4

Banco de Portugal Economic Studies

Volume IV

BANCO DE PORTUGAL
EUROSYSTEM

4

Banco de Portugal Economic Studies

Volume IV

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal
T +351 213 130 000 | estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisboa, 2018 • www.bportugal.pt

Banco de Portugal Economic Studies | Volume IV – no. 4 | Lisbon 2018 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt • Edition Economics and Research Department • Design Communication and Museum Department | Design Unit • ISSN (*online*) 2183-5217

Content

Editorial

Editorial

October 2018

The fourth issue of Banco de Portugal Economic Studies for 2018 contains three essays dealing with unconventional monetary and fiscal policies, credit markets and equity markets. The first essay, by Isabel Correia, is titled "From unconventional monetary to unconventional fiscal policies". The essay reflects on the changes that have occurred in monetary policies and how they may signal future changes in the design and use of fiscal policies.

The last decade of the last century and the earlier years of the current saw positive economic outcomes in most developed countries including declines in the average inflation rates, positive economic growth rates and low volatility of real aggregates. These outcomes were associated with deep changes in monetary policy. These included an increase in central bank independence and a movement away from discretionary interventions (and their associated effects on the volatility and uncertainty in the economy) towards credibility-enhancing, rules-based policies such as forward guidance. Developments in fiscal policy practice and research lagged behind those in monetary policy, despite some pioneering ideas from economists like Friedman or Solow on the central role that automatic stabilizers should play.

With the Great Recession, fiscal stimulus came back as the main way to stabilize the economy, with large increases in government expenditures that eventually left a few countries with high levels of debt. At the same time, monetary policy became exceptionally expansionary. Central banks started large-scale purchases of private and public assets, stimulating the economy by providing liquidity to financial markets and easing credit, lowering risk premia and driving interest rates down to the zero lower bound or even below. The consequences included changes in asset prices and redistributive effects, making these unconventional policies somewhat similar to fiscal policies.

Research carried out on the role that nominal rigidities and other distortions play in the way the economy reacts to shocks and to the policies that deal with those shocks deepened the connections between the two policies, by exposing the comparability of monetary and fiscal stabilization policy channels.

The area under research with probably the largest potential to generate improvements is the use of unconventional fiscal policies to achieve results traditionally pursued by monetary policies. The fundamental idea is that whatever monetary policy can achieve with the nominal interest rate that can also be done with a combination of consumption, labor and capital income taxes. The case where policy rates are constrained by the lower bound is interesting because state-contingent taxation of consumption can lead to desired inflation on the consumer side. A temporarily lower consumption tax relative to the future generates inflation in consumer prices and thus negative real interest rates without having significant effects on producer prices and thus avoiding the distortions associated with producer price inflation. The increase in consumption taxes would be counteracted by lowering taxes on labor and capital income. In this way the use of those three taxes can replicate the decline of the "tax" on money, that is the nominal policy interest rate. Other areas where unconventional fiscal policies have substantial potential for improvement are discussed, as in the case of credit subsidies.

The implementation of these types of policies requires unprecedented levels of flexibility from tax policy as well as stronger levels of coordination across institutions. So far, the experience of using these fiscal policies in Japan is still incomplete, so it is too soon for an evaluation. But the potential gains are large. All in all, these new policies may allow us to deal with aggregate shocks in a potentially much more efficient way compared with current policies.

The second essay, by Paulo Júlio and José Maria is entitled "Interest rate spreads hikes: What lies behind them?" In recent work, the authors have developed a novel new Keynesian Dynamic Stochastic General Equilibrium (DSGE) macroeconomic model for a small, open euro area economy. There are several agents in the model: households (who can be workers, entrepreneurs or bankers), foreign agents, the Government, entrepreneurial firms, banks and several industries (intermediate goods, final goods, capital goods, and retailers). Markets have imperfect competition and are endowed with important frictions in the form of staggered price-setting à la Calvo. There are industry specific quadratic adjustment costs for investment and for labor hours. Monetary policy is exogenous and unresponsive to domestic developments, a consequence of the small open economy framework.

The model has a detailed structure for the banking system. Banks mostly lend to entrepreneurs for the purposes of financing capital goods. Negative idiosyncratic shocks can drive entrepreneurs to default, but the model rules out bank failure for simplicity. Credit supply decisions are simultaneously driven by regulatory capital requirements, defaulted loans, and credit restrictions that may become binding under severe shocks. Hence, the model is capable of triggering asymmetric responses to potentially symmetric shocks. Bankers manage defaulted loans over time, optimally recognizing impairment losses. The model was calibrated for the specific case of Portugal. The model allows for a decomposition of the overall interest rate spread into the sum of three components: a capital requirements-driven spread (interest margin required to cover the expected costs of a possible violation of capital requirements), the credit restrictions-driven spread (interest margin induced by credit supply restrictions emerging when banks' value declines), and the retail-driven spread (the margin required by retail banks to cover the expected losses generated by corporate bankruptcy). The authors then simulate the consequences of three types of negative transitory shocks: a demand side shock (on public consumption), a supply side shock (on technology), and a financial shock (on risk).

The demand shock leads to transitory increases in the capital requirements spread and retail spread but leaves the credit restrictions spread unaffected. In the case of the supply shock both the capital requirements and the retail spreads come down as the shock is inflationary and firms have lower real costs of financing. As in the previous case the credit restrictions spread is unaffected. Finally, in the case of a financial shock, the hike in default rates triggers an increase in all three spreads, with the novelty that in this scenario the credit restriction-driven spread increases significantly. The risk shock leads to more expensive credit (via the retail-driven spread) as greater margins on performing loans are required to cover for higher bankruptcy probability. The concomitant accumulation of defaulted loans impacts bank returns and affect the ability of bankers to keep capital within the regulatory requirements, leading to a larger capital requirements-driven spread. Finally, the collapse in banks' value hinders credit supply and triggers the aforementioned hike in the credit restrictions-driven spread. More expensive credit pushes corporate loans down, resulting in fewer investments and less capital accumulation. The inertia in the banking system comes down slowly as time passes and thus the protracted recession takes time to wind down.

The third and last paper, by Nuno Silva, is titled "Monitoring the equity risk premium in the S&P500". In the last few years we have seen very low yields for bonds and a large growth of equity markets as measured by SP 500 and other indices. If investor growth expectations or required rates of return fluctuate with the business cycle, does the current situation precedes some sort of price correction? The answers to this question are certainly relevant to investors but also to those responsible for economic policy. With this motivation in mind, the goal of this essay is to measure the evolution of the equity risk premium implied by stock prices. This decomposition may help macroprudential authorities understand whether stock prices are fairly priced, and evaluate the risks to the financial system.

A starting point for the analysis is the traditional model where the value of a stock is given by adding up discounted future cash flows. The model, taught in any basic finance course, assumes a constant discount rate, including a risk premium component, and a constant growth rate for cash flows. The value of stocks is then given by the ratio of current smoothed cash flows to the difference between the discount rate and the cash flow growth rate. This model has some pedagogical virtues but it has serious limitations. Equity values are very sensitive to changes both on the discount rate and on the growth rate, making it easy to end up with bad estimates. Furthermore, the models used in practice typically ignore default risk and the effect on equity valuation of leverage dynamics. What this essay does is basically to replace the components of the traditional discounted cash flow model by more sophisticated components that accommodate explicitly the stochastic nature of the financial data involved and face up to the limitations mentioned earlier.

In the model the firms' cash flows are modeled as a geometric Brownian motion, the same stochastic process Black and Scholes used to model stock prices. In any given period the cash flows may be inferior to fixed costs, interest expenses and net borrowing. If that happens the shareholders must decide whether they are willing to inject capital in the firm, which they will do as long as the equity value after the capital increase is higher than the amount of cash flow they inject. Otherwise they give up and the firm defaults, closes and distress costs are incurred. These correspond to legal costs and value destruction caused by fire sales and loss of intangible value. The data for the calibration of the model comes from 1998 up to 2017. It includes annual accounting data and monthly market data for the 205 non-financial firms in the S&P500 that were continuously in the index throughout the years studied. Data from Credit Default Swaps (CDS) is also used. The yield on 30-year U.S. Treasury bonds is used as the risk-free rate and also as one of the estimates of the long-run corporate growth rate. The alternative estimate is based on the compounded growth rate of the 3 to 5-year forecasts by analysts (after some suitable adjustments given their excessive optimism).

The main results of the analysis are that the mean equity risk premium for the period between 1999 and 2017 was approximately 5.9%. Currently, the equity risk premium is in a downward trend reaching 4.6% by the end of 2017 when the risk free rate is used, and 5.6% when analysts' forecasts are used. The 4.6% is near the minimum value in its series for the sample period used, whereas the 5.2% is just under the mean value for it series. Maybe the difference comes from analysts' current optimism or maybe it comes from the abnormally low level of long-term interest rates.

What to make of these results? This seems to be a case where the proverbial (analytical) buck does not stop here, as a judgment on the sustainability of stock prices hinges on what will happen to other fundamentals. If the low long-term interest rates are here to stay, as embodied in the forecasts of analysts, that would indicate current stock prices are at sustainable levels. Otherwise some corrections may be on the way.

From unconventional monetary to unconventional fiscal policies

Isabel Correia
Banco de Portugal, Catolica Lisbon
SBE and CEPR

October 2018

"...the ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed the world is ruled by little else. Practical men, who believe themselves to be exempt from any intellectual influences, are usually the slaves of some defunct economist."

Keynes, in The General Theory

Introduction

The recent crisis has brought economic policy to the center of the public debate. If during the last decade reality shaped policy, at the end of last century we witnessed a period in which the interaction between economic policy and theory was stronger than ever. As usual, in practice causality runs in both directions, but nonetheless during this period this relation was closer and in large part very much driven by investments in research by central banks. Researchers worked on problems motivated by specific policy questions and specific policy experiences, and policy makers made use of theory to shape institutions, design rules, or simply to communicate policy decisions to the public. In contrast to Keynes´s quote above, lags between new theoretical results and their introduction into everyday policy making diminished as economists worked in close connection with policy makers.

To better understand this change we will focus on monetary policy. After strong changes in the 70's and 80's monetary policy recovered its glamour

---

Disclaimer: The opinions expressed in the article are those of the author and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. This article is based on a presentation made at the Riksbank conference "Rethinking the Central Bank's Mandate", and it is also available in the Sveriges Riksbank Economic Review", 2016-3 Special Issue.

