and the 1^{st} of February 2017. The red line stands for the OMT announcement date.

We find that, in general, the key threshold is substantially higher before the OMT announcement (the average transaction costs across countries is 922 bp) compared to the second period (384 bp). This result is consistent with the findings of Gyntelberg et al. (2017), who estimate a threshold more than twice higher during the Eurozone sovereign debt crisis compared to that of the pre-crisis period. Moreover, we find that the drop in the average transaction costs across the two periods is much more pronounced for Eurozone countries (from 930 bp to 247 bp), and in particular for the peripheral countries (from 1678 bp to 449 bp), with respect to non-Eurozone countries (from 912 bp to 548 bp).

Next, we compare the estimated transaction costs with the potential arbitrage profits across groups of countries by splitting our sample in two groups (Eurozone and non-Eurozone). However, our results hold if we again split the Eurozone countries in two sub-samples (core and peripheral). The plot in Figure 20 offers a straightforward interpretation of our results.

Before the OMT announcement, we estimate similar average transaction costs across groups of countries, and transaction costs are above the arbitrage profits for both groups of countries. Therefore, the arbitrageurs do not have an incentive to intervene and clear the arbitrage opportunities, as the riskless profits are not even sufficient to cover the costs to implement the strategy. Consequently, over this period, there is a persistent deviation from the zero-basis equilibrium condition.

After the OMT announcement, we estimate a strong reduction of the average transaction costs across the Eurozone countries. Then, the arbitrageurs find it profitable to enter the market and take advantage of the deviation from the equilibrium condition. Consequently, the arbitrage profits quickly converge to zero. In other words, the lower transaction costs have created the condition for the traders to profit from the arbitrage opportunities generated by the relative mispricing between CDS spreads and bond yields, thus leading the sovereign debt market back to equilibrium (zero basis).

On the other hand, this condition does not occur for the non-Eurozone countries. The reduction in the threshold, in fact, is not enough to create the condition for the traders to clear the arbitrage opportunities. Therefore, we observe a persistent mispricing between CDS spreads and bond yields even after the OMT announcement.

2.6. Conclusion

In the paper, we conduct an empirical investigation of the relationship between sovereign CDS spreads and sovereign bond yields. In summary, we document that, after the announcement of the OMT program by the ECB, the consistent cross-sectional relationship between CDS spreads and bond yields across Eurozone countries is restored.

We document a deviation from the no-arbitrage theoretical relationship between CDS spreads and bond yields over the time series for our sample countries. However, we show that such deviation does not affect the monotonicity in the cross-sectional relationship between CDS spreads and bond yields. Then, we show that the violation of the zero-basis equilibrium condition generates instead inconsistency in the cross section of the bond yields across countries with respect to the differences in terms of default risk priced in the CDS spreads. The differences across countries in terms of default risk, which is priced in the CDS spreads, are not consistently priced in the cross section of the bond yields. This inconsistent cross-sectional relationship vanishes after the OMT announcement for the Eurozone countries only.

Further investigation should focus on the big challenge of isolating the long-term effects of the OMT program on the relative pricing of the sovereign credit securities to prove and identify a robust causal relationship. The main issue in a sovereign analysis is created by the unavoidable interaction between external and internal factors that are simultaneously at work. With this paper, we want to highlight crucial evidence for the analysis of the risk-return relationship, linking this cornerstone of the financial theory with macro-economic and monetary events, awaiting further and deeper research.

3. Default Risk Premium in Credit and Equity Markets

3.1. Introduction

The default of a company may produce substantial losses for the investors holding securities issued by the company, and therefore investors seek a premium for bearing the default risk of the company. The premium arises from the difference between the actual default risk of a company and the valuation of the default risk implied by the pricing of the securities issued by the company, such as equity, bonds, or derivatives. We refer to the former as the real-world default probability, that is generally measured by using the company rating provided by rating agencies, and to the latter as the risk-neutral default probability, backed-out by the prices of the securities.

In this paper, we study the relation between the default risk premium and both stock prices and credit default swaps (CDS) spreads. We show that an increase in the default risk premium can be associated, at the same time, to either an increase in the stock price and a decrease in the CDS spread, or to a decrease in the stock price and an increase in the CDS spread.

We use CDS spreads following previous papers that show the better performance of the CDS market to provide a reliable proxy of credit risk with respect to bonds or loans spreads (Blanco et al. (2005), Almeida and Philippon (2007)). The CDS spreads are less affected by other sources of risk, in particular by liquidity risk (Longstaff et al. (2005)). The CDS spread is the price of an insurance against the default of the company. The higher is the default risk, the higher should be the price of the insurance.

The intuition behind our results is simple. We define the default risk premium as the ratio between the risk-neutral default probability and the real-world default probability. Then, the premium may increase either because the risk-neutral default probability increases more than the real-world default probability, or because the real-world default probability decreases more than the risk-neutral default probability. In the first case, the stock depreciates because of the higher default risk of the company, and the price of the CDS increases as an insurance against the default of the company is more costly. In the second case, the stock price moves up since the firm is safer, and the CDS depreciates as the demand for hedging against default is lower. On the other hand, the default risk premium may decrease either because the risk-neutral default probability decreases more than the real-world default probability, or because the real-world default probability increases more than the risk-neutral default probability. In the first case, the stock price moves up and the price of the CDS decreases. In the second case, the stock depreciates and the price of the CDS increases.

We show that the relation between the default risk premium and the pricing of the securities depends on whether the firm generates an expected return on the market value of the assets higher than the risk-free rate. We argue that the result does not depend on the particular choice of the model. In the Appendix, we prove it in a very simple and general pricing model, that features the pricing of a stock and an insurance contract by using both the real-world and the risk-neutral probabilities.

In the paper, we adopt a contingent claim model, where the default occurs as soon as the asset value of the firm crosses from above a default barrier. This framework, known as first-time passage model, is an extension of the seminal model of Merton (1974), where the default may occur only at the maturity of the debt. We estimate the model with a non-linear Kalman filter in conjunction with quasi-maximum likelihood, by using equity prices and CDS spreads. We first report that our estimates well fit the real data. Buying and selling equity shares and CDS, according to the one-step ahead forecasts based on our estimation of the firms' asset and debt, generates highly significant performance. Moreover, we show that the first-time passage model fits better the real data than the Merton (1974) model. Buying and selling stocks and CDS, according to the one-step ahead forecasts based on the first-time passage model estimation, generates larger and more significant performance compared to the forecasts based on the Merton (1974) model estimation.

In a contingent claim model, the securities issued by the company are all derivatives on the same underlying value, that are the assets of the company. As a consequence, the model provides an implicit relation across all the securities at each point in time and the corresponding pricing equations, that is the main advantage of adopting this class of models.

Then, we use our estimates to compute both the real-world and the risk-neutral probability of default, for each firm, and we combine the two measures to compute the default risk premium. We show that the equity and the credit markets display a relation with the default risk premium which is opposite to each other, and the type of the relation depends on whether the firm generates an expected return on the market value of the assets higher than the risk-free rate.

Therefore, our paper makes three contributions. First, we propose a novel methodological approach to estimate a corporate credit risk model, by using only market information, for a sample of non-financial firms, across different geographic regions. We generally find that the market value of the leverage, defined as the debt-to-asset ratio, has a peak around the 2008-2009 financial crisis. Also, we estimate the actual value of the default boundary for each firm in our sample, and we report values of the default barrier around 75% of the face value of the debt. This number is in line with the previous results of Davydenko (2012), who estimates a default boundary around 70% of the debt value, and Wong and Choi (2009), who report a default barrier equal to 74% of the debt.

As second contribution, we address the estimation of the default risk premium by using only market data, and combining information from both equity value and CDS spreads. The result is a measure of the distance between actual and risk-neutral default risk priced at the same time in both credit and equity markets. We find that the default risk premium is substantially time varying: the premium displays a lowest peak around the great recession, and an additional downturn during the sovereign debt crisis, following a strong increase in the actual default risk of our sample firms. We generally find that the default risk premium is greater than one for most of the firms, that is the risk-neutral default probability is higher than the corresponding real-world measure. However, we report the opposite result for a significant number of firms: more than one quarter of our sample, on average over time, has default risk premium lower than 1.

As last contribution, we bring new insight on the broad literature which links default risk and equity returns. We show that the dynamics of the default risk premium can be associated to opposite dynamics of equity value and CDS spreads. We split our sample firms in two sub-samples: high premium (default risk premium greater than 1) and low premium (default risk premium lower than 1) firms. We show that the difference in the correlation, between equity prices and default risk premium, across the two sub-samples is positive and significant, and the difference in the correlation, between CDS spreads and default risk premium, across the two sub-samples is negative and significant. We corroborate our findings by implementing long-short portfolio strategies of equity shares and CDS, based on the forecast on the default risk premium generated by the Kalman filter. We buy (sell) equity shares and we sell (buy) CDS when we predict an increase (a decrease) in the default risk premium, for the high premium firms. Instead, we buy (sell) equity shares and we sell (buy) CDS when we predict a decrease (an increase) in the default risk premium, for the low premium firms. We report highly significant performance for our long-short portfolio strategies, both in-sample and out-of-sample.

The paper closest to ours is by Friewald et al. (2014), who relate CDS spreads to equity returns in a Merton (1974) model. They show that the market risk premium, defined as the excess rate of return on the asset, depends on both the risk-neutral and the real-world default probabilities. They argue that both positive and negative relations between credit and equity risk premia are consistent with credit risk structural models. Compared to their results, we provide evidence that the dynamics of the relation between risk-neutral and real-world default probabilities may match with opposite dynamics of credit and equity securities across firms, depending on whether the firm generates positive or negative excess return on the asset.

The rest of the chapter is organized as follows. In the next Section, we relate our study to the previous literature. In Section 3.3, we detail the model and the estimation methodology. Moreover, we define the default risk premium, and we provide numerical evidence on the relation between the default risk premium and the market variables. The data are described in Section 3.4, and Section 3.5 presents the empirical analysis. The first part of this section reports our results on the contingent claim model estimation. In the second part, we focus on the default risk premium and the empirical relation between the premium and the market variables. Section 3.6 concludes and introduces directions for future research.

3.2. Related Literature

The first method to estimate a credit risk Merton (1974) model is based on a system of two equations linking equity value and volatility to the unobservable asset value and volatility. This method, proposed by Jones et al. (1984) and Ronn and Verna (1986), then amended by Bohn and Crosbie (2003) and Vassalou and Xing (2004), is known as volatility-restriction. Duan (1994) highlights the drawbacks of this approach and constructs a transformed likelihood function to estimate the asset values from the observable equity prices.

Brockman and Turtle (2003) address the estimation of a model with early default risk, and they proxy the value of the asset by adding the equity value to the book value of the debt. They report values of the default boundary above the nominal value of the debt. Therefore, they predict that the equity holders may liquidate the firm when the asset value is more than enough to guarantee the payment of the final debt. Perlich and Reisz (2007) argue that the results of Brockman and Turtle (2003) are driven by the mispricing of the debt when the nominal value is used in the estimation. Ericsson and Reneby (2004) and Li and Wong (2008) show that the maximum-likelihood approach of Duan (1994) outperforms both the volatility-restriction and the market proxy methods. By using a transformed maximum-likelihood approach to estimate a barrier-dependent model, Wong and Choi (2009) report values of the barrier around 70% of the face value of the debt. As in the original Duan (1994) formulation, they use equity prices to infer the asset value, by constructing a modified likelihood function adapted to the barrier framework. Similar results are documented in Davydenko (2012), where the market value of the asset just prior to the default is estimated for firms with only observable market value of debt. This is the case of firms with all liabilities traded on the market. Forte (2011) proposes a calibration approach to derive the asset value from the equity prices, and the default boundary from the CDS spreads, while Forte and Lovreta (2012) implement a double-stage maximization algorithm to extract both asset value and default barrier from the stock prices, under the assumption that the equity value incorporates the optimal level of the barrier. Both papers yield estimates of the default barrier below the nominal value of the debt.

Our paper nests this literature, with significant innovations. We estimate a barrier-dependent model by using only market data. In contrast with Brockman and Turtle (2003), Perlich and Reisz (2007), Wong and Choi (2009), and Forte and Lovreta (2012), we use both equity prices and CDS spreads. This approach allows to estimate the market value of debt, rather than using the book value. Moreover, by using equity value and CDS spreads at the same time, we estimate the asset value and the default barrier simultaneously.

We assume an exogenous boundary, as in Longstaff and Schwartz (1995). In this case, the barrier may represent certain covenants between creditors and debtors, where managers are committed to maintain some financial ratios above pre-specified levels in order to preserve the possibility of debt financing. Moreover, consistently with the previous literature, we specify a default barrier that is strictly positive but no greater than the book value of the liabilities. The intuition is simple. As long as the asset value is above the default boundary, even if is below the face value of the debt, the shareholders are willing to cover the promised payment and bear these losses in order to keep the firm operating. Otherwise, the firm defaults and the equity holders get nothing.

We focus our attention on the dynamics of the default risk premium which captures remarkable attention in both asset pricing and credit risk literatures. Different approaches have addressed the estimation of the premium. Hull et al. (2005) compute the ratio between the risk-neutral and the real-world default probability by using corporate bond spreads and historical data, while Driessen (2005) estimates a reduced-form model to derive the default risk premium implied by corporate bond spreads and rating-based default probability. Driessen (2005) finds that the premium is statistically and economically significant. In a pioneer work on the CDS-based default risk premium, Berndt et al. (2008) perform both panel regression and reduced-form model estimation by using 5-years CDS spreads and Moody's EDFs data, and they show that the premium is time, sector, and rating-varying. Diaz et al. (2013) adopt an approach similar to Driessen (2005) and Berndt et al. (2008), however using a wider term structure of CDS spreads on European firms. Huang and Huang (2012) calibrate different credit risk models with corporate bond spreads and default data from rating agencies, and find that the premium decreases with the credit quality and after the crisis periods.