E-mail: mihcarvalho@bportugal.pt

during the 90's, largely due to the positive economic outcomes perceived as the result of new ways of designing institutions and of using monetary policy instruments. We can summarize these outcomes as: a strong decline in the average inflation rate during that period, positive average growth in most economies, as well as low volatility of real aggregates over the business cycle frequency. Additionally, these results were not limited to some small area of the world but were widespread across developed countries. Trying to trace these results to a common trend, lessons from research suggest that the biggest change was the push for central banks to be independent of political power and the strong movement toward rule-based monetary policy.

By 2002, 22 countries had adopted monetary frameworks that emphasized inflation targeting as one, or the main, objective of its mandate.

This can be seen as the success of stabilization policy in normal times, and the widespread feeling was that the big hero was monetary policy. This idea was bolstered by the image of the central bank as an independent institution with a very concrete goal (low inflation), a very precise instrument (the short term policy rate), a rule and clear communication rule that governed the decisions on this instrument, and the credibility that came as a by-product of this institutional arrangement. It is therefore probably fair to say that the pre-crisis stability and growth was to a large extent explained by this new monetary design.

At the same time on the research front we went from showing that monetary policy shocks should be avoided since they just introduce volatility and uncertainty in the economy, to showing how good monetary policy should play a stabilizing role. This revealing step came from the ability to extend the traditional framework used to define optimal tax policy to monetary environments. In this way researchers could support the robustness of an average low inflation, and, at the same time, explain the gains of using monetary policy as a stabilization policy device, namely explaining how the so called "gaps" should be smoothed across time and states.

In contrast, developments in fiscal policy were often far from those reached in monetary policy, both in terms of real institutional design and research discussion. We should however remember that Milton Friedman (1948) earlier than anyone else wrote against the use of discretionary (both fiscal and monetary) policy to stabilize the business cycle. There Friedman instead defended the power of fiscal automatic stabilizers as a better instrument for countercyclical policy. In a sense, we now can recognize that Friedman was very ahead of his time. By recognizing that "changes in fiscal incentives may be more useful than traditional discretionary fiscal policies that increase budget deficits and work through income effects alone," Friedman launched the foundations of what later would be named unconventional fiscal policies, which are the main topic of this note.

From conventional to unconventional monetary policy

Solow (2005) return to the defense of automatic stabilizers as a way of designing efficient fiscal policies that could stabilize the economy. However, these remarks didn´t have an impact on research and the way they were transmitted to the policy arena is in no way comparable to the above description of the changes in the conduct of monetary policy. As a result, we arrived to the financial crisis with a consensus that stabilization policy was the responsibility of monetary institutions and that fiscal policies should be designed with built in automatic stabilizers. Additionally, fiscal policy should not amplify the cycle and there was some common understanding that it should satisfy some rules that would guarantee the sustainability of public debt.

In pre-crisis thinking of policy economists, there was a clear division of institutions, instruments, and objectives between these two set of policies. Still we can find in the research community some areas of intersection between monetary and fiscal policies, namely in the literature known as the Fiscal Theory of the Price Level, which discusses the multiplicity of equilibria associated with the conduct of monetary policy, and in the role of the central bank as a lender of last resort.

With the onset of the financial crisis, the Great Recession and the European debt crisis, this consensus was broken. Fiscal stimulus came back as a prescription to stabilize the economy. The development of government plans to increase aggregate demand marks a change both for the US and for Europe. Subsequently, governments in a large set of countries found themselves with very high levels of public debt. This was in part due to the aforementioned stimulus and lower tax revenues due to the economic downturn, and in part due to automatic stabilizers which reduced revenues and increased expenditures. Following the economic turmoil and the associated political stress we have seen a strong focus on fiscal consolidation through discretionary actions, but to my knowledge there has been little new analytical work on fiscal stabilization policy. On the question of designing better stabilizers, a recent answer was given by McKay and Reis (2016), who show that most of the measured welfare benefits from automatic stabilizers come from the provision of insurance (through changes in precautionary savings) and from redistribution. These authors also show that high transfers to the unemployed and poor can be quite effective at lowering volatility. However the effects on welfare, consumption, and output depend on the specific design of those automatic stabilizers, that is, on the way they avoid reducing the incentives to work or save and invest.

At the same time, the stance of monetary policy became exceptionally expansionary by historical accounts: the policy rate was lowered toward its effective zero lower bound and central banks began large scale purchases of private and public assets, with longer maturities than in normal times.

Consequently, central bank balance sheets grew to unprecedented levels. At the same time that these policies were used as stimulus to the economy, they served specifically to provide liquidity to the financial sector and to repair specific financial markets. In this way, monetary policy had a strong effect on the lowering of rates and risk premia.

Part of this new unconventional monetary policy works through credit easing through the so-called credit channel which affects credit allocation and relative yields. Through this channel monetary policy can have aggregate effects, but it also has strong redistributive effects. As quantitative (or credit) easing policies benefit the holders of financial assets, e.g. boosting prices of bonds and real estate, it is more difficult now to trace the dividing line between monetary and fiscal policy. Monetary policy transmission now looks like fiscal policy.

The abstraction typically used in research of not differentiating between the budget constraint of the government from that of the central bank began to be challenged these days. Modeling a separate balance sheet for the central bank and a constraint for the government obliges one to make explicit the restrictions of having an independent central bank as well as the vulnerabilities created by having the private sector holding an increasing amount of assets (reserves). Is it sustainable for the private sector to hold an increasing amount of assets when these are not associated with expected future taxes? If the associated risks materialize, should the treasury be ready to receive fewer remittances or to recapitalize central banks? On the question of why the central bank balance sheet matters, the recent work by Del Negro and Sims (2015) gives us good arguments to discuss the consequences of the lack of fiscal support for the central bank. For example, the commitment from the government to never recapitalize the central bank can impose a restriction on the ability of the central bank to satisfy its mandate to control inflation. I will not further comment on issues related to central bank balance sheets and risky assets, and instead refer interested readers to Benigno (2016).

So in practice we have arrived in a world where the new unconventional monetary policy has an increasing connection with the traditional fiscal policy, and policy makers are scrambling for additional instruments that could complement monetary policy and/or could be accommodated by governments with little fiscal room.

From research on optimal fiscal policy to optimal monetary policy

What happened on the research front that could help answer these questions? While we had a strong line of quantitative general equilibrium models which were largely used for fiscal policy over the long term and had helped to make progress on the analysis of fiscal (tax) policy, its extension to business cycle frequency and its interaction with monetary policy is very recent. Let me begin by describing the advances in monetary policy and return later to talk about its extension to unconventional fiscal policies and its potential value for the current situation. By extending general equilibrium models to stochastic and monetary environments, we were able to explain the gains of using monetary policy for stabilization purposes: agents may be restricted in the setting of prices, wages, or in the choice of portfolio composition. The severity of these restrictions determines the strength of the transmission mechanism of monetary policy. Even though monetary policy can have positive effects, it is not possible to use this policy systematically to take advantage of these effects. A new impetus for using monetary policy as a stabilizing mechanism occurred when research showed that policy can be used in response to shocks so that the negative welfare effects of the nominal rigidities, together with the other distortions in the economy, are minimized. This new strand of literature was able to address quite relevant questions such as: How should monetary policy be conducted in response to shocks in the economy? How relevant is the transmission mechanism of monetary policy for the conduct of this optimal policy? How costly can a single monetary policy be when countries don't share a single monetary transmission mechanism and are exposed to asymmetric shocks? Or in short, how should central banks conduct short run monetary policy?

Ireland (1996) is the first to extend the Ramsey concept of minimizing distortions in a general equilibrium model to a monetary model with nominal rigidities. The idea is to define the set of feasible allocations given the existing policy instruments and then to determine what characterizes the best solution, namely how policy should react to fundamental shocks and how prices and allocations react to the fundamental shocks and to the described policy changes. This new approach allows us to explain how the the so called Gaps can really be read as Triangles, or wedges, which policy should smooth across time and different states of the world (see Adão et al. (2003)).

This can be related with what, already in the end of the 80's, we read in Long and Summers (1988), that "demand management policies can and do affect not just the variance, but also the mean of output" and "... successful macroeconomic policies fill in troughs without shaving off peaks." That is, the role of policy is not to close gaps but to minimize wedges, implying that the criteria for stabilization policy should be identical to that of any other policy: a welfare criteria. When this framework, developed mainly for fiscal policy, is applied to monetary policy it has the advantage of making very clear the comparability between monetary and fiscal stabilization policy channels.

Even though the first series of papers had a strong focus on conventional monetary policy, with fiscal policy being reduced to lump sum taxes/transfers, that comparison was clear. The substitution of gaps by triangles showed that while the transmission of monetary policy shocks is extremely dependent of the type and degree of the frictions existing in the economy, the same is not true of the optimal reaction to a given shock: the design of optimal rules to various shocks have been shown to be much more robust.

One very instructive result from the early stage of this literature is that when price adjustment is slow, for example due to sticky prices, the planner is able to side step the zero bound restriction on nominal interest rates and achieve higher utility. This ability is driven by the reaction of policy to a particular fundamental shock that allows ex-post mark-ups to be state-contingent, contrary to what happens when prices are flexible, for the class of state-of-the-art monetary policy models with monopolistic competition with constant elasticity of substitution across goods. Therefore, we can write theoretical examples where the existence of nominal rigidities can improve the outcome of policy relatively to those with flexible prices. This is clearly a result similar to the well-known one in the second-best literature on fiscal policy: namely that in the face of several distortions the elimination of one of them is not necessarily welfare improving.

The optimal mix of policies

The next step in the research literature was to study the interaction of monetary and fiscal policy. The optimal policy was then a joint decision on the choice of both type of instruments. The way this interaction was developed was to limit fiscal instruments to proportional tax rates that can be state-contingent. In most of these papers government expenditures are exogenous and therefore cannot be used as a policy instrument. This methodological choice is very much driven by the difficulty of evaluating the welfare effects of a broad measure of public consumption.

These papers allow us to argue that, independently of the degree or type of price stickiness, it should be possible to implement the same relevant set of allocations (see Adão et al. (2004) and Correia et al. (2008)), and that each allocation in that set is implemented with policies that are also independent of the price stickiness. The intuition behind this result is that policy shocks have differing effects in the model economy depending on the type and degree of price rigidity, but the same is true for the exogenous shocks, e.g. technology or government expenditures. This leads to the important result that when a policy satisfies a minimum requirement of optimality, the combined effect of the exogenous shocks and the response of policy is invariant to the degree or type of price. In other words, the influence of price rigidities can be undone if policy makers can decide monetary and fiscal policy jointly. We can summarize these results by saying that transmission is very relevant when policy is discretionary or when it is very far from efficient. But in other environments, for example with different price setting restrictions, transmission can be observationally equivalent.

The necessary condition for this equivalence result of different environments is the existence of a sufficiently rich set of policy instruments. In particular, we show that within the confines of a standard business cycle model, state-contingent debt is a redundant policy instrument as long as policy makers can use both consumption and labor income taxes freely. The main policy lesson from our analysis is that when state-contingent fiscal and monetary policy are jointly decided, price stability is a requirement of efficiency, independent of preferences, as long as preferences are on the final goods from which the households extract utility. This is a normative statement, stronger than the Ramsey prescriptions. It also appears to be consistent with a generalized mandate and practice by central banks. In addition, this result tell us that it is not possible to distinguish whether the Great Moderation was due to lower volatility of outcomes from different transmissions of shocks or to better policy.

A related result is that the more you need to use monetary and fiscal policy instruments, the more effective they become. Therefore the question of the magnitude of the fiscal multiplier that has produced so many works in the post crisis period should be assessed carefully. What we have learned is that just as very different channels can be associated to different magnitudes of the multiplier, the same channels would lead to very different effects of the shock to which policy is reacting. When we join these two pieces, the total effect of the shock behind the recession and the policy response, the outcomes should be much more similar than those described in most of the literature.

We can now apply the lessons learned from this literature on the links between central banking and fiscal policy. To do that, let me present some results from what we can call the unconventional fiscal policies toolkit. I show how I believe we should complement monetary and fiscal policies in crisis times, when monetary policy exhausted its conventional instruments and fiscal space has no room for conventional stimulus of the economy.

In this way we can discuss really important, particularly topical questions. The first is the answer to the question of "How can we overcome the costs of the ZLB?" and the second is "How can we compare credit subsidies to credit easing?" It is well known that non-arbitrage between money and bonds restricts nominal interest rates from becoming (too) negative. How negative is currently an open question, but no one doubts that there is some lower bound. It is clear from recent experience that the Great Recession is one event in which it would be desirable for the central bank to lower the policy rate below that bound. Instead, alternative policies were put in place, namely the use of unconventional monetary policies, including forward guidance and the fiscal stimulus that lead to the public debt legacy that we face these days in a large number of countries. I want to stress that not only were there more options left unexplored but, more relevant for this note, these alternatives were precisely in the set of unconventional fiscal policies which include the interaction between fiscal and monetary policy.

The time for unconventional fiscal policy?

The cost of the zero bound is a major concern which leads to the suggestion of better integration between monetary and conventional fiscal policy (see e.g. Blanchard, Dell'Ariccia and Mauro (2010)). However, Correia et al. (2013) proposes the use of unconventional fiscal and monetary policy when the zero lower bound is reached. If the nominal interest rate is zero, proportional contingent taxes can substitute for the role the nominal interest rate would normally play. Whatever monetary policy can achieve with the nominal interest rate, fiscal policy can also be done with a combination of consumption, labor and capital income taxes. The intuition behind why this unconventional fiscal policy can neutralize the cost of the zero bound constraint is simple. The prices that matter for inter-temporal decisions are consumer prices, which are gross of consumption taxes. Therefore, the idea is to induce inflation in consumer prices, keeping producer price inflation at zero, to eliminate the costs associated with nominal frictions. The result is that we can reach negative real interest rates while avoiding the distortions associated with producer price inflation. A temporarily lower consumption tax relative to the future one generates expected inflation in consumer prices. To avoid changes in incentives, distinct from those usually associated with a lower interest rate, the change (increase) in the level of consumption taxes (or the equivalent VAT taxes) must be counteracted by a decline in the labor income tax. By the same reason a change (a decline) in the tax of capital income neutralizes the introduction of the increasing tax on consumption goods. In this way the use of those three taxes can replicate the decline of the tax on money, that is the nominal policy interest rate.

This policy recommendation requires flexibility of tax policy. It should be noted that this type of flexibility has been prescribed by several authors. Moreover, and perhaps even more relevant, some changes adopting these insights were introduced (partially) during this crisis. For example Feldstein (2002) says that "The Japanese government could announce that it will raise the current 5 percent value added tax by 1 percent per quarter and simultaneously reduce the income tax rates to keep revenue unchanged, continuing this for several years until the VAT reaches 20 percent." And in his presidential address to the 2011 American Economic Association Annual Meeting, Bob Hall (2011) reiterated Feldstein's ideas and encouraged further research to understand the viability and effects of unconventional fiscal policy, both theoretically and empirically. On the introduction of this instrument in reaction to the state of the economy we can point the Japanese experience: Japan announced in October 2013 an increase of the consumption tax in two phases (April 2014 and October 2015). Economic activity in Japan grew strongly in 2014Q1, particularly consumption, but contracted afterwards. The second plan was postponed to April 2017, not implemented then, and more recently announced for October 2019.

Therefore, I believe that the argument that fiscal instruments are not as flexible as monetary policy instruments should be revisited. While perhaps this conclusion can apply to stabilization policy during normal times, exceptional circumstances such as the recent crisis or the Japanese stagnation since the nineties can change this evaluation.

Another exercise is to compare the use of unconventional monetary versus fiscal policy, namely credit subsidies (Correia et al. (2016)). When considering the recent crises, financial and sovereign, the limitations of models without an explicit role for the financial sector and without financial frictions to evaluate both the causes of the recession as well as the policies to the recovery were self-evident. In the models used before the crisis, under the assumption of the absence of nominal rigidities, the zero bound on nominal interest rates is not a constraint to policy. In fact, it is the optimal policy. However those models were too simple to be able to take into account the financial channels which a large body of literature agrees were relevant during the recent crisis. One way to model the interest rate spreads as a simple extension to the existing models is to subject financial intermediaries to an enforcement problem, in the spirit of Gertler and Karadi (2011). Firms must borrow to pay wages, those loans must be intermediated, and banks can do that at low cost. This imposes not a resource cost, which for simplicity we assume to be zero, but rather an efficiency cost resulting from the incentive problem that bankers can divert part of the bank's assets. As banks must earn rents, they charge a differential between the deposit and the lending rates; this spread generates profits which are accumulated as internal funds. These lending spreads can be particularly high when banks' internal funds are low as a result of unfavorable exogenous shocks, which can be interpreted as shocks to the value of collateral. There is a sense in which lending rates may be too high in these economies: if they are too high or too volatile then policy can be used to lower or smooth them, increasing welfare. Although interest rate policy does not act directly on the spreads, monetary policy can be used to partially correct those distortions. The spreads are whatever they need to be to align the incentives of banks. Interest rate policy reduces the financing costs of banks, reducing the spreads and the lending cost of firms.

A very low policy rate, possibly zero, will minimize the lending rates and minimize the distortion that it causes on allocations. Nevertheless, because of the Zero Lower Bound, lending rates may still be too high and too volatile. If the policy rate could be negative and if it could be financed with lump sum taxes, then it would be possible to achieve the first best in these economies. A result analogous to the Friedman rule would be obtained, but this rule would be on the lending rate and not on the policy rate. When we introduce unconventional fiscal policy, in this case a credit subsidy, we can act directly on the existing distortions. Credit subsidies play the same role as the policy interest rate, even if acting through very different mechanisms. And, furthermore, they have the advantage that they are not subject to any restriction such as the zero bound constraint. With credit subsidies it is therefore possible to implement allocations that would be previously infeasible for monetary policy, because they would require negative interest rates. The policy rate could be set at some arbitrary level, possibly close to the zero bound. Banks would charge time varying spreads and lending rates. But the rates paid by borrowers net of credit subsidies could be smooth and very low. The paper also shows that the budget implications of the policy rate and tax subsidies are exactly the same if we take into account a consolidated budget constraint between the government and the central bank. This environment allows the comparison of this unconventional fiscal policy with the unconventional monetary policy in place after the crisis, namely the credit easing policies. It assumes that there is an alternative technology, which the central bank can use, in which the enforcement problem is solved by paying a resource cost, which allows the central bank to give credit directly to firms. The comparison of unconventional fiscal and monetary policies comes down to comparing a resource cost versus a deadweight loss. It can be shown that credit easing does not appear to be a good alternative to the already described unconventional fiscal policy.

Concluding remarks

We find ourselves in this post crisis period with a legacy that in addition to quite special economic and financial conditions is also characterized by a legacy coming from new policy tools and new experiments. The monetary toolkit was clearly reinforced and new, unconventional, monetary policies were implemented and are still in place in most developed economies. What I wanted to discuss in this short note was that the new world that monetary policy makers entered was not accompanied by a similar move in the fiscal sphere. And that it is difficult to say whether the return to the old normal of not very low policy interest rates will be there in the near future. The theoretical developments of the last two decades would point to more ambition and originality in the use of fiscal instruments such as the ones described here. This would not just give additional room of manoeuvre to tackle the ongoing prolonged recovery but, maybe more importantly, may allow us do so in a more efficient way compared with current policy actions.

It is true that this would require a stronger coordination across institutions compared to the pre-crisis period. But it is also the case that the continuation of the unconventional monetary policy has mechanisms very similar to those of fiscal policy and stronger re-distributional effects, which would imply such coordination may prove necessary in any case.

In this scenario, keeping the research agenda updated as well as a strong dialogue between policy and research is more important than ever.

References

Adão, B., I. Correia and P. Teles (2004). "The Monetary Transmission Mechanism: Is it relevant for policy?" JEEA, 2 (2-3), 310-319.

Adão, B., I. Correia and P. Teles (2003)."Gaps and Triangles." Review of Economic Studies, 70 (4), 699-714.

Benigno, P. (2017). ”A Central Bank Theory of Price Level Determination.” CEPR Discussion Paper No. 11966.

Blanchard, O., G. Dell'Ariccia and P. Mauro (2010). "Rethinking Macroeconomic Policy." Journal of Money, Credit and Banking, 42, issue s1, 199-215.

Correia, I. F. De Fiore, P. Teles and O. Tristani (2016). "Credit Subsidies." Working Paper BCE, 1877.

Correia, I., E. Farhi, J. P. Nicolini and P. Teles (2013). "Unconventional Fiscal Policy at the Zero Bound." The American Economic Review, 103, 1172-1211.