In terms of magnitude, Berndt et al. (2008) report a default risk premium around two, thus arguing that the investors price twice the expected default loss evaluated under the actual probability measure. Driessen (2005) estimates that the investors multiply the actual default probability by a factor close to 6 for pricing corporate bonds. With corporate bonds, the premium describes the amount required by the investors to bear the default risk of the company, which affects the bond repayment. Therefore, we report similar but lower premia compared to Berndt et al. (2008) and Driessen (2005), with the exception of the US sample, which displays default risk premia significantly larger.

3.3. The Model

We define the firm i as an entity fully financed with equity, with market value $E_{i,t}$ at time t, and a zero-coupon debt with face value F_i maturing at time T. Let $V_{i,t}$ be the asset value of the i-th firm and $Z_{i,t}$ its risky zero-coupon bond value at time t. Then, the following condition holds for every point in time and for every firm:

$$V_{i,t} = E_{i,t} + Z_{i,t}.$$

The asset value of the i-th firm is described by a geometric Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, \mathcal{P})$:

$$dV_{i,t} = \mu_{V_i} V_i dt + \sigma_{V_i} V_i dW_{i,t},$$

where μ_{V_i} and σ_{V_i} are the $\mathcal{P}$-drift and diffusion constant coefficients, and $W_{i,t}$ is a standard Brownian motion under the physical probability measure $\mathcal{P}$.

We define the market value of leverage as $L_{i,t} = \ln\left(\frac{F_i}{V_{i,t}}\right)$, following an arithmetic Brownian motion:

$$dL_{i,t} = \mu_{L_i} dt - \sigma_{L_i} dW_{i,t}, \tag{14}$$

where $\mu_{L_i} = -\left(\mu_{V_i} - \frac{1}{2}\sigma^2_{V_i}\right)$ is the $\mathcal{P}$-leverage drift coefficient, and $\sigma_{L_i} = \sigma_{V_i}$ is the leverage diffusion component. Because of the inverse relationship between asset and leverage, there is perfect and negative correlation between the two Brownian motions driving the asset and the leverage diffusion.

For pricing securities, we adopt a first-time passage framework, as in Black and Cox (1976) and Longstaff and Schwartz (1995), with constant and deterministic barrier below the face value of the debt. The firm defaults as soon as the asset value crosses from above the barrier C_i, at any time s, with $t \leq s \leq T$, where T is the debt maturity. Therefore, the equity value $E_{i,t}$ is a down-and-out European call (DOC) option on the asset value of the firm, with maturity T. We rearrange the pricing formula to express the equity value as function of the leverage:[11]

$$\begin{aligned} E_{i,t} = {} & \frac{F_i}{e^{L_{i,t}}}\Phi(d_1) - F_i e^{-r(T-t)}\Phi(d_1 - \sigma_{L_i}\sqrt{(T-t)}) - \frac{F_i}{e^{L_{i,t}}}\exp\left((K_i + L_{i,t})\left(\frac{2r}{\sigma^2_{L_i}} + 1\right)\right)\Phi\left(d_1^{C_i}\right) \\ & + F_i e^{-r(T-t)}\exp\left((K_i + L_{i,t})\left(\frac{2r}{\sigma^2_{L_i}} - 1\right)\right)\Phi\left(d_1^{C_i} - \sigma_{L_i}\sqrt{(T-t)}\right), \end{aligned} \tag{15}$$

where

[11] See Perlich and Reisz (2007) for original pricing equations of default probability and equity value in a first-time passage framework.

$$d_1 = \frac{-L_{i,t} + \left(r + \frac{1}{2}\sigma_{L_i}^2\right)(T-t)}{\sigma_{L_i}\sqrt{(T-t)}}, d_1^{C_i} = \frac{(2K_i + L_{i,t}) + \left(r + \frac{1}{2}\sigma_{L_i}^2\right)(T-t)}{\sigma_{L_i}\sqrt{(T-t)}}.$$

Φ stands for the cumulative distribution function of a standard normal variable, and $K_i = \ln\left(\frac{C_i}{F_i}\right)$. Since the default barrier is below the face value of the debt, K_i takes negative values only. The larger is the absolute value of K_i, the larger is the distance between the face value of the debt F_i and the default barrier C_i, and the larger is the potential amount the shareholders are willing to take out of their pocket in order to bail out the debt to keep the firm operating.

Default risk is also priced in the Credit Default Swaps (CDS), issued with different maturity τ_j, with j going from 1 to J, where the longest maturity τ_J matches the debt maturity T. In a CDS contract, the protection buyer pays a fixed premium each period until either the default event occurs or the contract expires, and the protection seller is committed to buy back from the buyer the defaulted bond at its par value. Therefore, the price of the CDS (i.e., the premium paid by the insurance buyer) is defined at the inception date of the contract to equate the expected value of the two contractual legs. Then, by assuming the existence of a default-free money market account appreciating at a constant continuous interest rate r, and M periodical payments occurring during one year, the CDS spread γ with time-to-maturity τ_j, priced at $t = 0$, solves the following equation:

$$\sum_{m=1}^{M} T\frac{\gamma}{M}\exp\left(-r\frac{m}{M}\right)\mathcal{E}_0^{\mathcal{Q}}[1_{t^*>\frac{m}{M}}] = \mathcal{E}_0^{\mathcal{Q}}[\exp(-rt^*)\alpha 1_{t^*<\tau_j}],$$

where t^* stands for the time of default, α is the amount paid by the protection seller to the protection buyer in case of default, and $\mathcal{E}_0^{\mathcal{Q}}$ indicates that the expectation is taken under the risk-neutral measure $\mathcal{Q}$. Therefore, $\mathcal{E}_0^{\mathcal{Q}}[1_{t^*<\tau_j}]$ is the probability that the firm defaults at any time before τ_j, that is the probability that the asset value crosses from above the barrier C_i. At t, this probability is equal to:

$$PD_{i,t}^{\mathcal{Q}}(\tau_j) = \Phi\left(\frac{K_i + L_{i,t} - \left(r - \frac{1}{2}\sigma_{L_i}^2\right)(\tau_j - t)}{\sigma_{L_i}\sqrt{(\tau_j - t)}}\right) + \exp\left((K_i + L_{i,t})\left(\frac{2r}{\sigma_{L_i}^2} - 1\right)\right)\Phi\left(\frac{(K_i + L_{i,t}) + \left(r - \frac{1}{2}\sigma_{L_i}^2\right)(\tau_j - t)}{\sigma_{L_i}\sqrt{(\tau_j - t)}}\right), \quad (16)$$

if $\tau_j < T$, otherwise

$$PD^{Q}_{i,t}(\tau_J) = 1 - \Phi\left(\frac{-L_{i,t} + \left(r - \frac{1}{2}\sigma^2_{L_i}\right)(\tau_J - t)}{\sigma_{L_i}\sqrt{(\tau_J - t)}}\right) + \exp\left((K_i + L_{i,t})\left(\frac{2r}{\sigma^2_{L_i}} - 1\right)\right)\Phi\left(\frac{(2K_i + L_{i,t}) + \left(r - \frac{1}{2}\sigma^2_{L_i}\right)(\tau_J - t)}{\sigma_{L_i}\sqrt{(\tau_J - t)}}\right), \quad (17)$$

since we need to consider both early bankruptcy risk (equation (16)) and the risk that the firm may not be able to pay back the outstanding debt F_i at time T, even though the asset value never crossed the default boundary before T.

3.3.1. The Default Risk Premium

The default risk premium is defined as the ratio between the risk-neutral default probability and the real-world default probability:

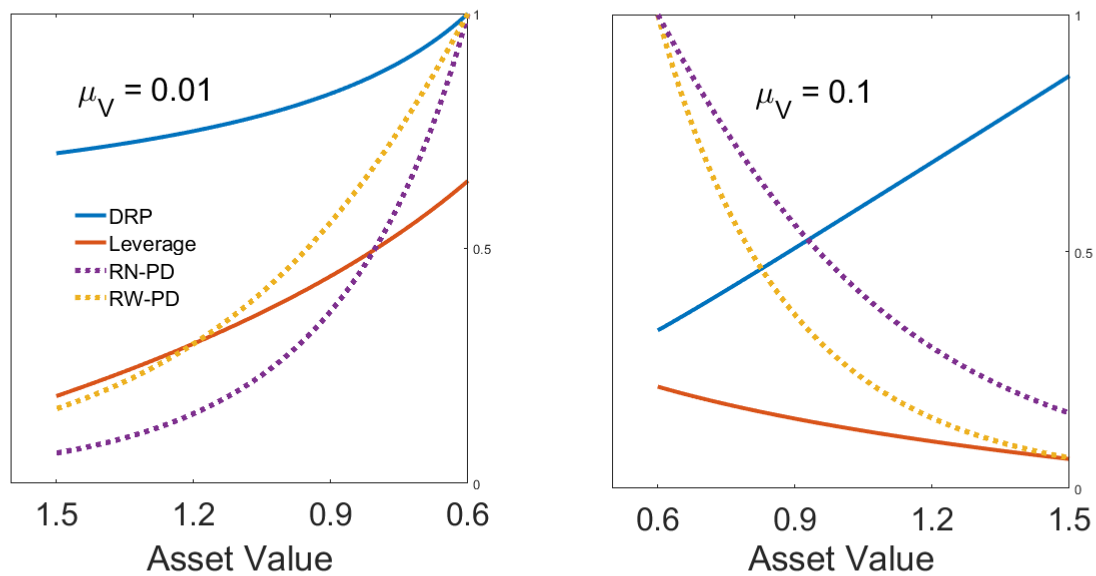
$$DRP_{i,t}(\tau_j) = \frac{PD^{Q}_{i,t}(\tau_j)}{PD_{i,t}(\tau_j)} \quad (18)$$

The default risk premium can be computed for each firm i, each point in time t, and each time horizon τ_j. Our target is to investigate the relation between the dynamics of the default risk premium and the dynamics of both the stock price and the CDS spreads. There is a straightforward relation between leverage and both stock price and CDS spreads. Since the leverage increases when the asset value of the firm decreases, an increase in the leverage is associated to both a decrease in the equity value and an increase in the CDS spreads.

Therefore, the natural procedure to achieve our target is to study the sensitivity of the default risk premium with respect to the leverage. We study here the sign of the derivative of the premium with respect to the leverage numerically. We show that, *ceteris paribus*, the sign of this derivative depends on the difference between μ_V and r, and in particular whether this difference is positive or negative. The plots in figure 21 illustrate this result.

The left plot shows that the default risk premium increases when the leverage increases, so the equity value of the firm depreciates, and the CDS spread increases. The default risk premium increases, in fact, because the risk-neutral probability of default increases more than the real-world probability of default. This is the case with $\mu_V < r$, and so default risk premium

Figure 21. Default Risk Premium and Leverage. Numerical Example



The left plot shows the variation of the leverage, default risk premium, real-world and risk-neutral default probabilities, when the asset value of the firm decreases, up to the default boundary. The right plot shows the variation of the leverage, the default risk premium, real-world and risk-neutral default probabilities, when the asset value of the firm increases, starting from the default boundary. The value of the model parameters common to both plots are: $r = 0.05$, $\sigma_V = 0.2$, $C = 0.6$. On the left plot, we use $\mu_V = .01$. On the right plot, we use $\mu_V = 0.1$.

lower than 1 (i.e., $PD^Q < PD^P$). The right plot shows that the default risk premium increases when the leverage decreases, so the CDS spread decreases, and the equity value increases. The default risk premium increases, in fact, because the real-world probability of default decreases more than the risk-neutral probability of default. This is the case with $\mu_V > r$, and so default risk premium larger than 1 (i.e., $PD^Q > PD^P$). Therefore, the increase in the default risk premium can be associated to opposite dynamics of stock price and CDS spreads. This result does not depend on the particular choice of the model. In the appendix, we present a very simple and general pricing model, where we show the relation between the default risk premium and the dynamics of the stock price, and the relation between the default risk premium and the dynamics of an insurance premium (such as the CDS spread).

3.4. Data

Our dataset consists of daily CDS spreads traded on four different maturities (1, 3, 5, and 10 years), which are the most liquid on the CDS market, and daily market capitalization (i.e., the product between the number of outstanding shares and the share price), between the 20th

December 2007 and the 19th December 2013. We use only reference entities with active CDS contracts before the 20th December 2007, that is the starting point of our time series. The choice of this time window is driven by the so-called International Monetary Market dates.[12]

The CDS premia are expressed in basis points. The premium that the protection seller receives is expressed as an annualized percentage of the notional value of the transaction, and this value is recorded as the market price. The details of a CDS transaction are recorded in the CDS contract, which is usually based on a standardized agreement prepared by the International Swaps and Derivatives Association (ISDA), an association of major market participants.[13]

The event triggering the payment of the protection leg in a CDS contract ranges from no-restructuring to full-restructuring. As priority rule, we select the contracts that adopt the no-restructuring (MR) clause, that is the standard convention after the CDS Big Bang protocol in April 2009. Otherwise, we include the contracts that adopt the modified-restructuring (MR) clause, that was the standard convention before the protocol. As last options, we include the contracts with full-restructuring or modified-modified restructuring clause.