Correia, I., J.P. Nicolini and P. Teles (2008). "Optimal Fiscal and Monetary Policy: Equivalence Results." Journal of Political Economy, 116-1, 141-170.

Del Negro, M. and C. A. Sims (2015). "When does a central bank's balance sheet require fiscal support?" Journal of Monetary Economics, 73(C), 1-19.

Feldstein, M. (2002). "Commentary: Is There a Role for Discretionary Fiscal Policy?" Rethinking Stabilization Policy. Kansas City, Missouri: Federal Reserve Bank, 151-62.

Friedman, M. (1948). "A Monetary and Fiscal Framework for Economic Stability." The American Economic Review, 38- 3, 245-264.

Gertler, M. and P. Karadi (2011). "A model of Unconventional Monetary Policy," Journal of Monetary Economics, 58(1), 17-34, January.

Hall, B. (2011). "The Long Slump" The American Economic Review, 101, 431–469.

Ireland, P. (1996). "The Role of Countercyclical Monetary Policy," Journal of Political Economy, 104(4), 704-723.

Long, J. B. de and L. H. Summers (1988). "How Does Macroeconomic Policy Affect Output?" Brookings Papers on Economic Activity, 2, 433-494.

McKay, A. and R. Reis (2016). "The Role of Automatic Stabilizers in the U.S. Business Cycle." Econometrica, 84, 141-194, January.

Solow, R. (2005). "Rethinking Fiscal Policy", Oxford Review of Economic Policy, 21, 4, 1, December.

Interest rate spreads hikes: What lies behind them?

Paulo Júlio
Banco de Portugal and CEFAGE

José Maria
Banco de Portugal

October 2018

Abstract
The 2007–2009 financial turmoil and the euro area sovereign debt crisis that followed were characterized by severe interest rate spread hikes. In recent work we developed a novel general equilibrium model for a small-open euro-area economy, endowed with a rich characterization of the banking system that allows for regulatory capital requirements, defaulted loans and occasionally binding endogenous credit restrictions. In this article, we use our model to offer a model-based explanation of the endogenous mechanisms associated with sharp interest rate increases based on macroeconomic fundamentals. After briefly describing the model, we analyze the concomitant interest rate dynamics and decompose the overall spread into three components: a capital requirements-driven spread, a credit restrictions-driven spread, and a retail-driven spread. Results suggest that defaulted loans and occasionally biding credit restrictions—two of our novel mechanisms—contribute to severely amplify spread hikes under financial disturbances, but play lesser roles under non-financial ones. (JEL: E12, E32, E44)

Introduction

The 2007–2009 financial turmoil and the euro area sovereign debt crisis that followed were characterized, among other factors, by severe interest rate spread hikes. The explanations for the sharp increase in the price of credit are manifold and include not only a deterioration in the macroeconomic fundamentals, but also reactions based on other factors such as market fears and panic responses. Interbank market freezes, fears of default in the private and public sectors, spillover effects and rises in the nationwide risk premium are amongst the commonly discussed topics.

In Portugal, both the spread on bank loans and debt securities of non-financial corporations faced a double-stage increase, first on the verge of the financial turmoil and thereafter over the sovereign debt crisis period (see Figure 1). During this period, debt security interest rates reached similar levels

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the views of Banco de Portugal or the Eurosystem. Any errors and mistakes are ours. Paulo Júlio acknowledges financial support from *Fundação para a Ciência e Tecnologia*, (grant UID/ECO/04007/2013), and FEDER/COMPETE (POCI-01-0145-FEDER-007659). We are grateful to Rodrigo Barrela for excellent research assistance.

E-mail: pfjulio@bportugal.pt; jrmaria@bportugal.pt

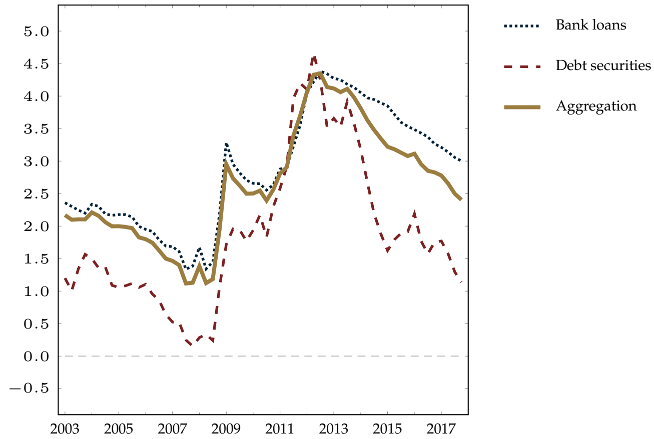


FIGURE 1: Interest rate spreads of Non Financial Corporations.

Notes: Interest rate spreads, calculated as the difference *vis-à-vis* the 3-month EURIBOR, use bank loans (dotted line), debt securities (dashed) and an aggregation of both. The cost of financing with bank loans and debt securities are measured, respectively, by the interest rates on new loans granted by resident banks, and interest rates on commercial paper and long-term bonds issued by Portuguese corporations.

as those of new loans granted by resident banks—in sharp contrast with the first part of the sample—which suggest different operating conditions in these market segments. More recently we notice a downward trend of interest rates *vis-à-vis* the 3-month EURIBOR, but particularly when based on commercial paper and long-term bonds.

A key criticism often addressed to general equilibrium models is their inability to identify the cumulative vulnerabilities before the worst recession of the postwar period (Christiano *et al.* 2018), let alone to signal meaningful policy warnings. The inability of financial frictions-based models to properly take into account rare or extreme events and to provide a convincing improvement over simpler and more standard alternatives—including interest rate spread dynamics such as those depicted in Figure 1—suggests that some work must be targeted to this area.

In a recent paper, Júlio and Maria (2018) develop a novel Dynamic Stochastic General Equilibrium (DSGE) model for a small-open euro-area economy, endowed with a rich characterization of the banking system. Such work has obvious policy implications. First, the model improves the explanatory power of the mechanisms behind interest rate spreads, particularly on the supply side. Second, the distribution of shocks matter to explain output fluctuations and in particular output downfalls; the mean is not sufficient. Third, a narrow set of negative small-sized financial-based shocks can trigger a deep and protracted recession, which may contribute decisively to enhance the predictive density of DSGE model in crisis periods. However, the opposite is not true: positive financial shocks may not trigger a sizable expansion. Fourth, our model predicts that defaulted loans mostly accumulate in banks balance sheet on the aftermath of financial shocks, which is in line with facts recorded in a number of euro area economies in the aftermath of the financial crisis. Fifth, the model provides a completely novel framework to analyze policy-oriented measures aimed at increasing the robustness of the financial and banking system, especially during crisis periods.

Credit supply decisions in Júlio and Maria (2018) are simultaneously driven by regulatory capital requirements, defaulted loans, and credit restrictions that become binding under shocks severely depleting banks' value. The banking system proposed therein intertwines two strands in the literature with two novel features. Capital requirements follow the approach in Benes and Kumhof (2015), and are coupled to a moral hazard-inspired credit constraint mechanism in the spirit of Gertler and Karadi (2011), Gertler *et al.* (2012), and Gertler and Karadi (2013). However, contrary to these studies, which assume an always binding incentive compatibility constraint, we propose and develop an occasionally binding mechanism which is slack in the steady state but endogenously affects credit supply decisions when banks' capital is severely affected. Simultaneously, we bring forth into the model a theory of optimal impairment loss recognition, which gives rise to an endogenous defaulted loans stock that bankers manage over time.

In this article we offer a model-based explanation of the mechanisms behind interest rate spread hikes triggered by macroeconomic fundamentals, *i.e.* intrinsic and endogenous transmission mechanisms. We leave aside non-fundamental issues (such as "sunspot events"). We calibrate the model for the specific case of Portugal, and decompose the overall interest rate spread into the sum of three components: the capital requirements-driven spread, the credit restrictions-driven spread, and the retail-driven spread. The capital requirements-driven spread is the interest margin required to cover the expected costs of a possible violation of capital requirements. We assume that the bank is allowed to remain in activity at all times, but has to pay some cost to place the capital back on track. This can be for instance a cost for restructuring the bank or some fraction of assets, or even a reputation cost. The credit restrictions-driven spread is defined as the interest margin induced by credit supply restrictions, which emerge when banks' value declines significantly. Specifically, during a financial meltdown the value of banks' capital collapses and bankers' must tighten credit conditions in order to limit the leverage position, triggering important spread hikes. Finally, the retail-driven spread is the margin required by retail banks to cover the expected losses generated by corporate bankruptcy. We define the wholesale interest rate spread in this framework as the sum of the capital requirements-driven spread and the credit-restrictions driven spread.

To disentangle the role played by defaulted loans and credit restrictions on the spread decomposition, we simulate versions of the model in which these mechanisms are deactivated. For illustrative purposes, we carry out this decomposition for three shocks with negative macroeconomic impacts—a supply-side shock, a domestic demand shock and a financial shock. Specifically, we simulate a contraction in technology, a decline in government consumption, and an increase in the riskiness of investment projects. All shocks have a temporary nature, and therefore the equilibrium reverts to the steady-state values in the long run.

For the financial shock—an risk increase—greater corporate default rates impose losses for the banking sector, to which banks respond by increasing the wholesale spread, specifically the capital requirements component. This generates larger margins, required to cope with greater expected losses arising from the larger likelihood of not complying with capital requirements. The wholesale spread hike is severely amplified by defaulted loans and credit restrictions. The sizable increase in the amount of due loans that follows the sharp increase in corporate default imposes extra losses in the banking system, severely contributing to the wholesale spread hike *via* the capital requirements component. The collapse of banks' value on the aftermath of credit losses leads bankers to impose tighter lending conditions, unfolding an additional contribution to the wholesale spread hike *via* the credit restrictions component.

Whereas the demand-side shock imposes a spread hike, the supply-side shock fosters a decline due to its inflationary effects, which reduce the real cost of credit and benefit corporate leverage. In both cases, credit restrictions remain slack at all times and the corresponding spread component is nil, since banks' value is barely affected. Changes in the retail spread reflect shifts in the corporate default rate, whereas changes in the wholesale spread—corresponding in this case to the capital requirements component—echo the banks' leverage position and the probability of not complying with regulatory capital requirements. Defaulted loans amplify variations in the wholesale spread, since they affect banks' costs and the risk of violating regulatory levels.

Defaulted loans and credit restrictions are strongly intertwined. On the one hand, a large increase in the former imposes extra losses in the banking system and depletes banks' value, leveraging the effects of credit restrictions and thus magnifying wholesale spreads hikes. On the other, more expensive credit that follows tighter lending conditions negatively impacts corporate balance sheets by rising interest expenses. This boosts the wholesale spread hike as banks cope with the increased amount of due loans that follows the higher corporate bankruptcy chance.

Results are quantitatively dependent on the model calibration, which is naturally subject to some uncertainty. Robustness checks suggest however that the above conclusions are qualitatively valid under plausible alternative values.

A DSGE model for a small-open euro area economy

This section briefly overviews the model presented in Júlio and Maria (2018). The domestic economy is composed of nine types of agents: households, intermediate goods producers (manufacturers), final goods producers (distributors), retailers, capital goods producers, entrepreneurs, banks, the government, and foreign agents.

Households are composed of three member types: workers, entrepreneurs and bankers. There is full consumption insurance within the family. When exiting activity, the last two member types transfer accumulated earnings back to the household. In each period and for each activity the number of entries and exits is the same. Households follow an infinitely-lived structure, renting labor services to manufacturers, paying lump-sum taxes to the government, and earning interest on money holdings. They also receive income for services provided in the repossession of assets from bankrupt corporations and in reducing the defaulted loans portfolio of banks—activities performed at no personal effort—and receive dividends from firms, in addition to the accumulated earnings from exiting entrepreneurs and bankers. Households are not allowed to hold foreign financial assets. A representative household derives utility from consumption and real money holdings, measured by the real value of deposits, and disutility from working. Money holdings do not affect the intratemporal consumption-leisure choice, and thus money is superneutral in the steady state. Households are also wage setters, setting the wage according to their disutility from working and taking into account their market power and the demand for labor. They face Calvo-type frictions when setting wages and are therefore unable to reoptimize in every period.

Capital goods producers combine the undepreciated installed productive capital stock bought from entrepreneurs with investment goods bought from retailers, to produce new installed productive capital. They face quadratic adjustment costs to reproduce the sluggishness in investment. Capital is sold to entrepreneurs, who will own it during the production cycle. Manufacturers combine capital—rented from entrepreneurs—with labor services—hired from households—to produce intermediate goods. They face Calvo-staggered price adjustment and quadratic costs when deciding to adjust the number of hired hours. Distributors combine intermediate goods—bought from manufacturers— with imported goods—bought from abroad—to produce the final good. They face Calvo-type price staggering and quadratic adjustment costs on import content changes. Finally, perfectly competitive retailers acquire the final good from distributors and reallocate it to different costumers—households, capital goods producers, government, and foreign distributors.

The government keeps the budget balanced at all times, financing public consumption with lump-sum taxes, levied on households. The foreign economy corresponds to the rest of the euro area. The domestic economy interacts with the foreign economy *via* goods and financial markets. In the goods market, domestic distributors buy imported goods from abroad to be used in the production stage. Likewise for foreign distributors, who buy export goods from domestic retailers for the same purpose. In the financial market, banks can finance balance sheet operations by trading assets with the foreign economy. Monetary policy is exogenous and unresponsive to domestic developments, a consequence of the small-open economy framework. Hence, developments in euro-area interest rates are orthogonal to domestic developments, as in Adolfson *et al.* (2007). The nominal exchange rate *vis-à-vis* the rest of the euro area is irrevocably set to unity.

The financial sector: entrepreneurs and banks

The financial transmission mechanism linking entrepreneurs to banks is modeled along the lines in Bernanke *et al.* (1999), Christiano *et al.* (2010), and Kumhof *et al.* (2010). Financial frictions affect the return on capital and therefore capital demand. Before each production cycle, capital goods producers buy the undepreciated productive capital stock from entrepreneurs, combining it with investment goods bought from retailers to produce new installed productive capital. This capital is then sold to entrepreneurs, which will own it during the next production cycle. Entrepreneurs do not have access to sufficient internal resources to finance desired capital purchases, but can borrow the difference from retail banks at a cost. They face an idiosyncratic shock that changes the value of the firm after decisions have been made. When hit by a severe shock, the value of assets collapses and the entrepreneur must declare bankruptcy, handing over the value of the firm to the bank.

The banking system builds on Benes and Kumhof (2015), and is composed of retail branches and wholesale banks. Retail branches operate in a perfectly competitive environment, celebrating loan contracts with entrepreneurs. These contracts set an unconditional, non-state contingent lending rate. Since entrepreneurs are risky, so are the individual loans of retail banks, who therefore charge a spread over the wholesale lending rate—the cost of obtaining funds from the wholesale bank or households—to cover for the losses incurred in the mass of entrepreneurs that declare bankruptcy. We hereinafter term this margin the retail-driven spread. Since a given retail branch lends to many entrepreneurs, by the law of large numbers the aggregate loan portfolio is risk-free, and hence *ex-ante* profits are zero. Retail branches are however exposed to non-diversifiable aggregate risk given the non-state contingent lending rate, and thus *ex-post* profits—to be transferred to wholesale banks—may differ from zero. When a corporate firm goes bankrupt, retail branches must pay a service to households—*i.e.* a repossession cost—in order to take possession of corporate assets.

Wholesale banks finance their activities, i.e. loans to retail branches, through equity, deposits, and foreign funds. We assume that repossessed assets are illiquid and accumulate as defaulted loans on the balance sheet. Over time, an exogenous fraction of defaulted loans is automatically transformed from illiquid into liquid at no cost, but banks can increase the pace of this transformation by requesting a liquidation service—henceforth interpreted as impairment losses—from households. Wholesale banks face an idiosyncratic shock affecting the return on their loan portfolio which, coupled with potential losses from retail branches, may trigger balance sheet effects and/or credit supply restrictions. They are subject to regulatory capital requirements and non-compliance results in adjustment costs and reputational losses. Banks therefore endogenously set capital buffers, which allow them to cushion adverse shocks that negatively affect the value of capital. For simplicity, we rule out bank failure.

Credit supply restrictions arise endogenously from a modified moral hazard/costly enforcement problem inspired in Gertler and Karadi (2011), Gertler *et al.* (2012), and Gertler and Karadi (2013). The banker has the option to divert a fraction of funds, though this only becomes attractive when the bank's value collapses well below the steady-state level, *i.e.* when the bank is and is expected to remain excessively leveraged. Households recognize this fact and restrain the amount of deposits placed at the bank, pushing down leverage and aligning their interests with the banker's incentives. In this way, wholesale banks become supply constrained with respect to the resources they can make available to the entrepreneurial sector.

The occasionally binding nature of credit restrictions is able to generate powerful asymmetric responses to financial or banking shocks—those whose nature is endowed with important effects on the banking system. Specifically, under "good shocks," *i.e.* those increasing banks' value, credit restrictions remain slack and play no role whatsoever. Under "bad shocks" depleting banks' value, they may become binding for some period of time and greatly affect the model dynamics, and particularly spreads.

The wholesale interest rate spread is the bank's margin, simply defined as the interest received by lending to retail branches minus the cost of raising funds, *i.e.* the interest rate paid to depositors. In equilibrium, this margin is driven by the possible violation of capital requirements and by credit supply restrictions. While the former—termed capital requirements-driven spread—implies a pecuniary or reputational cost, the latter—named credit restrictions-driven spread—triggers a *ceteris paribus* decline in banks' income. Defaulted loans impose extra losses in the banking sector, increasing the probability of non-compliance with regulatory requirements and depleting banks' value. They therefore affect the wholesale interest rate spread *via* both components, and interact strongly with credit restrictions. Figure 2 illustrates the relationship between interest rate spreads and agents in our small-open euro-area model.

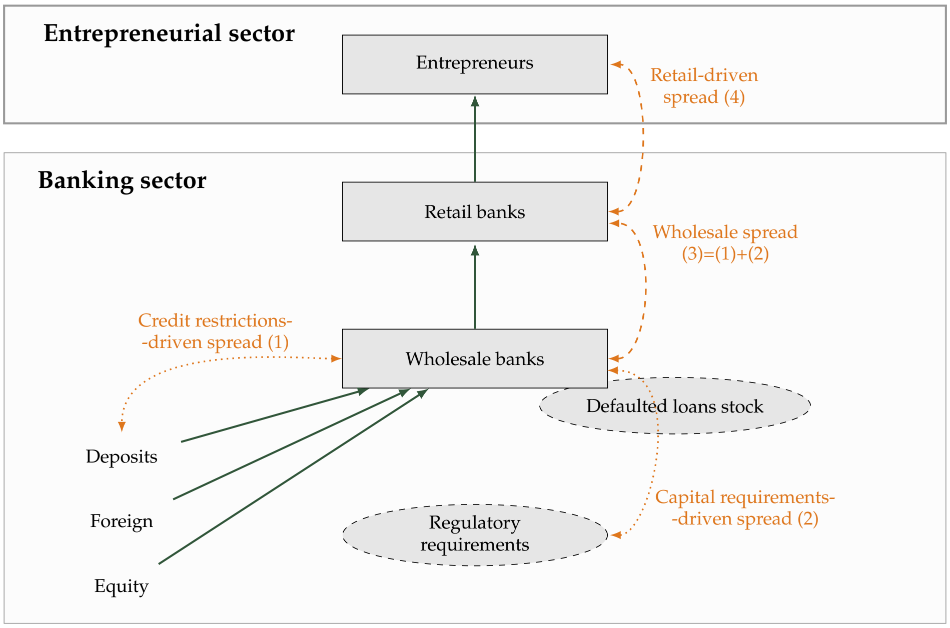


FIGURE 2: Interest rate spreads and agents.

Notes: The defaulted loans stock is managed by wholesale banks. Before the end of each period, retail banks are assumed to transfer all their defaulted loans to wholesale banks.

Calibration

The model is calibrated to match long-run data or studies for Portugal and euro area economies. Some parameters are exogenously set by taking into consideration common options in the literature, available historical data, or empirical evidence, whilst others are endogenously determined to match great ratios or other measures. We now describe briefly the main calibration features. Tables 1 and 2 present a selection of the model's calibrated parameters, whereas Table 3 exhibits implied key steady-state relationships.[1]

The interest rate target is set at 3.2 percent per year, matching the pre-crisis average for the 3-month Euribor. Steady-state inflation is set at 2 percent per year, in line with the ECB's price stability target. The inverse Frisch elasticity is set to 0.276. The discount factor is 0.996, resulting in a net foreign asset position of around -40 percent of GDP for a target ratio of -30 percent. The

1. A more complete and exhaustive description of the calibration is found in Júlio and Maria (2018). The calibration on defaulted loans has been revised herein and differs from the one depicted in the paper.