We download CDS data from Thomson Reuters Datastream, which provides information on CDS spreads from December 2007. The sample firms are chosen from the reference entities listed on the Markit indexes, where the most liquid companies in terms of CDS transactions are quoted. In particular, we refer to the iTraxx indexes for Australia, Japan, and Europe. In the last index, we select only the Eurozone entities to ensure homogeneity in currency. The CDX North America - Investment Grade index is adopted for the United States sample, where we select only the US companies for the same homogeneity reason.[14]

We exclude the financial firms, which are characterized by a peculiar capital structure in terms of assets and liabilities. Also, we control for stale prices, where a price is defined as stale

[12] This term is also used for the conventional quarterly termination dates of credit default swaps, which fall on 20 March, 20 June, 20 September and 20 December. From late 2002, the CDS market began to standardize credit default swap contracts so that they would all mature on one of the four days of 20 March, 20 June, 20 September and 20 December. These dates are used both as termination dates for the contracts and as the dates for quarterly premium payments. So, for example, a five-year contract traded any time between 20 September 2005 and 19 December 2005 would have a termination date of 20 December 2010.

[13] See Longstaff et al. (2005) for an extensive description of the CDS contractual structure, and functioning

[14] The Markit iTraxx and CDX indices are constructed every six months, according to specified criteria and selections rules, which ensure the eligibility of an entity to be a constituent of the index, especially in terms of CDS contracts liquidity. We refer to the 20th list for the iTraxx indexes, and the 21th list for the CDX index. All have been issued on September 2013

when does not change for many consecutive days. Stale prices may cause issues when we extract risk-neutral default probabilities from the CDS spreads curve. In fact, a large discrepancy in the number of active trading days across CDS with different maturities may produce non-monotone default probabilities for a given reference entity, while the default probability for the longer time horizon must be at least equal to the default probability for the shorter time horizon. In fact, while it is possible to have a CDS downward sloping term structure when the company is perceived to be riskier in the short term than in the long term (similarly to a downward sloping term structure of interest rates), this curve must be not too steep. We drop out from the dataset the companies with non-monotone default probabilities for more than 1% of the time series for at least one time horizon.

At the end, we are left with a dataset of 172 firms. The data on the daily market capitalization are downloaded from Bloomberg. We delete 8 firms with no data on the market capitalization. Our final dataset consists of 164 firms, divided as follows: Australia (9), Japan (10), United States (89), Europe (56).

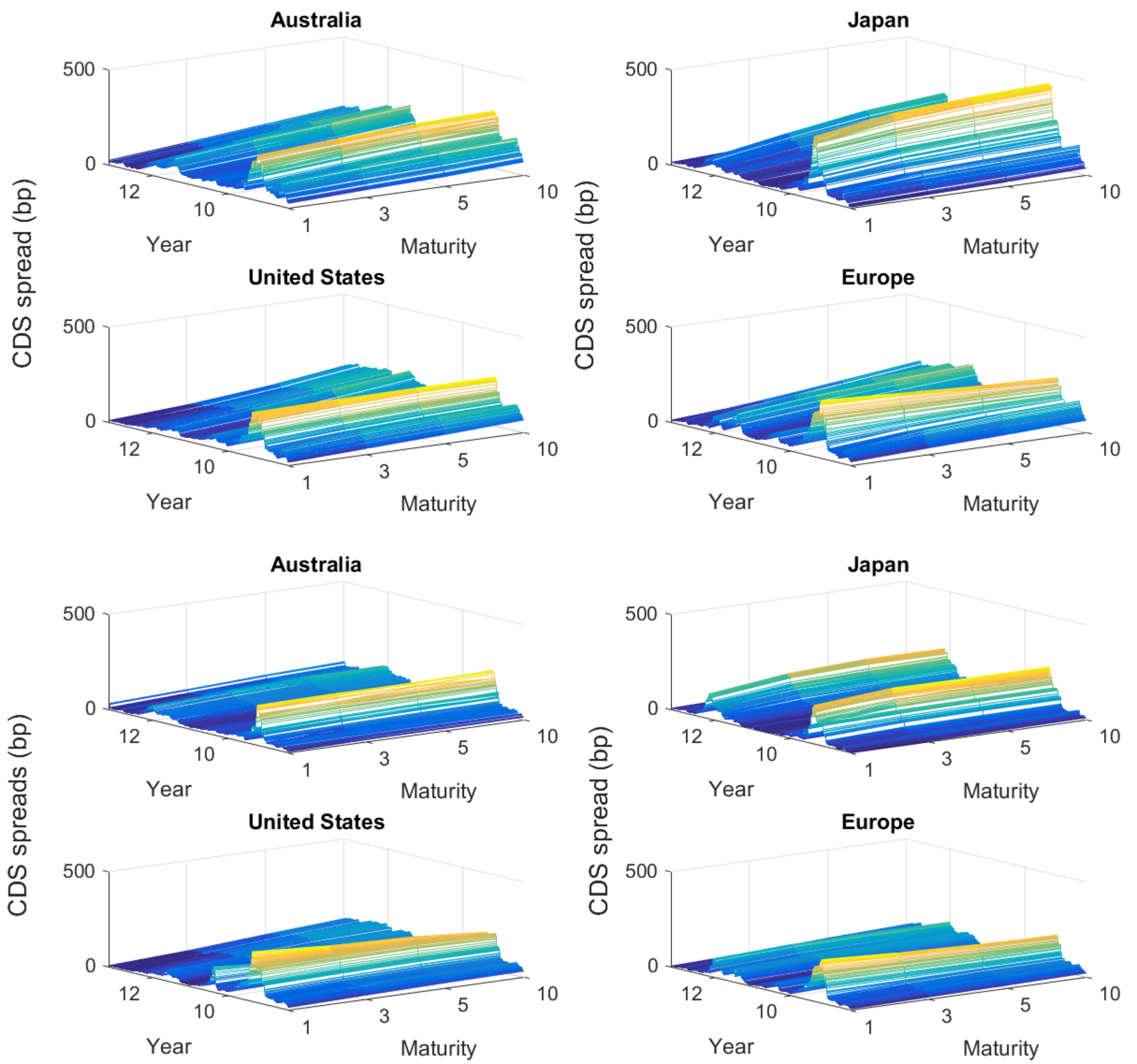
As proxy of the risk-free rates, we use the sovereign bonds curve constructed by Bloomberg for Unites States, Australia, and Japan, respectively. The European curve is the result of the aggregation of the sovereign bonds issued by France, Germany, and Holland. Given the CDS spreads, the termination date, the standard assumption on the recovery rate (i.e., 40% of the debt notional value), and the risk-free rate term structure, we extract the risk-neutral default probability from the CDS spreads for the corresponding maturity, and for each day. There is one only arbitrage-free default probability implied by the CDS spreads, that equates the starting value of the premium leg (protection buyer) and the protection leg (protection seller) in a CDS contract.[15]

3.4.1. Summary Statistics

We report in Table 14 the descriptive statistics on CDS spreads for all sample firms. The upper plot of figure 22 shows the term structure of the average CDS spreads for each region. The worldwide propagation of the 2008-2009 financial crisis generates the peak in the CDS premia around that time interval. The difference in the CDS spreads across years is very pronounced

[15]See O'Kane and Turnbull (2003) for an extensive overview of the CDS valuation

Figure 22. CDS Spreads. Regional Average and Standard Deviation

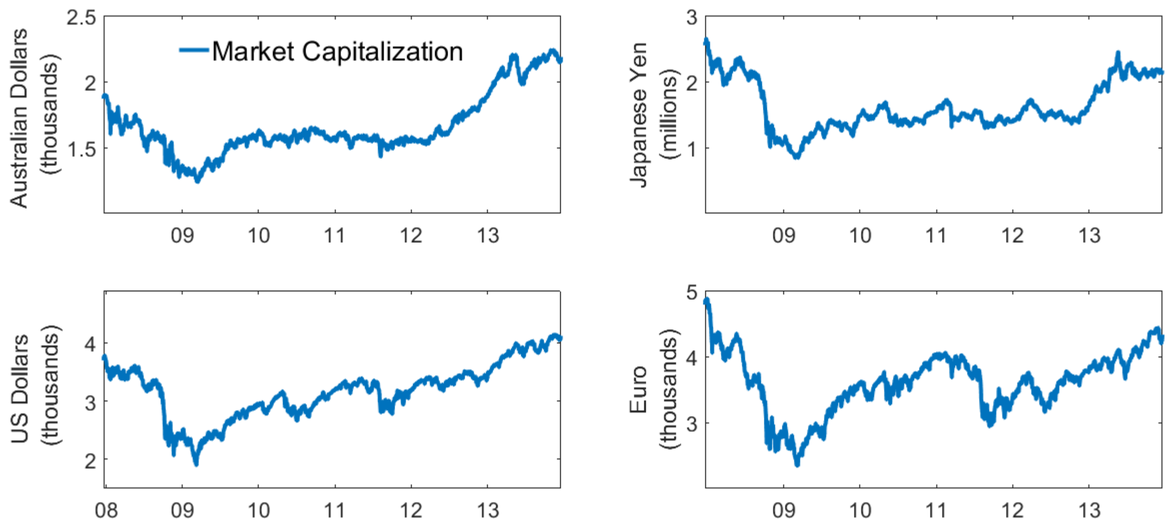


The four upper plots show the cross-sectional average of the CDS spreads across firms for the four geographic regions, between the 20th December 2007 and the 19th December 2013 (X-axis), on the four different maturities (Y-axis). The four lower plots show the cross-sectional standard deviation of the CDS spreads across firms for the four geographic regions, between the 20th December 2007 and the 19th December 2013 (X-axis), on the four different maturities (Y-axis). Both the average and the standard deviation are expressed in basis points.

for the shorter time horizons, and in particular for the one-year maturity, while the CDS premia paid on longer time horizons are more stable over time. It turns out that the conjecture of the market on the default risk in the short-term dramatically increased during the crisis.

This pattern is homogeneous across regions. However, the strong concentration of the

Figure 23. Market Capitalization. Regional Average



The plots show the cross-sectional average of the market capitalization across firms for the four geographic regions, between the 20th December 2007 and the 19th December 2013. Market capitalization is computed as number of outstanding shares times share price, and is expressed in the regional currency.

financial turmoil in the US market causes large premia paid on the US reference entities in the first part of the time series, and the European sovereign debt crisis causes large premia paid on the Eurozone companies in 2011-2012.

The financial crisis of 2008-2009 is also characterized by a large cross-sectional standard deviation of the CDS spreads (figure 22, lower plot), thus documenting great heterogeneity across entities in terms of CDS premia. The combination of both large spreads and great heterogeneity across firms reflects the incentive of the insurance sellers, even more emphasized during a crisis period, to ask for large premium on risky entities and to accept lower premium on safer companies.

Summary statistics on market capitalization are calculated for each region, and reported in Table 15. Data are expressed in millions of the local currency.

Market capitalization substantially drops during the 2008-2009 crisis across all regions, and generally increases afterwards. Some peculiarities still appear across regions. The market capitalization swiftly grows up for US and Eurozone firms from 2009 to 2010. Then, the positive trend is more persistent for US companies compared to the Eurozone firms. Also, this shift at the end of the crisis is not very remarkable for Australian and Japanese firms, that show instead big jumps during the 2013.

3.5. Estimation Results

3.5.1. State-Space model

We formulate our estimation problem in a state-space model, where the measurement equations are the equations (15), (16), and (17). We observe equity value and CDS spreads for four time horizons (1, 3, 5, and 10 years), so that we have five measurement equations. Let g and h to define the non-linear relationships between the observable variables and the latent variable ($L_{i,t}$), and $\epsilon_{i,t}(\tau_j)$ the measurement noise associated to the CDS implied-default probability at the time horizon j, then the set of measurement equations of the state-space model is the following:

$$PD^{Q}_{i,t}(\tau_j) = g\left(L_{i,t}; K_i, \sigma_{L_i}\right) + \epsilon_{i,t}(\tau_j), [j = 1, 3, 5, 10], \tag{19}$$

and

$$E_{i,t} = h\left(L_{i,t}; K_i, \sigma_{L_i}, F_i\right) + \varsigma_{i,t}, \tag{20}$$

where the time to maturity is expressed in years, and $j = 10$ stands for the maturity T of the outstanding debt F_i (i.e., 10 years). The measurement noises in equation (19), for each firm i, are assumed to follow a multivariate normal distribution, with zero mean, and diagonal covariance matrix R_i. We assume a homoscedastic covariance matrix, which is firm-varying. $\varsigma_{i,t}$ is the measurement noise associated to the equity observation. $\varsigma_{i,t}$ is normally distributed, with zero mean and variance ω_i.

On the other side, the transition equation describes the evolution over time of the leverage. It follows from the discretization of the stochastic process defined in equation (14):

$$L_{i,t+\delta t} = L_{i,t} + \mu_{L_i}\delta t + \eta_{i,t+\delta t},$$

where $\eta_{i,t+\delta t} = \sigma_{L_i}(W_{i,t} - W_{i,t+\delta t}) \backsim \mathcal{N}(0, Q_i)$ is the transition error, and $Q_i = \sigma^2_{L_i}\delta t$.

We combine the default probability extracted from the CDS spreads and the equity value to compose the set of observable variables, thus combining information from both credit and equity markets to estimate the unobservable dynamics of the market value of the leverage and the model parameters (i.e., F_i, μ_{L_i}, σ_{L_i}, and K_i).

The face value of the debt, F_i, appears in both the equations (15) and (17), where it appears in terms of the parameter K_i. We estimate F_i since we are mapping companies with complex debt structure, in terms of debt instruments and maturities, into companies issuing one single zero-coupon bond. This simplification does not affect the definition of default risk. In fact, what only matters in terms of default risk is the threshold triggering the default. As long as the asset value is above the default boundary, the shareholders are willing to cover the firm's losses, even if the asset value is not enough to meet the promised debt payment at maturity.

We perform the estimation for each firm. We infer the dynamics of the unobservable variable by iteration of the predicting and the updating equations of the Kalman filter, and we estimate the model parameters by performing a maximum likelihood algorithm, under the assumption of normality for the measurement errors. We describe in details the non-linear Kalman filter and the likelihood function in the appendix. The measurement errors are given by the difference between the actual and the fitted value of the observable variables. The fitted value of the observable variables is computed by using the equations (15), (16), and (17), respectively, using the current estimate of the state variable.

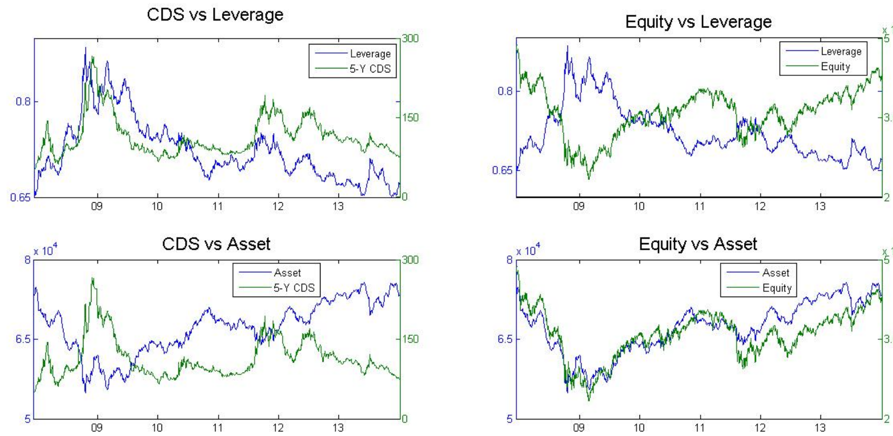
As a result, we reconstruct the time series of the market value of the leverage for each firm in our sample. Estimating the dynamics of the asset value $V_{i,t}$, the dynamics of the market value of the debt $Z_{i,t}$, and the level of the default boundary C_i, is then straightforward. In fact, from the definition of $L_{i,t}$, and K_i, it follows that:

$$\begin{aligned} V_{i,t} &= \frac{F_i}{e^{L_{i,t}}}, \\ C_i &= e^{K_i} \times F_i, \\ Z_{i,t} &= V_{i,t} - E_{i,t}. \end{aligned} \tag{21}$$

3.5.2. The Dynamics of Leverage and the Default Boundary

We report in Table 16 mean and standard deviation of our estimates of the model parameters, for each region. To provide a graphical idea of our estimation results, we show in Figure 24

Figure 24. Leverage and Assets - Eurozone



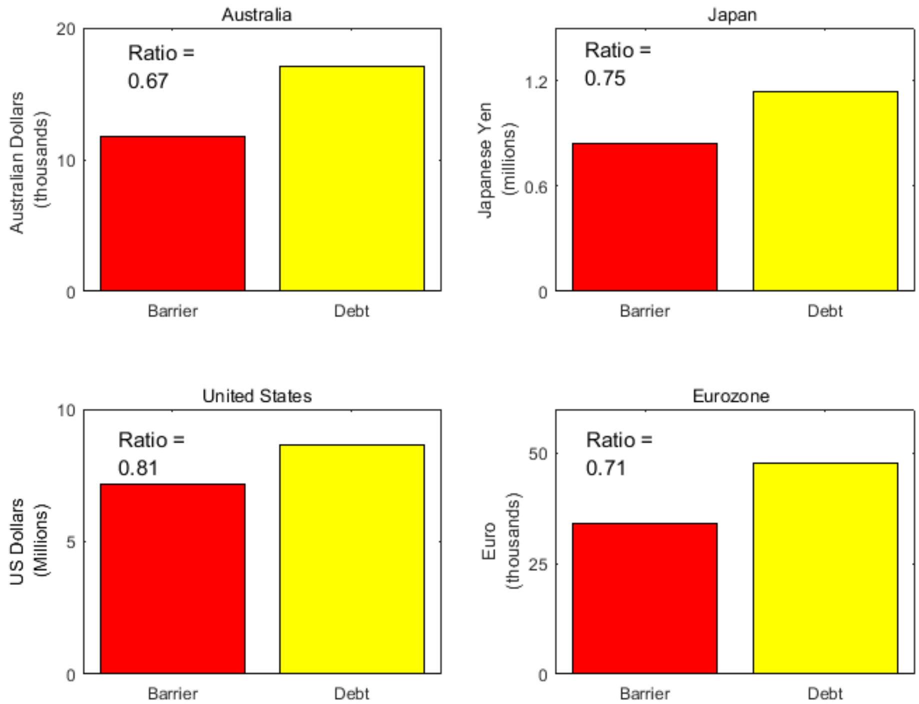
The top-left plot shows the average estimated leverage against the average observed CDS spreads, across the Eurozone firms, on the 5-years maturity, between the 20th December 2007 and the 19th December 2013. The leverage is computed as $e^{L_{i,t}} = \frac{F_i}{V_{i,t}}$, where F_i is the estimated value of the firm's total debt, and $V_{i,t}$ is the estimated market value of the firm's assets. The CDS spreads are expressed in basis points. The top-right plot shows the average estimated leverage against the average market capitalization, across the Eurozone firms, between the 20th December 2007 and the 19th December 2013. Market capitalization is expressed in millions of Euro. The bottom-left plot shows the average estimated market value of assets against the average CDS spread, on the 5-years maturity, across the Eurozone firms, between the 20th December 2007 and the 19th December 2013. CDS spreads are expressed in basis points. The bottom-right plot shows the average estimated market value of assets against the average market capitalization, across the Eurozone firms, between the 20th December 2007 and the 19th December 2013.

the dynamics of the average asset value and leverage across the Eurozone firms. The leverage is expressed in terms of debt-to-asset ratio. In Figure 24, we compare asset value and leverage with market capitalization and the 5-years CDS spreads.

The average market value of leverage displays a peak around the 2008-2009 financial crisis, when the debt counts for more than 80% of the market value of the asset. Following a decreasing trend until 2011, the leverage increases again during the sovereign debt crisis. By construction, the value of the asset displays the opposite dynamics, showing a lowest peak around the great recession and an increasing trend in the last part of the time series, partially interrupted during the sovereign debt crisis of 2012.

The market value of the leverage has very similar dynamics to the CDS spreads dynamics. On the other side, the leverage moves in the opposite direction compared to the dynamics of the market capitalization. Therefore, our estimation of the market value of the leverage reflects information coming from both markets. Specular discussion applies to the estimation of the

Figure 25. Debt and Barrier



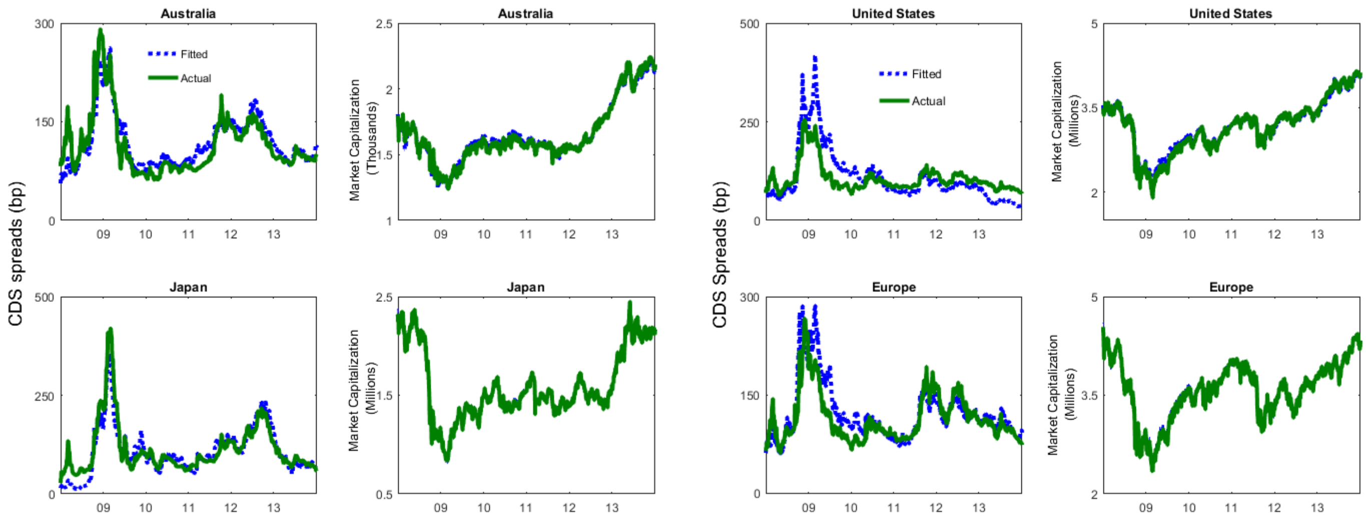
The plots show the cross-sectional average of the total value of debt, and of the default barrier, across firms, for the four geographic regions. The values are expressed in the regional currency. In the plot, we report the average ratio, across firms, between the value of the default boundary and the total value of debt.

asset value, that follows a trend that is opposite to the CDS spreads dynamics and in line with the market capitalization dynamics. We derive similar results for the other geographic regions. However, the downturn of the asset value in 2012 is less pronounced for the other regions, due to a lower impact of the sovereign debt crisis.

A key quantity in our framework is the willingness of the shareholders to rescue the firm in case of distress to avoid default, that is measured by K_i, which proxies the distance between the default boundary and the face value of the debt. By using equation (21), we quantify the default barrier and the outstanding debt, for each firm. Figure 25 shows the average barrier and debt across firms, for each region, and reports the barrier-to-debt ratio.

We document similar barrier-to-debt ratios across regions, between 67% (Australia) and 81% (United States). This result suggests that the Australian firms are more likely to be rescued by the shareholders compared to the US firms. These findings are in line with the estimates of

Figure 26. Estimation Performance: Fitting Real Data



The plots compare the model implied dynamics of the 5-years CDS spreads and equity value against the observed 5-years CDS spreads and equity value, for the average firm across regions, between the 20th December 2007 and the 19th December 2013. The CDS spreads are expressed in basis points, and the equity value is expressed in the regional currency.

Davydenko (2012), who argues that firms on average default when the asset value drops down to 70% of the debt value, and of Wong and Choi (2009), who estimate a default boundary around 74% of the nominal debt.

We show that our estimation results allow the model to match real data quite well. We generate a fitted time series of equity value and CDS spreads, and we compare in Figure 26 with the actual observation of the market data, for each region.

To test the goodness of fit of our estimation results, we construct long-short portfolio strategies using the fitted dynamics of equity value and CDS spreads. At each day, we buy stocks and CDS for which we forecast an increase in the price on the next day, and we sell otherwise. Table 17 reports significant and positive performance. As a robustness, we perform long-short portfolio strategies with randomly selected firms. At each day, we randomly split the sample in two sub-samples, and we buy stock and CDS for the firms belonging to the first sub-sample, and we sell stock and CDS for the firms belonging to the second sub-sample. We repeat this exercise for 1000 simulations. Table 17 reports the maximum performance achieved across the 1000 simulations, which is not significant for none of the securities.

3.5.3. The Default Risk Premium

We calculate the default risk premium as the ratio between the risk-neutral and the real-world default probability, for each firm i, for each point in time t, and each time horizon τ_j. We compute the risk-neutral and the real-world default probability using our estimates in the equations (16) and (17). The result is a daily term structure of the default risk premium for each firm. The difference between the two measures of default probability is given by the drift coefficient of the asset value dynamics. We use the risk-free rate to estimate the risk-neutral default probability, and we use the actual leverage drift, μ_{L_i}, to estimate the real-world default probability. μ_{L_i} is related to the asset drift coefficient as follows:

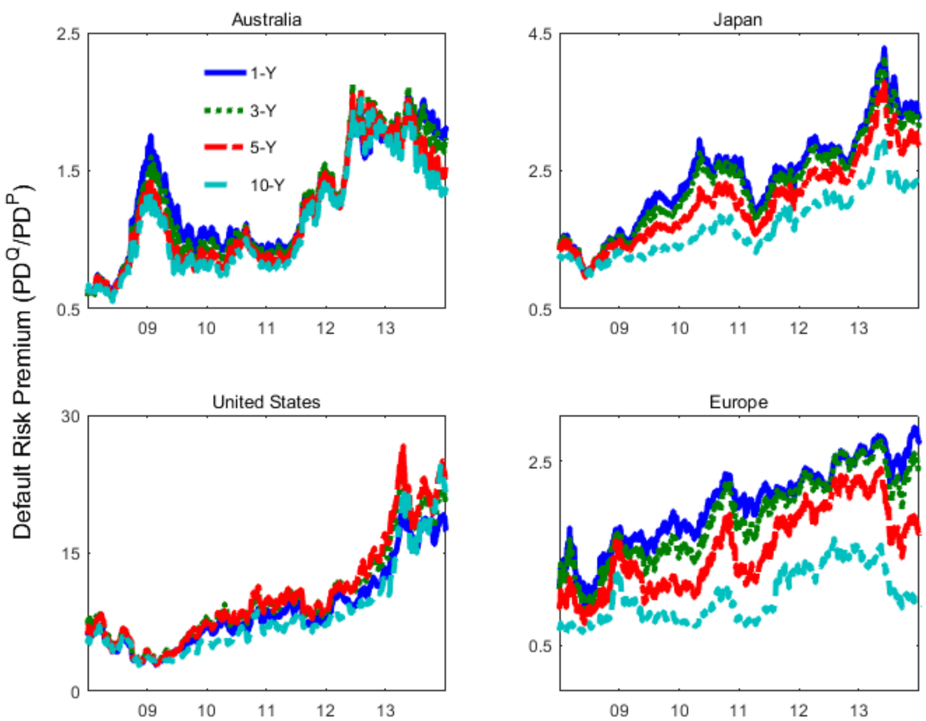
$$\mu_{V_i} = -\left(\mu_{L_i} - \frac{1}{2}\sigma^2_{L_i}\right),$$

where σ_{L_i} is the result of the quasi-maximum likelihood estimation, and μ_{L_i} is the result of a matching moment approach. In fact, after performing the quasi-maximum likelihood algorithm to obtain the estimates of the model parameters, we perform non-linear squares minimization to calibrate the variance of the measurement errors to the observed market data, and we reconstruct the dynamics of the market value of the leverage which is able to fit simultaneously the CDS spreads and the equity value. Finally, we compute the mean of the first differences, divided by the time step. μ_{L_i} is a measure of the average growth of the leverage over the time series, and our estimation of μ_{L_i} therefore contains information on the growing trend of the leverage from both credit and the equity markets. [16]