	Value
Households	
Inverse Frisch elasticity	0.276
Discount factor	0.996
Wage and price markups	
Wage markup	0.32
Intermediate goods price markup	0.21
Final goods price markup	0.09
EoS and technology	
EoS, intermediate goods	0.99
EoS, final goods	1.50
EoS, exports	1.50
Quasi-labor income share	0.60
Home bias in domestic distributors	0.67
Export market share	0.03
Calvo parameters	
Wage	0.75
Intermediate goods	0.75
Final goods	0.50
Miscellaneous	
Depreciation rate (annualized)	0.1
Interest rate target (annualized)	0.032
Inflation target (annualized)	0.02
Target NFA-to-GDP ratio	-0.30

TABLE 1. Selected main parameters (non-financial).

Sources: *Banco de Portugal* data, National accounts data, several studies on the Portuguese and euro area economies, and authors' own calculations.
Notes: EoS—Elasticity of Substitution; NFA—Net Foreign Assets.

deposits-to-GDP ratio is roughly 40 percent. Steady-state price markups are set at 6/19 for wage setting, 4/19 for the intermediate goods sector, and 1/11 for the final goods sector. The elasticity of substitution between capital and labor is nearly 1, whereas for domestic and foreign goods distributors the elasticity of substitution between inputs is 1.5. The depreciation rate of capital is calibrated at 10 percent per year. The labor quasi-share and the home bias parameters are endogenously calibrated to take into account the actual labor income share and the import share, whereas the export market share is adjusted according to the exports-to-GDP ratio. The investment and labor adjustment costs are parameterized to ensure plausible dynamics. Likewise for the parameter assessing the cost of under- or over-utilization of capital. The import content adjustment costs ensures plausible real exchange

	Value
Entrepreneurs	
Repossession costs (% of firm value)	0.40
Average lifetime (years)	6.25
Banks	
Average lifetime (years)	5
Capital ratio requirement	0.14
Defaulted loans	
Recovery fraction	0.04
Credit restrictions	
Fraction of corporate loans that can be diverted	0.16

TABLE 2. Selected main parameters (financial).

Sources: *Banco de Portugal* data, National accounts data, several studies on the Portuguese and euro area economies, and authors' own calculations.

rate fluctuations. Calvo parameters imply an average contract duration and intermediate goods average price duration of 1 year, and a final goods average price duration of half a year. We assume no indexing.

On the entrepreneurial side, we calibrate parameters to match a target leverage (net worth-to-debt ratio) of 1.2, a yearly default probability of 3.6 percent, and a yearly retail lending rate spread of 1.6 percentage points. An entrepreneur stays on the job on average around 6 years. For the banking sector, we set the capital requirement to 14 percent and let banks build an endogenous capital buffer of 3 percentage points, yielding a steady-state capital-to-loans ratio of 17 percent. The probability of non-complying with capital requirements is set at 4 percent, and the spread between the wholesale interest rate—matched by the 6-month Euribor—and the deposits rate is 0.5 percentage points. A banker stays in the job on average around 5 years. We consider that 4 percent of total loans are recovered in each quarter, and adjust parameters to obtain a defaulted loans-to-credit ratio of approximately 6.8 percent. New defaulted loans in each period amount to 0.56 percent of total credit, which in the steady state approximately matches the amount that is withdrawn from the balance sheet—0.29 percent is recovered and 0.23 percent is recognized as impairment loss and written off.[2] Repossession costs amount

2. The match is only approximate and not exact due to inflationary effects. Defaulted loans recovered respect previous period amounts, and thus they lose value to inflation.

	Model	Data	Period
Expenditure (GDP ratio)			
Private consumption	0.61	0.65	1995-2016
Private investment	0.19	0.18	1995-2016
Public consumption & investment	0.23	0.23	1995-2016
Exports	0.35	0.32	1995-2016
Imports	0.38	0.39	1995-2016
Shares (output ratio)			
Import share	0.28	0.30	1995-2008
Labor income share	0.60	0.67	1995-2016
External account (GDP ratio, in %)			
Net foreign assets (annualized)	-41.5	-83.5	1995-2016
Current and capital accounts	-0.8	-5.3	1995-2016
Trade balance	-3.0	-3.1	1995-2016
Financial sector, ratios			
Deposits-to-GDP ratio	0.41	0.46	1995-2016
Financial sector, Entrepreneurs			
Leverage ratio	1.2	1.2	1999-2008
Probability of default (in %)	3.6	3.6	1999-2008
Retail-wholesale interest rate spread (in p.p.)	1.6	1.7	1999-2008
Financial sector, Banks			
Probability of not fulfilling capital requirements (in %)	4.0	n.a.	
Capital-to-loans ratio (in %)	17.0	n.a.	
Endogenous capital buffer (in %)	3.0	n.a.	
Wholesale-deposits interest rate spread (in p.p.)	0.5	0.6	1999-2008
Financial sector, defaulted loans			
Defaulted loans-to-credit ratio (in %)	6.76	n.a.	
New defaulted loans-to-credit ratio (in %)	0.56	n.a.	
Defaulted loans recovered (in %)	0.29	n.a.	
Impairment-to-credit ratio (in %)	0.23	n.a.	
Repossession costs (credit ratio, in %)	0.37	n.a.	

TABLE 3. Key steady-state relationships.

Sources: *Banco de Portugal* data, National accounts data, and authors' own calculations.
Notes: Repossession costs are endogenously calibrated according to the retail-wholesale interest rate spread. We adjust the impairment-to-credit ratio to yield an overall loss-given default around 42 percent.

to 0.37 percent of total credit. This calibration results in a loss given default around 42 percent.[3]

3. Loss given default is understood herein as total losses in each period over the amount at risk given default.

The mechanism behind credit restrictions is endogenously calibrated so that agency problems do not arise in the steady state, but are triggered in the presence of shocks with large negative impacts on banks' terminal wealth. This results in a potential diversion of 16 percent of total loans.

Decomposing interest rate spreads

In this section we decompose the overall interest rate spread into the contribution of three well identified elements: the capital requirements-driven spread, the credit restrictions-driven spread, and the retail-driven spread. The wholesale interest rate spread is simply the sum of the two former components. To disentangle the role played by defaulted loans and credit restrictions on the spread decomposition, we simulate the complete model described in Section 3, henceforth named "banking model & DL & CR," and two additional models: the banking model deprived of credit restrictions (henceforth "banking model & DL"), and the banking model deprived of both defaulted loans and credit restrictions (hereinafter "banking model"). To appropriately compare the spreads across different models, we first calibrate and run the "banking model & DL & CR" and then successively deactivate parts of the model while fixing the values for all common parameters.

We perform the spread decomposition for three unexpected shocks of distinct natures: a demand side shock (on public consumption), a supply side shock (on technology), and a financial shock (on risk). All shocks have negative macroeconomic impacts and follow a standard autoregressive process of order 1 with a half-life of around 4 quarters. Shock sizes are merely illustrative.

Figure 3 depicts the interest rate spreads that follow an exogenous decline in public consumption. Credit restrictions remain slack at all time as the shock has minor impacts on banks' value, and therefore the corresponding spread component remains nil. Hence, the dynamics of the "banking model & DL & CR" and of the "banking model & DL" are identical.

The shock implies an increase in the capital requirements-driven spread, necessary to cope with an increased probability of violating regulatory requirements. Higher default rates that naturally arise during a demand-triggered recession lead to a decline in banks' returns, due to higher unexpected losses. Banks' capital is pushed downwards, therefore raising the probability of non-compliance with the regulatory requirements. The spread increase is empowered by defaulted loans, particularly over the medium term, as becomes visible when comparing the dynamics of the "banking model" with those of the "banking model & DL." Specifically, the downturn generated by the shock boosts corporate default and hence the amount of due loans. These impact negatively the income statement of banks, placing them closer to the minimum capital requirement and forcing them to charge

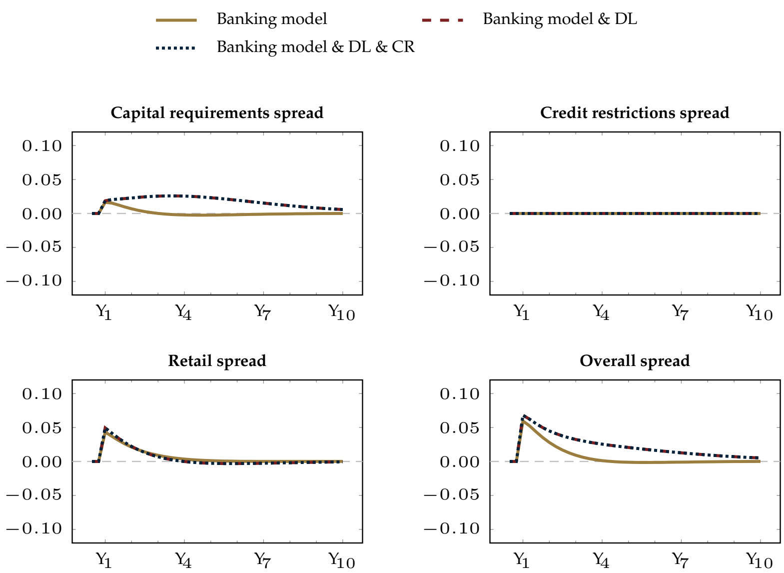


FIGURE 3: Government consumption shock.

Notes: The figure represents a negative government consumption shock of 1 percent of GDP.

even larger spreads to cover additional expenses and lost revenues while defaulted loans do not return to the pre-shock level. The effect is protracted in time, since defaulted loans have their own inertia and banks optimally prefer to spread impairment losses throughout time rather then recognizing them immediately in their income statement. The retail spread hike mostly reflects larger margins, required to cover higher corporate default rates.

For the technology shock, credit restrictions also remain slack at all times since banks' value is barely affected. The corresponding spread components is therefore nil at all times. However, the remaining spreads decline in this case as opposed to the previous shock, a direct consequence of a supply-triggered recession and hence of higher inflation. The decline in corporate default rates triggered by the lower real cost of credit that results from higher inflation is reflected into better banks' returns. This in turn implies a lower probability of violating regulatory capital requirements—and thus a decline in the corresponding spread. The downfall is strengthened by the decline in defaulted loans, since lower corporate default rates diminish the amount of due loans. As before, retail spreads are fairly identical across all models, and reflect lower corporate default rates associated with the lower real cost of credit.