Figure 27 shows the term structure of the default risk premium for the median firm across regions. The default risk premium is substantially time varying. The variation over time of the premium is primarily due to the evolution of the market value of asset. The premium displays a lowest peak around the great recession, and an additional downturn during the sovereign debt crisis, although less pronounced. This pattern is the consequence of the larger actual default risk measured during the crisis periods. This result is also consistent with Huang and

[16]By adopting this procedure we also avoid the issue related to the insensitivity of the likelihood function with respect to the drift coefficient. Previous works have widely reported huge complexity in estimating the drift component in stochastic processes. We find that even trying to solve this problem by calibration is pretty useless. The impact of changing the drift coefficient of the state variable dynamics on the value of the sum of the squared errors is quite negligible. Instead, the likelihood function is very sensitive to the diffusion component, and we can directly estimate σ_{L_i} in the first step

Figure 27. Term Structure of Default Risk Premium



The plots show the term-structure of the default risk premium, for the median firm across the four regions, between the 20th December 2007 and the 19th December 2013. The default risk premium is expressed as ratio between the risk-neutral and the real-world default probability, where the two measures of default probability are computed for each firm with daily frequency, using the estimation results.

Huang (2012), who argue that the default risk premium decreases as the credit quality declines, since the real-world default risk is larger for the lower-rating firms. In terms of magnitude, we report similar but lower premia compared to Berndt et al. (2008) and Driessen (2005), with the exception of the US sample, which exhibits default risk premia significantly larger (Table 18). Our lower estimates of the default risk premium are probably due to our sample period spanning over the two recent crises, when the actual default risk has substantially increased.

The premium is generally higher for the shorter time horizons. It is well-known that the credit risk models that do not include jumps in the dynamics of the asset value tend to underestimate the actual default probability for short time horizons. However, Huang and Huang (2012) argue that including jumps do not substantially affect the fraction of spread explained by credit risk, unless if the model is calibrated with rather extreme values. Moreover, Huang and Huang (2012) document that the mean reversion of credit quality increases the credit risk for longer maturities, in particular for investment grade companies.

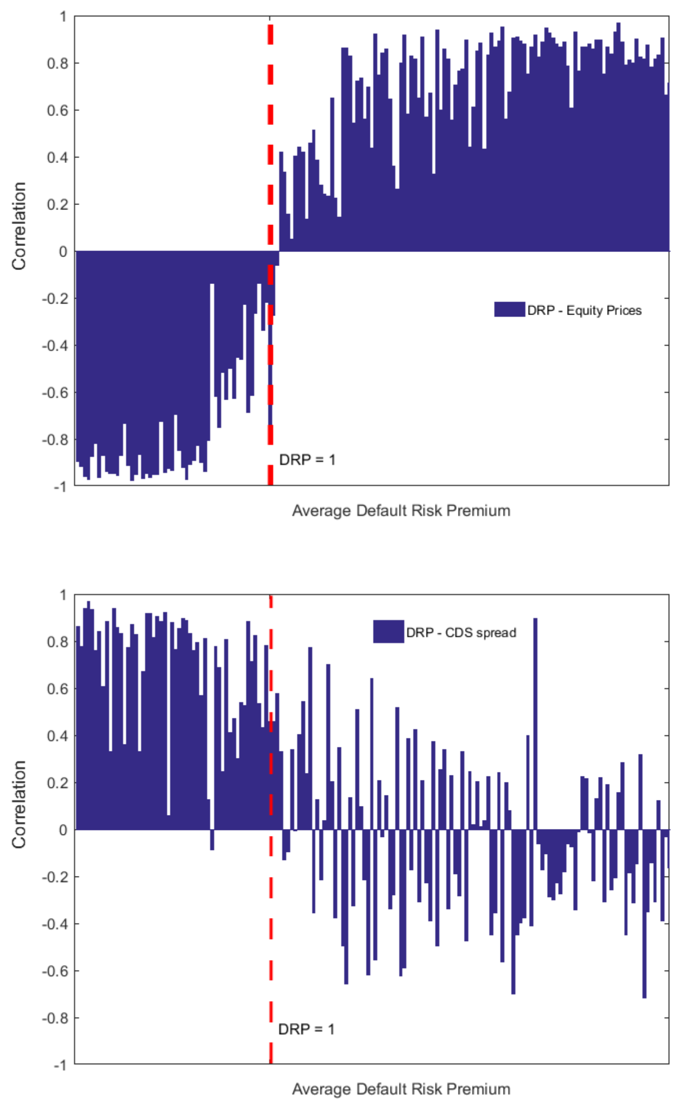
We generally find that the default risk premium is greater than one, that is the risk-neutral default probability is higher than the corresponding real-world measure. However, we report for a few firms an actual default probability higher than the corresponding risk-neutral measure, and therefore a default risk premium lower than 1. For instance, narrowing our attention to the 10-years maturity, we find that the 32% of our sample firms (53 out of 164) is characterized by an average default risk premium lower than 1. Looking at the 5-years maturity, the percentage of firms with average default risk premium lower than 1 drops to 26% of the sample (43 firms). Similar results are obtained for the 3-years maturity (41 firms), and the 1-year maturity (40 firms). In the next section, we explore the relation between this result and the dynamics of the market data, by implementing long-short portfolio strategies of equity and CDS, based on the default risk premium.

3.5.4. Long-Short Portfolio Strategy

We start the analysis on the relation between the default risk premium and the observable market data by studying the sample correlation between the premium and both CDS spreads and equity value for each firm. We calculate the average premium for each firm over time, and we rank the firms according to the average premium, in ascending order. In Figure 28, we show the sample correlation coefficient between the default risk premium and both the equity value and the 10-years CDS spread.

The top panel of Figure 28 shows that the firms with average default risk premium higher than 1 display positive correlation with the equity value, and the firms with average premium lower than 1 display negative correlation with the equity prices. On the other hand, the low premium firms display positive correlation with the CDS-spreads (bottom panel of Figure 28). This correlation substantially decreases as the premium increases, even though this opposite relation is not as uniform as for the low premium firms. Table 19 supports the results statistically. The difference in the correlation between equity value and default risk premium across high premium and low premium firms is positive and very significant, and the difference in the correlation between CDS spreads and default risk premium across high premium and low premium firms is negative and very significant. Therefore, we document opposite relation be-

Figure 28. DRP and Market Variables



The top plot shows the correlation between equity prices and default risk premium computed for each firm between the 20th December 2007 and the 19th December 2013. The bottom plot shows the correlation between 10-years CDS spread and default risk premium computed for each firm between the 20th December 2007 and the 19th December 2013. The firms are ranked according to the average default risk premium over the time series. In both graphs, the dotted line stands for default risk premium equal to 1.

tween the default risk premium and the observable market data across high premium and low premium firms.

We strength this result by using portfolio analysis. We construct long-short portfolios of equity shares and CDS spreads. We implement our trading strategy in the last year of the time series (i.e., 252 daily observations). In a first step, we perform an in-sample analysis, where we sort firms according to the t-average default risk premium by using the market data information up to t, and the estimation of the model parameters is based on full sample information, following the approach of Cochrane and Piazzesi (2005).

At each day t, and for each firm, we compute the average default risk premium until $t-1$. We split the sample in two sub-samples: the firms with average premium higher than 1 (HP-firms), and the firms with average premium lower than 1 (LP-firms). For the HP-firms, an expected increase (decrease) in the premium should predict higher (lower) price of the stock, and lower (higher) CDS spread. For the LP-firms, an expected increase (decrease) in the premium should predict lower (higher) price of the stock, and higher (lower) CDS spread.

Then, at t, we buy (sell) the stocks and we sell (buy) the CDS of the HP-firms for which we predict an increase in the premium at $t+1$. On the other hand, we sell (buy) the stocks and we buy (sell) the CDS of the LP-firms for which we predict an increase in the premium at $t+1$. We predict the value of the premium at $t+1$ by using the one-step ahead forecast of the state variable obtained with the Kalman filter. The top-left panel of Table 20 reports the average daily performance of the trading strategy for the equity and the CDS portfolios, respectively. The daily mean performance of the trading strategy is positive and very significant for both portfolios.

3.5.5. Out-of-Sample and Robustness

We support our results with a set of robustness checks and an out-of-sample test. First, we exclude firms with average default risk premium higher than 1000. Huge values of the premium are due to very low estimates of the real-world default probability, which is the denominator of the ratio. Table 20 shows that the previous results are not altered by narrowing the set of firms available for the trading strategy.

Second, we perform the strategy by trading in the opposite direction compared to the benchmark approach: we buy (sell) the stocks and we sell (buy) the CDS of the LP-firms for which we predict an increase in the premium at $t+1$. On the other hand, we sell (buy) the stocks and we buy (sell) the CDS of the HP-firms for which we predict an increase in the premium at $t+1$. The bottom part of Table 20 shows that this strategy generates negative and significant performance.

Finally, we corroborate our findings with an out-of-sample analysis, where we implement our trading strategy sorting the firms according to the t-average default risk premium, by using the market prices information up to t, and the estimation of the model parameters is based on the information up to t. All the results are substantially confirmed.

3.6. Conclusion

In this paper we estimate a barrier-dependent contingent claim model, with only market data. By gathering information from both credit and equity markets, we infer the dynamics of the market value of assets and debt, and the default boundary, for a sample of non-financial firms. Then, we exploit our estimation results to measure the default risk premium, for each company. Our estimate of the default risk premium, then, reflects the information conveyed by the two markets at the same time. We find that the default risk premium is substantially time varying, and decreases during the crisis periods because of an increase in the actual risk of default.

Our main finding is that the dynamics of the default risk premium can be associated to completely different dynamics of equity value and CDS spreads, and this result depends on whether a firm generates expected rate of return on the market value of assets higher than the risk-free rate.

We believe that our estimation approach may also be addressed for different targets. Combining the result on the default boundary and the market value of the asset, we supply a new insight in the academic research on the corporate credit risk based on the well-known distance-to-default. In addition, we equip the corporate finance research with a new approach for the evaluation of the market value of the debt.

3.7.1. A Simple Asset Pricing Model

Default Risk Premium and Stock Price

Consider two states of the world, *high* and *low*, and a stock that pays out S^h and S^l in the two states, at the payout date T, respectively. We refer to the low state as the default. We define S_t the price of the stock at time t, μ the discounting rate of return under the real-world probability, and r the discounting rate of return under the risk-neutral probability (i.e., the risk-free rate). Hence, at time $t = 0$,

$$S_0 = e^{-\mu} E_0^P[S_T] = e^{-\mu}[p_0 S^l + (1 - p_0)S^h], \tag{22}$$

and

$$S_0 = e^{-r} E_0^Q[S_T] = e^{-r}[q_0 S^l + (1 - q_0)S^h], \tag{23}$$

where E^P indicates the expectation taken under the real-world probability, E^Q indicates the expectation taken under the risk-neutral probability, p is the real-world default probability, and q is the risk-neutral default probability. Therefore,

$$\mu = -\ln\left(\frac{S_0}{E_0^P[S_T]}\right), r = -\ln\left(\frac{S_0}{E_0^Q[S_T]}\right),$$

and

$$p_0 = \frac{S^h - \frac{S_0}{e^{-\mu}}}{S_l}, q_0 = \frac{S^h - \frac{S_0}{e^{-r}}}{S_l}$$

At time $t = 1$,

$$p_1 = \frac{S^h - \frac{S_1}{e^{-\mu}}}{S_l}, q_1 = \frac{S^h - \frac{S_1}{e^{-r}}}{S_l},$$

and

$$\frac{p_1}{p_0} = \frac{e^{-\mu}S^h - S_1}{e^{-\mu}S^h - S_0}, \frac{q_1}{q_0} = \frac{e^{-r}S^h - S_1}{e^{-r}S^h - S_0}$$

Indeed, when $S_1 > S_0$, then $p_1 < p_0$ and $q_1 < q_0$, provided that $S^l < S_t < S^h$. The opposite holds when $S_1 < S_0$.

Moreover, by equating (1) and (2), we obtain

$$\mu - r = \ln\left(\frac{E_0^P[S_T]}{E_0^Q[S_T]}\right)$$

.