In the case of a risk shock, the spread increase is severely amplified by both defaulted loans and credit restrictions, as banks' returns and value

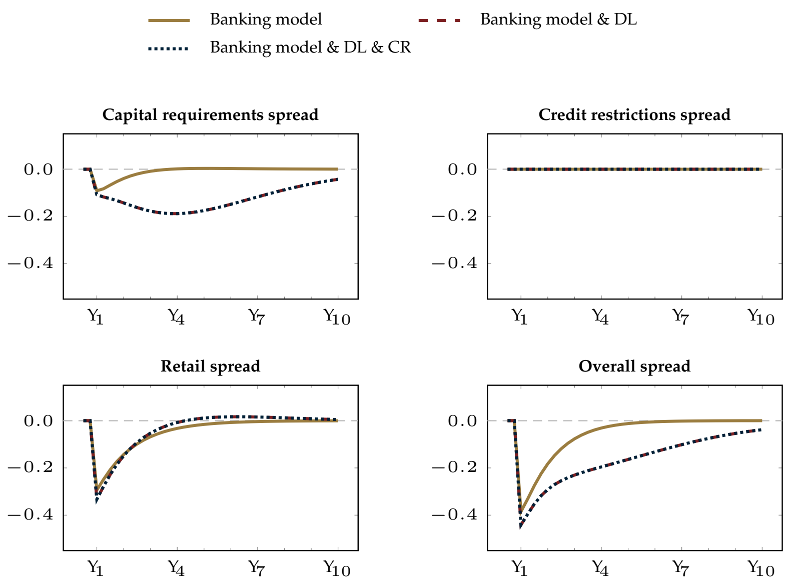


FIGURE 4: Technology shock.

Notes: The figure represents a 5 percent negative technology shock.

are severely affected in this case (Figure 5). The shock impacts directly the corporate default rate, leading to more expensive credit through a higher retail and wholesale spreads. The former translates greater margins on performing loans, required to cover higher losses on defaulting loans. The latter translates greater margins for banks, required to cope with higher losses on the aftermath of the income decline, triggered by the unexpected increase in corporate default.

Greater corporate default rates also lead to a substantial accumulation of defaulted loans. The wholesale spread is therefore further pushed upwards in the "banking model & DL," as banks require higher income to cope with regulatory requirements and defaulted loans opportunity and management costs. The powerful impact on bank returns and thus on their value forces bankers to hold back on credit, unfolding a large credit restrictions-driven wholesale spread hike, visible in the "banking model & DL & CR." Additionally, by negatively impacting returns, credit restrictions raise banks' leverage position. The risk of non-compliance with regulatory requirements therefore increases, feeding back to a higher capital requirements-driven spread.

Naturally, more expensive credit pushes corporate loans down, resulting in fewer investments and less capital accumulation. Entrepreneurs are forced to withhold investment decisions and hinder capital accumulation as external

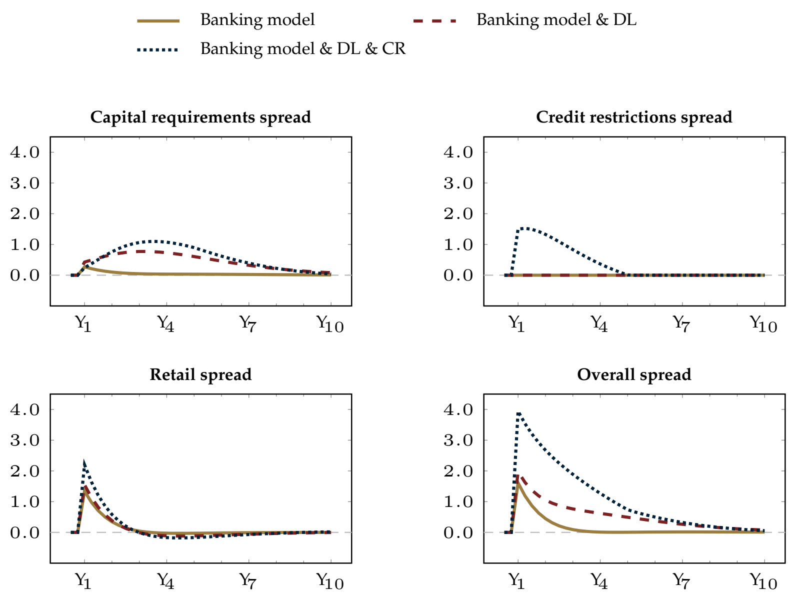


FIGURE 5: Risk shock.

Notes: The figure represents a 10 percent increase in risk.

finance collapses. In addition, corporate loans recover gradually as banks strive to keep the loans ratio and thus the probability of not complying with the regulatory requirements under control, generating a protracted recession.

Final remarks

In this article we borrow our work developed in Júlio and Maria (2018) and offer a model-based explanation of severe interest rate spread hikes, in line with those observed during the Great Recession. Our model proposes two novel features that are able to endogenously generate large interest spread movements: occasionally binding credit restrictions and defaulted loans. We borrow from the literature the endogenous capital requirements and the moral hazard-inspired credit constraint mechanisms. We then propose and develop an occasionally binding version of the latter mechanism, which is slack in the steady state but endogenously affects credit supply decisions when banks' capital is severely affected. As a result, credit is mostly demand/price driven but in some situations endogenously becomes supply/quantity driven. Simultaneously, we bring forth into the model a theory of optimal impairment loss recognition, which gives raise to an endogenous defaulted loans stock that bankers manage over time. Defaulted loans interact with regulatory capital requirements and credit restrictions.

In this article, we use the model to decompose the interest rate spreads in several components of interest, and analyze that decomposition under three shocks of distinct natures: a demand shock, a supply shock, and a financial shock. It is under the latter shock that the model provides a deeper insight on the interest rate spread decomposition, highlighting an higher and more persistent contribution of defaulted loans and credit restrictions to spread dynamics. We implemented an entrepreneurial risk shock as an illustrative financial disturbance, but conclusions qualitatively carry out for any type of financial disturbance affecting entrepreneurial or banking sectors, such as a nationwide risk shock, an entrepreneurial net worth shock, or a banks' capital shock. These type of disturbances have important impacts on banks' returns and hence banks' valuation, potentially triggering restrictive credit conditions and thus interest spread hikes.

References

Adolfson, Malin, Stefan Laseén, Jesper Lindé, and Mattias Villani (2007). "Bayesian estimation of an open economy DSGE model with incomplete pass-through." *Journal of International Economics*, 72, 481–511.

Benes, Jaromir and Michael Kumhof (2015). "Risky bank lending and countercyclical capital buffers." *Journal of Economic Dynamics and Control*, 58(C), 58–80.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist (1999). "The financial accelerator in a quantitative business cycle framework." In *Handbook of Macroeconomics*, vol. 1, Part C, edited by J. B. Taylor and M. Woodford, 1 ed., chap. 21, pp. 1341–1393.

Christiano, Lawrence, Roberto Motto, and Massimo Rostagno (2010). "Financial factors in economic fluctuations." Working Paper 1192, European Central Bank.

Christiano, Lawrence J, Martin S Eichenbaum, and Mathias Trabandt (2018). "On DSGE Models." *Journal of Economic Perspectives, Forthcoming*.

Gertler, Mark and Peter Karadi (2011). "A model of unconventional monetary policy." *Journal of monetary Economics*, 58(1), 17–34.

Gertler, Mark and Peter Karadi (2013). "QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool." *International Journal of Central Banking*, 9(1), 5–53.

Gertler, Mark, Nobuhiro Kiyotaki, and Albert Queralto (2012). "Financial crises, bank risk exposure and government financial policy." *Journal of Monetary Economics*, 59, S17–S34.

Júlio, Paulo and José R. Maria (2018). "An integrated financial amplifier: the role of NPL and occasionally binding constraints in output fluctuations." Working Papers 13/2018, Banco de Portugal.

Kumhof, Michael, Dirk Muir, Susanna Mursula, and Douglas Laxton (2010). "The Global Integrated Monetary and Fiscal Model (GIMF) - Theoretical structure." IMF Working Paper 10/34, International Monetary Fund.

Monitoring the equity risk premium in the S&P500

Nuno Silva
Banco de Portugal

October 2018

Abstract
This article uses a structural contingent claims model based on free cash flows to equity (FCFE) to derive the equity risk premium implicit in S&P500 stocks. This is done at the aggregate level for the period between 1999 and 2017. The results obtained are compared with those that come out from the traditional single-stage FCFE model. Two assumptions regarding long-term corporate growth expectations are made leading to slightly different results. Setting cash flow growth expectations based on 30-year U.S. bond yields the equity risk premium in December 2017 is found to be 4.6%, very close to the minimum value of the series. When a multiple of analysts forecasts on corporate 3 to 5-year earnings growth is used, the equity risk premium is found to be 5.2%, somewhat closer to the average equity risk premium estimated, which is approximately 5.9% in both cases. Under both cases the implied equity risk premium is found to be currently on a downward trend. The higher equity risk premium obtained in the second case is justified by the recent decoupling between analysts forecasts and the long-term risk free rate. This can be the result of analysts optimism on future firm performance but can also be related with the current abnormally low level of long-term interest rates. (JEL: G12, G13, G32)

Introduction

What discount rate is implicit in current stock prices? What expectations about a firm's future performance are consistent with its current market capitalization? These are questions equity analysts often try to answer before issuing recommendations on whether to buy or sell a firm stock. With bond yields sticking around very low levels and the S&P500 staying close to its all-time maximum in the longest bull market in its history, answering these questions has become increasingly relevant not only for financial analysts and academics but also for regulators and macroprudential authorities all over the world. Implied growth expectations not compatible with economic projections or an implied equity risk premium significantly below their historical average signal that investors are either

Acknowledgements: I would like to thank António Antunes, Nuno Alves, Isabel Horta Correia, Diana Bonfim, Luisa Farinha, José Faias, Pedro Moreira and António Santos for their comments. The opinions expressed in this article are those of the author and do not necessarily coincide with those of Banco de Portugal. Any errors and omissions are the sole responsibility of the author.