Hence, if

$$\mu > r \Leftrightarrow E_0^P[S_T] > E_0^Q[S_T] \Leftrightarrow p(S^l - S^h) > q(S^l - S^h) \Leftrightarrow p < q,$$

as $(S^l - S^h) < 0$. The opposite relationship holds when $\mu < r$.

Let define now Z as the default risk premium, that is the ratio between the risk-neutral and the real-world default probability:

$$Z_0 = \frac{q_0}{p_0},$$

then we have that,

$$\mu > r \Leftrightarrow Z_0 > 1 \tag{24}$$

$$\mu < r \Leftrightarrow Z_0 < 1 \tag{25}$$

We refer to (3) as the *normal* case, and to (4) as the *distressed* one. At every time t,

$$Z_t = e^{r-\mu}\frac{e^{-r}S^h - S_t}{e^{-\mu}S^h - S_t}$$

Hence, the gross rate of growth of the default risk premium $\dot{Z}$ between 0 and 1 is

$$\frac{Z_1}{Z_0} = \frac{e^{-r}S^h - S_1}{e^{-r}S^h - S_0} \cdot \frac{e^{-\mu}S^h - S_0}{e^{-\mu}S^h - S_1} = \dot{q} \cdot 1/\dot{p} \tag{26}$$

Therefore, the default risk premium increases, that is $\dot{Z} > 1$, either because $\dot{q} > \dot{p} > 1$ if $S_1 < S_0$, or because $1/\dot{p} > 1/\dot{q} > 1$ if $S_1 > S_0$. In the first case there is an increase in the risk-neutral default probability higher than the increase in the real-world default probability. This picture is associated to a decrease in the stock price. In the second case there is a decrease in the real-world default probability higher than the decrease in the risk-neutral default probability. This picture is associated to an increase in the stock price.

Moreover, by simply rearranging (5), we can show that

$$\dot{Z} > 1 \Leftrightarrow e^{-r}(S_1 - S_0) > e^{-\mu}(S_1 - S_0)$$

It turns out that $S_1 > S_0$ when $\mu > r$, and $S_1 < S_0$ when $\mu < r$. By combining this result with the statement following the equation (5), we can conclude that an increase in the default risk premium can be associated to an increase in the stock price, when the real-world default probability decreases more than the risk-neutral default probability, and this is the case when $\mu > r$. Otherwise, an increase in the default risk premium can be associated to a decrease in the stock price, when the risk-neutral default probability increases more than the real-world default probability, and this is the case when $\mu < r$.

Default Risk Premium and Insurance Price

Consider again the two states of the world, *high* and *low*, and a contingent claim security that pays out B^h and B^l in the two states, at the payout date T, respectively. We refer again to the low state as the default. We think the security B as a type of insurance, that pays out more in case of default, that is $B_l > B_h$. We define B_t the price of the insurance at time t, a the discounting rate under the real-world probability, and c the discounting rate under the risk-neutral probability. Hence, at time $t = 0$,

$$B_0 = e^{-a}E_0^P[B_T] = e^{-a}[p_0 B^l + (1 - p_0)B^h], \tag{27}$$

and

$$B_0 = e^{-c}E_0^Q[B_T] = e^{-c}[q_0 B^l + (1 - q_0)B^h], \tag{28}$$

Then,

$$a = -\ln\left(\frac{B_0}{E_0^P[B_T]}\right), c = -\ln\left(\frac{B_0}{E_0^Q[B_T]}\right),$$

and

$$p_t = \frac{B^h - \frac{B_t}{e^{-a}}}{B_l}, q_t = \frac{B^h - \frac{B_t}{e^{-c}}}{B_l}$$

Hence,

$$\frac{p_1}{p_0} = \frac{e^{-a}B^h - B_1}{e^{-a}B^h - B_0}, \frac{q_1}{q_0} = \frac{e^{-c}B^h - B_1}{e^{-c}B^h - B_0}$$

Indeed, when $B_1 > B_0$, then $p_1 > p_0$ and $q_1 > q_0$, provided that $B^l > B_t > B^h$. The opposite holds when $B_1 < B_0$.

Moreover, by equating (6) and (7), we obtain

$$a - c = \ln\left(\frac{E_0^P[B_T]}{E_0^Q[B_T]}\right)$$

.

Hence, if

$$a > c \Leftrightarrow E_0^P[B_T] > E_0^Q[B_T] \Leftrightarrow p(B^l - B^h) > q(B^l - B^h) \Leftrightarrow p > q,$$

as $(B^l - B^h) > 0$. The opposite relationship holds when $a < c$.

Now, the gross rate of growth of the default risk premium $\dot{Z}$ between 0 and 1 is

$$\frac{Z_1}{Z_0} = \frac{e^{-c}B^h - B^l}{e^{-c}B^h - B_0} \cdot \frac{e^{-a}B^h - B_0}{e^{-a}B^h - B^1} = \dot{q} \cdot 1/\dot{p} \tag{29}$$

Therefore, the default risk premium increases, that is $\dot{Z} > 1$, either because $\dot{q} > \dot{p} > 1$ if $B_1 > B_0$, or because $1/\dot{p} > 1/\dot{q} > 1$ if $B_1 < B_0$. In the first case there is an increase in the risk-neutral default probability higher than the increase in the real-world default probability. This picture is associated to an increase in the insurance price. In the second case there is a decrease in the real-world default probability higher than the decrease in the risk-neutral default probability. This picture is associated to a decrease in the insurance price.

Again, by simply rearranging (8), we can show that

$$\dot{Z} > 1 \Leftrightarrow e^{-c}(B_1 - B_0) > e^{-a}(B_1 - B_0)$$

It turns out that $B_1 > B_0$ when $c < a$, and $B_1 < B_0$ when $c > a$. By combining now with the statement following the equation (8), we can conclude that an increase in the default risk premium can be associated to a decrease in the insurance price, when the real-world default probability decreases more than the risk-neutral default probability, and this is the case when $c > a$. Otherwise, an increase in the default risk premium can be associated to a increase in the insurance price, when the risk-neutral default probability increases more than the real-world default probability, and this is the case when $a > c$.

We can finally state the simultaneous and two-fold relationship between default risk premium and stock price, and between the default risk premium and a contingent claim security such as an insurance (e.g., a CDS spread). An increase in the default risk premium can be associated to an increase in the stock price and a decrease in the insurance price, when the real-world default probability has decreased more than the risk-neutral default probability. This result holds with *normal* securities in the context of the classical risk-return framework, that is $\mu > r$ (a risk-averse investor seeks for an excess rate of return on the top of the risk-free rate), and $a < c$ (a risk-averse investor is willing to pay more to hedge against a risky scenario).
On the other hand, an increase in the default risk premium can be associated to a decrease in the stock price and an increase in the insurance price, when the risk-neutral default probability has increased more than the real-world default probability. This result holds with *distressed* securities, that are in a counterintuitive risk-return framework (ie., $\mu < r$, and $a < c$).

3.7.2. Kalman filter and Quasi-Maximum Likelihood Estimation

In a general formulation, with a non-linear relationship between the measurement and the state variables, the state-space model is defined by two sets of equations, the transition and the measurement equation, respectively:

$$X_{i,t+\delta t} = X_{i,t} + c_i + \epsilon_{i,t+\delta t},$$

$$Y_{i,t+\delta t} = \psi(X_{i,t+\delta t}) + u_{i,t+\delta t},$$

where $X_{i,t+\delta t}$ is the i-th observation of the state variable at time $t + \delta t$, c_i is the time-invariant component driving the evolution of the state variable, $\epsilon_{i,t+\delta t}$ is the transition error on the i-th observation of the state variable at time $t + \delta t$. On the other hand, $Y_{i,t+\delta t}$ is the i-th observation of the measurement variable at time $t + \delta t$, ψ is the measurement function which links the observable and the latent variable, and $u_{i,t+\delta t}$ is the measurement error.

For a Gaussian state-space model, under standard assumptions, the discrete Kalman filter is proved to be the minimum mean squared error estimator. However, in the case of non-linear relation between the measurement and the state variable, the classic linear Kalman filter is not longer optimal. One possible solution is to linearize the estimation around the current estimate by using the partial derivatives of the process and measurement functions.[17] To linearize the measurement process, we need to compute the derivatives of ψ with respect to

(a) the state variable: $H_{i,j} = \frac{\partial \psi_i}{\partial X_j}(\tilde{X}_t, 0)$,

where H is the Jacobian matrix of partial derivatives of the generic measurement function $\psi(\cdot)$ with respect to the state variable X, and $\tilde{X}_t$ is the current estimate of the state.

(b) the measurement noise: $\breve{H}_{i,j} = \frac{\partial \psi_i}{\partial \nu_j}(\tilde{X}_t, 0)$,

where $\tilde{H}$ is the Jacobian matrix of partial derivatives of $\psi(\cdot)$ with respect to the noise term ν.

Once the linearization has been completed, we can implement the discrete Kalman filter in the usual steps. First, we need to set the *initial conditions*:

$$\lambda_{i,0} \quad P_{i,0},$$

where $P_{i,t} := var[X_{i,t} - \lambda_{i,t}]$ is the variance of the estimation error, and $\lambda_{i,t}$ is the estimate of the state at time t based on the information available up to time t. Then, the filter implementation

[17] Instances of application of this filter deal in particular with interest rates term structure estimation. Duan and Simonato (1999) implement this technique to estimate an exponential-affine term structure models, while more recently Duffee and Stanton (2012) prove the robustness of the non-linear Kalman filter for a dynamics term-structure model estimation

is based upon two sets of equations, the *predicting* equations, and the *updating* equations, that must be repeated for each time step in the data sample.

- *State Prediction*

$$\lambda_{i,t+\delta t/t} = \lambda_{i,t} + c_i,$$

and

$$P_{i,t+\delta t/t} = P_{i,t} + Q_i,$$

where $\lambda_{i,t+\delta t/t}$ is the estimate of the state at time $t + \delta t$ based on the information available up to time t, and Q_i is the covariance of the transition noise.

- *Measurement Update*

$$\lambda_{i,t+\delta t} = \lambda_{i,t+\delta t/t} + P_{i,t+\delta t/t} H'_{i,t+\delta t} Z^{-1}_{i,t+\delta t} \left(Y_{i,t+\delta t} - \psi(\lambda_{i,t+\delta t/t}) \right)$$

$$P_{i,t+\delta t} = P_{i,t+\delta t/t} - P_{i,t+\delta t/t} H'_{i,t+\delta t} Z^{-1}_{i,t+\delta t} H_{i,t+\delta t} P_{i,t+\delta t/t}$$

$$Z_{i,t+\delta t} = H_{i,t+\delta t} P_{i,t+\delta t/t} H'_{i,t+\delta t} + R_i,$$

where H stands for the Jacobian matrix of partial derivatives of the generic measurement function ψ with respect to the state variable X, $Z_{i,t+\delta t}$ is the covariance matrix of the prediction errors at time $t + \delta t$. The prediction errors are defined as $v_{i,t+\delta t} = Y_{i,t+\delta t} - \psi(\lambda_{i,t+\delta t/t})$, where $Y_{i,t+\delta t}$ is the observation of the measurement variable at time $t + \delta t$.

The parameters that describe the dynamics of the transition and the measurement equations (i.e., *hyperparameters*) are unknown, and need to be estimated.

Let rewrite the state-space model as follows:

$$(y_{t+\delta t}, x_{t+\delta t}) = (x_t, \{\theta\}), \quad \{\theta\} = \{\theta^{(f)}; \theta^{(g)}\}$$

, where $y_{t+\delta t}$ is the observable variable at time $t+\delta t$, $x_{t+\delta t}$ is the state variable at time $t+\delta t$, $\{\theta^{(f)}\}$ is the set of unknown parameters in the transition equation, and $\{\theta^{(g)}\}$ is the set of unknown parameters in the measurement equation. The measurement and transition equations of the system are:

$$g(y_{t+\delta t},\alpha)=\varphi(x_{t+\delta t},\beta)+\epsilon_{t+\delta t},\quad \epsilon_t\backsim\mathcal{N}(0,\sigma_\epsilon^2)$$

$$x_{t+\delta t}=f(x_t,\gamma)+\eta_{t+\delta t},\quad \eta_t\backsim\mathcal{N}(0,\sigma_\eta^2)$$

Then,

$$\{\theta^{(f)}\}=\{\gamma,\sigma_\eta^2\}$$

$$\{\theta^{(g)}\}=\{\alpha,\beta,\sigma_\epsilon^2\}$$

We assume that the nonlinear regression disturbance, ϵ_t, is normally distributed:

$$f(\epsilon_t)=\frac{1}{\sqrt{2\pi\sigma_\epsilon^2}}\exp\left[-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}\right]$$

By transformation of variable, the density of y_t is given by

$$f(y_t)=f(\epsilon_t)\left|\frac{\partial\epsilon_t}{\partial y_t}\right|,\quad \frac{\partial\epsilon_t}{\partial y_t}=\frac{\partial g(y_t,\alpha)}{\partial y_t}$$

Then, the density of y_t is

$$f(y_t)=\frac{1}{\sqrt{2\pi\sigma_\epsilon^2}}\exp\left[-\frac{(g(y_t,\alpha)-\varphi(x_t,\beta))^2}{2\sigma_\epsilon^2}\right]\left|\frac{\partial g(y_t,\alpha)}{\partial y_t}\right|$$

The log-likelihood function for observation t is

$$\ln\Omega_t\left(y_t;\{\theta\}\right)=-\frac{1}{2}\ln(2\pi)-\frac{1}{2}\ln(\sigma_\epsilon^2)-\frac{(g(y_t,\alpha)-\varphi(x_t,\beta))^2}{2\sigma_\epsilon^2}+\ln\left|\frac{\partial g(y_t,\alpha)}{\partial y_t}\right|,$$

and the log-likelihood function for $t=1,2,...,T$ observations (i.e., $\delta t=1$) is

$$\ln\Omega=\sum_{t=1}^{T}\ln\Omega_t\left(y_t;\{\theta\}\right)=-\frac{T}{2}\ln(2\pi)-\frac{T}{2}\ln(\sigma_\epsilon^2)-\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^{T}(g(y_t,\alpha)-\varphi(x_t,\beta))^2$$

$$+\sum_{t=1}^{T} \ln \left| \frac{\partial g(y_t, \alpha)}{\partial y_t} \right|$$

As long as $g(y_t, \alpha) = y_t$, then

$$f(y_t) = f(\epsilon_t) \Rightarrow \ln \Omega_t \left(y_t; \{\theta\}\right) = \ln \Omega_t \left(\epsilon_t; \{\theta\}\right)$$

The last term in the log-likelihood function is equal to zero, and the space of the hyperparameters to be estimated is reduced to:

$$\{\theta^{(f)}\} = \{\gamma, \sigma_\eta^2\}$$

$$\{\theta^{(g)}\} = \{\beta, \sigma_\epsilon^2\}$$

In practice, the iteration of the filter generates a measurement-system prediction error, and a prediction error variance at each step. Under the assumption that measurement-system prediction errors are Gaussian, we can construct the log-likelihood function as follows:

$$\ln \Omega(y_t; \{\theta\}) = \ln \prod_{t=0}^{T-\delta t} p\left(y_{t+\delta t/t}\right) = \sum_{t=0}^{T-\delta t} \ln p\left(y_{t+\delta t/t}\right) =$$

$$= -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=0}^{T-\delta t} \ln |Z_{t+\delta t}| - \frac{1}{2} \sum_{t=0}^{T-\delta t} v_{t+\delta t}{}' Z_{t+\delta t}^{-1} v_{t+\delta t},$$

where N is the number of time steps in the data sample. Finally, this function is maximized with respect to the unknown parameters vector $\{\theta\}$. This is known as the *Quasi-Maximum Likelihood* estimation, in conjunction with the non-linear Kalman filter.

3.7.3. Prices Observation Noise: Minimization by Model Calibration

A useful interpretation of the Kalman filtering is to think of it as an updating process, where we first form a prior preliminary guess about the unobservable state variable and then we perform a correction to this guess. It is a typical Bayesian procedure. The correction depends on how good has been the guess in predicting the next observation, and the weight of the correction to update our guess about the state variable primarily depends on how confident we are about the reliability of the observation. It is measured by the variance of the errors in the measurement equation, which links the state and the observable variable.

It follows that we are more confident about the observation of the CDS-PD and the equity prices when lower is the variance of the measurement errors. In the extreme case, when the error associated with the CDS-PD measurement equation has zero variance and the equity equation has high measurement error variance, we take into account only the CDS prices to reconstruct the dynamics of the market value of the leverage, and vice versa. We interpret this variance as a measure of the noise in the observation of the market data. This noise may be a consequence of the market microstructure effects, the model misspecification, or even the recording price procedure for the CDS spreads, which are traded on an OTC market, and the quotes may vary according to the data provider. It turns out that the impact of the market data on the estimated hyperparameters and state variable dynamics is inversely proportional to the observation noise measured by the measurement error variance.

In principle, we could impose an arbitrary prior on the variance of the measurement error either of the CDS-PD equation or the equity equation according to the weight that we want to assign to a particular market variable. It would be the case if we have an ex-ante conjecture on the informativeness of the credit or the equity market. However, we do not impose any prior constraint, and we include the variance of the measurement errors in the hyperparameters that we estimate, by using the quasi-maximum likelihood algorithm based on the non-linear filtering. Then, in a second stage, we implement a non-linear squares minimization to calibrate the values of the measurement error variances associated with the risk-neutral PD and the equity pricing equation, in order to fit simultaneously the observed dynamics of the risk-neutral PD implied by CDS spreads, and the equity value.

Driessen (2005) adopts a similar algorithm to derive the risk-premium component which translates the real-world probability measure of default, extracted from rating agency data, into the risk-neutral intensity. In practice, he performs a non-linear squares minimization by using the observed and the model-implied default probabilities, after estimated the hyperparameters with a Kalman filter. We search the value of the measurement errors variances which minimize the sum of the squared differences between the observed and the model-implied market data, where the sum is calculated over the sample time series.

$$\min_{R_i,\omega_i} \sum_{t=1}^{T} \alpha_{i,t}^2,$$

where $\alpha_{i,t} = \begin{bmatrix} \hat{PD}^{Q}_{i,t}(\tau_a) - PD^{Q}_{i,t}(\tau_a) \\ \hat{E}_{i,t} - E_{i,t} \end{bmatrix}$

Finally, we iterate again the updating and the predicting equations of the non-linear Kalman filter to reconstruct the dynamics of the firm's market value of the leverage. Now, we do not perform the quasi-maximum likelihood estimation as we have already obtained the hyperparameters in the first step, and the measurement errors variances in the second stage. The result is the dynamics of the market value of the leverage which minimizes the distance between the implied and the observed market variables. It may be interpreted as the 'best-we-can-do' to reduce the noise in the prices observation, due to the model misspecification. We can attribute the residual error to the systematic noise implicitly contained in the market quotes observation.

Table 14. Summary Statistics, CDS Spreads

		1 year						3 years					
STAT		**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Mean		**86.69**	**140.74**	**43.36**	**39.80**	**41.74**	**18.81**	**107.93**	**151.43**	**69.40**	**72.73**	**84.67**	**52.12**
Standard Deviation	**TS**	71.84	88.29	11.21	19.18	11.12	3.61	65.10	78.97	11.23	23.48	12.13	7.63
	CS	166.38	300.91	37.34	27.09	48.26	30.79	167.02	274.38	50.89	40.43	70.35	44.10
Median	**TS**	60.38	107.44	40.92	28.08	39.78	18.83	86.00	121.56	65.95	57.75	81.14	52.18
	CS	56.81	83.70	34.34	32.98	27.20	13.38	75.61	93.73	54.98	64.14	62.78	41.50

		5 years						10 years					
STAT		**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Mean		**121.91**	**156.27**	**93.87**	**106.15**	**122.18**	**92.24**	**128.75**	**148.60**	**111.62**	**136.59**	**154.57**	**136.96**
Standard Deviation	**TS**	60.49	70.16	11.36	24.68	12.45	9.24	54.42	58.39	9.79	25.19	12.13	7.74
	CS	159.98	251.92	60.66	52.03	85.45	63.83	143.89	216.25	63.77	58.13	88.68	74.67
Median	**TS**	102.52	127.58	93.34	90.71	118.41	92.56	112.20	119.15	111.89	125.61	154.81	138.43
	CS	89.88	101.93	75.55	94.53	98.07	74.61	99.54	101.27	91.88	122.19	131.77	117.72

The table reports the summary statistics of the CDS spreads for all the 164 firms in the sample, and for the four time horizons. The statistics are calculated over two dimensions, the time series (TS), and the cross section (CS). The years refer to the time period going from the 20th December of the previous year to the 19th December of that year. The CDS spreads are expressed in basis points as annualized percentage of the notional value of the transaction.

Table 15. Summary Statistics: Market Capitalization

STAT		Australia 2008	2009	2010	2011	2012	2013	Japan 2008	2009	2010	2011	2012	2013
Mean		**16106.98**	**14269.40**	**15830.27**	**15558.27**	**16615.72**	**20973.81**	**3306770.16**	**2187244.71**	**2374068.84**	**2312945.03**	**2354328.36**	**3650067.03**
Standard Deviation	**TS**	1388.69	1082.65	336.20	330.65	1026.57	934.35	702575.93	249284.28	167946.76	210075.19	150128.26	404024.02
	CS	16552.35	13902.67	14003.61	14192.08	16159.59	20936.58	4449327.34	3234826.62	3015261.81	2841162.46	2981459.13	5360640.13
Median	**TS**	16079.42	13988.59	15836.51	15630.89	16327.32	21121.01	3482359.81	2250207.16	2354792.22	2350307.91	2327871.27	3825030.35
	CS	6971.62	5591.93	7826.98	8502.39	8996.27	10797.14	1594845.22	1047443.75	1301167.07	1355906.53	1270331.88	1559423.72

STAT		United States 2008	2009	2010	2011	2012	2013	Europe 2008	2009	2010	2011	2012	2013
Mean		**31736.19**	**24972.26**	**29632.92**	**31879.01**	**33599.42**	**38801.01**	**37476.06**	**29748.36**	**36045.39**	**36186.81**	**35124.99**	**39957.96**
Standard Deviation	**TS**	4395.72	2568.67	1269.05	1566.59	815.32	1945.23	6002.41	3337.60	1556.23	3476.93	1597.79	1934.45
	CS	38637.15	32837.07	35168.41	36649.62	41794.40	44666.25	42760.72	33453.42	42475.56	42916.63	42344.26	47475.09
Median	**TS**	33456.44	24420.28	29546.20	32264.76	33676.55	39087.60	39113.69	28951.05	35924.19	37473.11	35617.56	39421.99
	CS	18109.02	14066.85	17251.33	18248.14	18756.83	22731.48	21511.27	16238.91	20440.47	19406.24	18639.01	22046.98

The table reports the summary statistics of the market capitalization across regions. The same structure in terms of time intervals as for the CDS table applies. The market capitalization is expressed in millions of the local currency.

Table 16. Empirical Results - Parameters Estimation

			Volatility	Barr/FV	Drift	Debt	Barrier
Regions	**AUS**	**Mean**	0.19	-0.39	-0.08	17119	11746
		St. Dev.	0.03	0.06	0.04	17378	12133
	JAP	**Mean**	0.19	-0.30	-0.10	841215	1137242
		St. Dev.	0.04	0.10	0.09	676442	474817
	USA	**Mean**	0.08	-0.21	-0.04	71826	86770
		St. Dev.	0.02	0.08	0.02	84963	99737
	EUR	**Mean**	0.15	-0.34	-0.06	33972	47644
		St. Dev.	0.03	0.05	0.04	40132	55945

The table reports the mean and the standard deviation of real data estimation on 164 firms divided, respectively, in four geographic regions, nine GICS sectors, and three Rating classes, for the unknown parameters σ_{L_i}, K_i, μ_{L_i}, F_i, C_i, R_i, ω_i, by using the univariate non-linear Kalman filter in conjunction with quasi-maximum likelihood estimation. The statistics for F_i, C_i are reported only for the four regions as the respective values are expressed in local currency.

Table 17. Goodness of Fit

	Equity			5-Y CDS			10-Y CDS		
	SME	Merton	Random	SME	Merton	Random	SME	Merton	Random
Mean	0.011***	-4.76E-05	5.41E-05	0.008***	0.005***	-3.00E-05	0.007***	4.00E-03***	-4.00E-05
t-test	48.47	-0.28	1.03	24.46	16.16	-0.4	21.92	15.88	-0.58

The table reports the mean, and the t-test statistic, for the daily performance of the long-short portfolio trading strategy described in section 5, for the Equity and the CDS portfolios. The stars on the mean stand for the level of significance according to the t-test outcome (*** = 99% significance level, ** = 95%, * = 90%). The SME columns refer to the portfolio strategy constructed by using the barrier-dependent model estimation, the Merton columns report the performance achieved by using the estimation with the Merton model, while the Random columns refer to the maximum performance achieved across the 1000 randomly constructed portfolios.

Table 18. The Default Risk Premium

	Real-World PD				Risk-Neutral PD				Default Risk Premium			
	1Y	**3Y**	**5Y**	**10Y**	**1Y**	**3Y**	**5Y**	**10Y**	**1Y**	**3Y**	**5Y**	**10Y**
AUS	0.001	0.031	0.071	0.166	0.001	0.033	0.073	0.179	1.27	1.25	1.18	1.13
JAP	0.000	0.020	0.051	0.139	0.000	0.029	0.074	0.186	2.38	2.29	2.05	1.70
USA	0.000	0.006	0.013	0.032	0.001	0.028	0.068	0.173	8.78	10.03	10.35	8.28
EUR	0.000	0.023	0.063	0.181	0.001	0.035	0.082	0.186	2.02	1.88	1.53	1.02

The table reports the mean over the entire time series of the estimated median-firm risk-neutral and real-world default probability, and default risk-premium, across the four time horizons (1 year, 3,5, and 10 years), and the four geographic regions.

Table 19. Correlation DRP - Market Data

	5-th	**25-th**	**50th**	**75-th**	**95th**
corrEQ	-0.95	-0.28	0.31	0.88	0.89
corrCDS	0.77	0.08	-0.35	-0.49	-0.44

	DRP-Equity		**DRP-CDS**	
	DRP<1	DRP>1	DRP<1	DRP>1
Mean	-0.76	0.68	0.69	-0.03
t-test Diff	-33.39		15.11	

The upper table reports the sample correlation between the default risk premium and the equity prices, and the 10-years CDS spread, respectively, in correspondence of different firm-percentiles of the default risk premium distribution, after sorting the firms according to the average default risk premium. The lower table reports the mean correlation, across the high and low default risk premium sub-samples, between the default risk premium and the market variables. Moreover, the table reports the t-test outcome on the difference between the mean correlations across the high and low premium samples, for each market variable.

Table 20. Trading Strategy - Long-Short Portfolios

Correct Trading								
	All Firms				**Excluding Outliers**			
	In-Sample		**Out-of-Sample**		**In-Sample**		**Out-of-Sample**	
	Equity	CDS	Equity	CDS	Equity	CDS	Equity	CDS
Mean	0.0015***	6.29E-04***	1.98E-04**	2.55E-04**	7.47E-04***	4.03E-04***	2.76E-04***	2.50E-04***
t-test	16.89	5.34	2.31	2.34	8.29	3.64	3.17	2.49
Opposite Trading								
	All Firms				**Excluding Outliers**			
	In-Sample		**Out-of-Sample**		**In-Sample**		**Out-of-Sample**	
	Equity	CDS	Equity	CDS	Equity	CDS	Equity	CDS
Mean	-0.0023***	-0.0011***	1.75E-04	-6.73E-04**	-8.47E-04***	-8.21E-04***	8.24E-04	-4.97E-04*
t-test	-12.72	-3.65	0.91	-2.13	-5.11	-2.82	0.45	-1.64

The table reports the mean, and the t-test statistic, for the daily performance of the long-short portfolio trading strategy described in section 6, for the Equity and the CDS portfolios. The stars on the mean stand for the level of significance according to the t-test outcome (*** = 99% significance level, ** = 95%, * = 90%). The table reports the results for the in-sample, and the out-of-sample approach, as described in the paper. The upper part of the table reports the results obtained by following the trading strategy scheme as described in the paper, while the lower part reports the results obtained by trading in the opposite direction. The left-side reports the results for the trading strategy performed by using all the sample firms, while the right-side reports the results for the trading strategy performed after excluding the firms with default risk premium higher than 1000.

Admati, A., P. Pfeiderer, and J. Zechner, 1994, Large shareholder activism, risk sharing, and financial market equilibrium, *Journal of Political Economy* 102, 1097–1130.

Albuquerque, R., and E. Schroth, 2010, Determinants of the block premium and of private benefits of control, *Journal of Financial Economics* 96, 33–55.

Albuquerque, R., and E. Schroth, 2015, The value of control and the costs of illiquidity, *Journal of Finance* 70, 1405–1455.

Almeida, H., and T. Philippon, 2007, The risk-adjusted cost of financial distress, *Journal of Finance* 62, 2557–2586.

Anginer, D., and C. Yildizhan, 2017, Is there a distress risk anomaly? pricing of systematic default risk in the cross section of equity returns, *Review of Finance* 22, 633–660.

Arce, O., S. Mayordomo, and J. Pena, 2013, Credit-risk valuation in the sovereign cds and bonds markets: Evidence from the euro area crisis, *Journal of International Money and Finance* 35, 124–145.

Aretz, K., C. Florackis, and A. Kostakis, 2018, Do stock returns really decrease with default risk? new international evidence, *Management Science* 64, 3821–3842.

Bai, J., and P. Collin-Dufresne, 2013, The cds-bond basis, *Financial Management* .

Barclay, M., and C. Holderness, 1989, Private benefits from control of public corporations, *Journal of Financial Economics* 25, 371–395.

Benos, E., and M.S. Weisbach, 2004, Private benefits and cross-listings in the united states, *Emerging Markets Review* 5, 217–240.

Berndt, A., R. Douglas, D. Duffie, M. Ferguson, and D. Schranz, 2008, Measuring default risk premia from default swap rates and edfs, Working Paper*, Tepper School of Business at Carnegie Mellon University* .

Black, F., and J. Cox, 1976, Valuing corporate securities: some effects of bond indenture provisions, *Journal of Finance* 31, 351–367.

Blanco, R., S. Brennen, and I. Marsh, 2005, An empirical analysis of the dynamic relationship between investment grade bonds and credit default swaps, *Journal of Finance* 60, 2255–2281.

Bohn, J., and P. Crosbie, 2003, Modeling default risk, *Moody's KMV* .

Bolton, P., and E.L. von Thadden, 1998, Blocks, liquidity, and corporate control, *Journal of Finance* 53, 1–25.

Brockman, P., and H.J. Turtle, 2003, A barrier option framework for corporate security valuation, *Journal of Financial Economics* 67, 511–529.

Burkart, M., D. Gromb, and F. Panunzi, 2000, Agency conflicts in public and negotiated transfers of corporate control, *Journal of Finance* 55, 645–677.

Campbell, J., J. Hilscher, and J. Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Chan, W. S., 2003, Stock price reaction to news and no-news: drift and reversal after headlines, *Journal of Financial Economics* 70.

Chava, S., and A. Purnanandam, 2010, Is default risk negatively related to stock returns?, *Review of Financial Studies* 23, 2523–2559.

Cochrane, J.H., and M. Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Collin-Dufresne, P., and V. Fos, 2015, Do prices reveal the presence of informed trading?, *Journal of Finance* 70, 1555–1582.

Corradin, S., and M. Rodriguez-Moreno, 2014, Limits to arbitrage: empirical evidence from euro area sovereign bond markets, *Working Paper* .

Dammon, R.M., K.B. Dunn, and S.S. Chester, 1993, The relative pricing of high-yield debt: the case of rjr nabisco holdings capital corporation, *American Economic Review* 83, 1090–1111.

Davydenko, S.A., 2012, When do firms default? a study of the default boundary, *Working Paper* .

DeMarzo, P.M., and B. Urosevic, 2006, Ownership dynamics and asset pricing with a large shareholder, *Journal of Political Economy* 1, 774–815.

Dhillon, A., and S. Rossetto, 2014, Ownership structure, voting, and risk, *Review of Financial Studies* 28, 521–560.

Diaz, A., J. Groba, and P. Serrano, 2013, What drives corporate default risk premia? evidence from the cds market, *Journal of International Money and Finance* 37, 529563.

Donelli, M., B. Larrain, and I. Urzsua, 2013, Ownership dynamics with large shareholders: An empirical analysis, *J. Financial Quantitative Analysis* 48, 579–609.

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

Duan, J.C., 1994, Maximum likelihood estimation using price data of the derivative contract, *Mathematical Finance* 4, 155–157.

Duan, J.C., and J.G. Simonato, 1999, Estimating exponential-affine term structure models by kalman filter, *Review of Quantitative Finance and Accounting* 13, 111–135.

Duffee, G.R., and R.H. Stanton, 2012, Estimation of dynamic term structure models, *The Quarterly Journal of Finance* 2, 1–54.

Dyck, A., and L. Zingales, 2004, Private benefits of control: an international comparison, *Journal of Finance* 59, 537–600.

Eisedorfer, A., A. Goyal, and A. Zhdanov, 2017, Distress anomaly and shareholder risk: international evidence, *Financial Management* 47, 553–581.

Ericsson, J., and J. Reneby, 2004, An empirical study of structural credit risk models using stock and bond prices, *Journal of Fixed Income* 13, 38–49.

Faccio, M., and L.H.P. Lang, 2002, The ultimate ownership of western european corporations, *Journal of Financial Economics* 65, 365–395.

Faccio, M., M.T. Marchica, and R. Mura, 2011, Large shareholder diversification and corporate risk-taking, *Review of Financial Studies* 24, 3601–3641.

Ferreira Filipe, S., T. Grammatikos, and M. Dimitra, 2016, Pricing default risk: The good, the bad, and the anomaly, *Journal of Financial Stability* 26, 190–213.

Fleckenstein, M., 2013, The inflation-indexed bond puzzle, *Working Paper* .

Fleckenstein, M., F.A. Longstaff, and H. Lustig, 2014, The tips-treasury bond puzzle, *Journal of Finance* 69, 2151–2197.

Foley-Fisher, N., 2010, Explaining sovereign bond-cds arbitrage violations during the financial crisis 2008-09, *Working Paper* .

Fontana, A., 2010, The persistent negative cds-bond basis during the 2007/08 financial crisis, *Working Paper* .

Fontana, A., and M. Scheicher, 2016, An analysis of euro area sovereign cds and their relation with government bonds, *Journal of Banking and Finance* 62, 126–140.

Forte, S., 2011, Calibrating structural models: A new methodology based on stock and credit default swap data, *Quantitative Finance* 11, 1745–1759.

Forte, S., and L. Lovreta, 2012, Endogenizing exogenous default barrier models: The mm algorithm, *Journal of Banking and Finance* 36, 1639–1652.

Friewald, N., C. Wagner, and J. Zechner, 2014, The cross-section of credit risk premia and equity returns, *Journal of Finance* .

Gao, P., C.A. Parsons, and J. Shen, 2017, The global relation between financial distress and equity returns, *Review of Financial Studies* 31, 239–277.

Gapen, M., D. Gray, C.H. Lim, and Y. Xiao, 2011, Measuring and analyzing sovereign risk with contingent claims, *IMF Staff Papers* 55, 109–148.

Garlappi, L., and H. Yan, 2011, Financial distress and the cross-section of equity returns, *Journal of Finance* 66, 789–822.

Gomes, A., 2000, Going public without governance: managerial reputation effects, *Journal of Finance* 55, 615–645.

Guiso, L., P. Sapienza, and L. Zingales, 2018, Time-varying risk aversion, *Journal of Financial Economics* 128, 403–421.

Gyntelberg, J., P. Hordahl, K. Ters, and J. Urban, 2017, Arbitrage costs and the persistent non-zero cds-bond basis: Evidence from intraday euro area sovereign debt markets, *BIS Working Paper* 631, 1–45.

Hansen, B., and B. Seo, 2002, Testing for two-regime threshold cointegration in vector error-correction models, *Journal of Econometrics* 110, 292–318.

Hilli, A., D. Laussel, and N. Van Long, 2013, Large shareholders, monitoring, and ownership dynamics: Toward pure managerial firms?, *Journal of Econimic Dynamics and Control* 37, 666–679.

Huang, J.Z., and M. Huang, 2012, How much of the corporate-treasuty yield spread is due to credit risk?, *Review of Asset Pricing Studies* 2, 153–202.

Hull, J., M. Predescu, and A. White, 2004, The relationship between credit default swap spreads, bond yields, and credit rating announcements, *Journal of Banking and Finance* 28, 2789–2811.

Hull, J., M. Predescu, and A. White, 2005, Bond prices, default probabilities, and risk premiums, *Journal of Credit Risk* 1, 53–60.

Jarrow, R., H. Li, X. Ye, and M. Hu, 2018, Exploring mispricing in the term structure of cds spreads, *Review of Finance* 23, 161–198.

John, K., L.P. Litov, and B.Y. Yeung, 2008, Corporate governance and risk taking: Theory and evidence, *Journal of Finance* 63, 1607–1728.

Jones, E., S. Mason, and E. Rosenfeld, 1984, Contingent claims analysis of corporate capital structures: An empirical investigation, *Journal of Finance* 39, 611–625.

La Porta, R., F. Lopez-De Silanes, and A. Shleifer, 1999, Corporate ownership around the world, *Journal of Finance* 54, 471–517.

Laeven, L., and R. Levine, 2008, Complex ownership structures and corporate valuations, *Review of Financial Studies* 21, 579–604.

Lamont, O.A., and R.H. Thaler, 2003, Can the market add and subtract? mispricing in tech stock carve-outs, *Journal of Political Economy* 111, 227–268.

Li, K.L., and Y.H. Wong, 2008, Structural models of corporate bond pricing with maximum likelihood estimation, *Journal of Empirical Finance* 15, 751–777.

Lippi, F., and F. Schivardi, 2014, Corporate control and executive selection, *Quantitative Economics* 5, 417–456.

Longstaff, F., S. Mithal, and E. Neis, 2005, Corporate yield spreads: Default risk or liquidity? new evidence from the credit default swap market, *Journal of Finance* 60, 2213–2253.

Longstaff, F.A., and E.S. Schwartz, 1995, A simple approach to valuing risky fixed and floating rate debt, *Journal of Finance* 50, 789–819.

Merton, R., 1974, On the pricing of corporate debt: the risk structure of interest rates, *Journal of Finance* 29, 449–470.

Nenova, T., 2003, The value of corporate voting rights and control: a cross-country analysis, *Journal of Financial Economics* 68, 325–351.

Nicodano, G., and A. Sembenelli, 2004, Private benefits, block transaction premiums, and ownership structure, *International Review of Financial Analysis* 13, 227–244.

Odegaard, B.A., 2016, The diversification cost of large, concentrated equity stakes. how big is it? is it justified?, *Finance Research Letters* 6, 56–72.

O'Kane, D., and S. Turnbull, 2003, Valuation of credit default swaps, *Fixed Income Quantitative Credit Research - Lehman Brothers* .

Perlich, C., and A. Reisz, 2007, A market-based framework for bankruptcy prediction, *Journal of Financial Stability* 3, 85–131.

Roger, P., and A. Schatt, 2016, Idiosyncratic risk, private benefits, and the value of family firms, *Finance Research Letters* 17, 235–245.

Ronn, E.I., and A.K. Verna, 1986, Pricing risk-adjusted deposit insurance: An option-based model, *Journal of Finance* 41, 871–895.

Rossetto, S., and R. Staglian, 2018, Ownership concentration and firm risk. evidence from the us, *Working Paper* .

Schneider, P., C. Wagner, and J. Zechner, 2019, Low risk anomalies?, *Working Paper* .

Vassalou, M., and Y. Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831–868.

Wong, H.Y, and T.W. Choi, 2009, Estimating default barriers from market information, *Quantitative Finance* 9, 187–196.

Zingales, L., 1995, The value of the voting right: A study of the milan stock exchange experience, *Review of Financial Studies* 7, 125–148.

Zwiebel, J., 1995, Block investment and partial benefits of control, *Review of Economic Studies* 62, 161–185.