E-mail: nrsilva@bportugal.pt

too optimistic on firms future performance or have a risk appetite above the one observed on average in the past. As investors long term projections and risk appetite often move with the business cycle, both cases are usually interpreted as early warning indicators of a potential future reappraisal of asset prices. Equity analysts often answer the aforementioned questions by reverse engineering discounted cash flow models (DCF). The procedure is simple. In the case of the constant growth free cash flow to equity (FCFE) model, equity value corresponds to the perpetual sum of future expected cash flows available to shareholders discounted at a rate μ_E that takes into account equity risk. Assuming that FCFE grows forever at a constant rate g below the discount rate, the usual perpetuity formula gives us the equity value:

$$E_0 = \frac{FCFE_0}{\mu_E - g}. \tag{1}$$

Assuming a discount rate based on some asset pricing model as the CAPM, analysts can back out implied FCFE growth rates and compare them with their projections. Alternatively, using their growth projections they can back out the implied discount rate and compare it with the outcome from their preferred asset pricing model. This type of exercise is very popular among practioners and there is a great number of academic papers on this (e.g. Gebhardt *et al.* (2001), Easton *et al.* (2009) and Ohlson and Juettner-Nauroth (2005)). However, it has two important weaknesses. First, equity value is very sensitive to changes both on the discount rate and on growth expectations. Second, there is substantial model risk. In this regard, it is noteworthy that most models used in practice ignore default risk and the effect on equity valuation of leverage dynamics.

This paper does an exercise similar to the one just explained. In this case, long-term growth expectations are assumed and the implied equity risk premium is derived. This is done however using a contingent claims model able to take into account default risk, operating leverage and financial leverage. The approach here proposed also benefits from incorporating information on credit default swap spreads. The exercise in this article is done at the aggregate level using accounting and market data of 205 firms belonging to the S&P500 for the period between 1999 and 2017. Two alternative growth rate assumptions are considered. First, long-term corporate growth expectations are set based on U.S. 30-year bond yields. Second, normalized analysts 3 to 5-year earnings growth forecasts are used. Growth expectations in these second case have a mean value equal to the ones in the first case, but they are better able to capture analysts' optimism in firm fundamentals. Under both cases the implied equity risk premium is found to be on a downward trend, but still at a level above the one observed in the late nineties. The equity risk premium derived using the structural approach proposed in this article is shown to be more stable than the one that comes out from the application of the traditional single-stage FCFE model.

Related literature and contribution

Contingent claims models, also known as structural models of corporate liabilities, started with Merton (1974). In this model, a firm financed by equity and a single pure discount bond is considered to honour its commitments if the market value of its assets at debt maturity is higher than debt's nominal value. If not, the firm defaults and shareholders receive 0. In Merton's model equity can thus be seen as a call option on the firm assets with strike equal to nominal debt. Empirical applications of this seminal model showed poor results, but its tremendous insights opened the door to a huge list of academic and non-academic papers that tried to relax its initial restrictive assumptions in order to better fit the data.[1] In most of the models that followed Merton (1974), the market value of a firm assets has been seen as an exogenous traded asset. Breaking with this tradition, Goldstein *et al.* (2001) propose a model where the asset value is seen as a fictive non-traded security whose value corresponds to the perpetual sum of all future earnings before interest and taxes (EBIT). The latter is assumed to follow a geometric Brownian motion implying that the underlying asset is lognormally distributed. In this framework contingent contracts such as equity, bonds and options are all interrelated through the same market price of risk. The lognormal EBIT assumption in this model is not compatible however with negative EBIT values, something often observed. In addition, EBIT is an income account and thus its relation with the firm capacity to generate cash flow is not direct. The model presented in this article overcomes these issues by defining the state variable as the sum of the cash flow from operating and investment activities, interest expenses and any costs termed fixed. The latter is seldom negative and thus more suitable to be modelled as a geometric Brownian motion. Adding up non-financial fixed costs such as selling, general and administrative expenses (SG&A) allows us to consider operating leverage in addition to financial leverage. Debt dynamics are also different. While Goldstein *et al.* (2001) consider that debt only increases when the market value of assets goes up to a level where the firm wants to restore its optimal capital level, in this paper debt is continuously sold at market price meaning that net borrowing contribution to the FCFE is lower whenever the firm is performing poorly.[2] This debt dynamics has already been assumed by Ericsson and Reneby (2003) in a very similar model.[3] The estimation procedure is nevertheless very different. Though the project value in their paper derives its value from

1. Popular industry applications include Moody's EDF, the CreditGrades model from Deutsche Bank, Goldman Sachs, JPMorgan and the RiskMetrics Group and Credit Suisse CUSP model.

2. The roll-over process of the initial stock of debt is nevertheless not taken into account. See He and Xiong (2012) on this regard.

3. The model in this paper differs from theirs only on the state variable definition, the addition of operating leverage and the division of debt between interest-bearing and non-interest bearing.

firm fundamentals (earnings before taxes in their case), this is not relevant in their estimation procedure. As a result, their asset value estimates are not compatible with observed fundamentals for the estimated model parameters. In addition, while Ericsson and Reneby (2003) use the model for bond pricing, the objective of this study is to measure the evolution of the equity risk premium implied by stock prices. This decomposition may help analysts and macroprudential authorities understand whether stock prices are fairly priced, and evaluate the risks to the financial system.

The model

The FCFE model of equity valuation is one of the most popular among equity analysts. FCFE is a measure of how much cash is available to shareholders after all expenses, investment and net borrowing is taken into account. Firms can distribute this as dividends, buy back stocks or do nothing leading to an increase in cash accounts. Negative FCFE means that the firm has either to decrease its cash reserves, sell own shares held in its portfolio or issue additional equity to finance its activities. In contrast to dividend discouting, FCFE-based valuation models recognize that firms can also compensate their shareholders by repurchasing stocks, something that has become increasingly popular in the last decades. Taking the cash flow statement as the starting point to compute the FCFE we have that

$$FCFE_t = CFO_t + CFI_t + d_t, \tag{2}$$

where CFO_t refers to the cash flow from operations, CFI_t is the cash flow from investment and d_t corresponds to net borrowing. CFO comprises all cash flow the firm receives from its regular business activities. This includes all cash flow received from customers net of all expenditures with suppliers, fixed costs, corporate taxes and interest expenses. CFO is generally positive, though during recessions it may become negative, even for firms not in financial distress. In contrast to CFO, CFI is usually negative as it comprises investments in long-term assets such as property, plant and equipment (PP&E) and long term investments in other companies. However, it can also be positive when a firm sells its investments. Net borrowing is very irregular, but it tends to be positive over time following firm growth. As explained in the introduction, in the single-stage FCFE model, this is assumed to follow an infinite horizon discrete time deterministic trend process. In this article it is considered instead that $FCFE_t$ is a continuous time stochastic process with a finite horizon. This difference will turn the model significantly more complex but it will also allow us to better take into account the effect of business risk, default risk, operating leverage and financial leverage on the value of future FCFE. The reader less interested in how this is done can skip to the next section.

Before presenting the free cash flow to equity dynamics, for reasons that will become clear soon, consider adding and subtracting in equation (2) fixed costs, q_t, and after-tax interest expense, which is hereafter presented as the product of the firm after-tax coupon rate c and total liabilities L_t:

$$FCFE_t = (CFO_t + CFI_t + q_t + cL_t) - q_t - cL_t + d_t \tag{3}$$

The first term in brackets will be hereafter denoted as δ_t and assumed to follow a geometric Brownian motion with drift μ_δ and volatility σ (see equation (A.1) in the Appendix). The geometric Brownian motion is the same stochastic process Black and Scholes (1973) used to model stock prices. In this case it states the idea that in each moment in time the continuous compounding growth rate of our state variable δ_t follows a normal distribution with mean $\mu_\delta \Delta t$ and variance $\sigma^2 \Delta t$. This leads to a highly persistent process, which cannot take negative values.[4] For positive μ_δ and σ, the longer the time interval the higher is the expected value of our state variable and the uncertainty around its value. q_t and L_t are assumed to grow deterministically $\alpha q_t \Delta t$ and $\alpha L_t \Delta t$, respectively (see equations (A.2) and (A.3) in the Appendix). For simplicity, it is considered that debt is perpetual and gives right to a constant coupon rate c, which should be seen as a weighted average of interest expenses in interest-bearing and non-interest-bearing debt. The latter corresponds to a share φ of L_t. All new debt issues are considered to be perpetual. Non-interest-bearing debt is issued at nominal value, while interest bearing debt is issued at market value. The latter implies that the total cash inflow from new debt issues, i.e. d_t, is a function of the firm financial position at each moment in time. The lower the probability of the firm defaulting the higher the amount of cash flow it receives for the same level of additional nominal debt. Figure 1 (Panel A) shows examples of different δ_t paths along with total costs with coupon payments and fixed costs. Cash inflows arising from new debt issues under each δ_t path are presented in Figure 1 (Panel B).

In the traditional single-stage FCFE model, FCFE never takes negative values as it is assumed to grow at a constant rate up to infinity. In this model, however, δ_t may become less than $q_t + cL_t - d_t$ implying a negative $FCFE_t$. This is the case of the red path in Figure 1 (Panel C). Whenever the FCFE is negative, shareholders must decide whether they are willing to inject capital in the firm. They will do it until time τ, the first time δ_t hits a lower boundary $\overline{\delta}_t$, which is determined by solving equation (A.8) in the Appendix. This condition is known in the optimal stopping time literature as the smooth pasting condition. The intuition behind this is that shareholders are willing to inject capital as long as the equity value after the capital increase is higher than the amount of cash flow they inject. q_t and cL_t are crucial in shareholders

4. The sum of fixed costs to CFO significantly mitigates this problem.

default decision. Everything else equal, the higher the fixed costs the firm runs in its productive process (i.e. the higher its operating leverage) and its financial duties (i.e. its financial leverage) the earlier shareholders will give up the firm. It is important to emphasize that even if shareholders are liquidity constrained, in a world with no information problems and restrictions to capital movements, as long as the market value of equity after the capital increase is above the capital increase there will always be a price at which the firm will be able to raise capital. This occurs because, no matter the consequences in terms of dilution, it is always better for shareholders to raise capital than to lose the firm and receive nothing. The default barrier in our simulation exercise is presented in Figure 1 (Panel A) along with potential δ_t trajectories. Similar to L_t and q_t, $\overline{\delta}_t$ grows at rate α. Whenever the barrier is hit, the firm is closed and distress costs are incurred. These correspond to legal costs and value destruction caused by fire sales and loss of intangible value. In this case the firm stakeholders receive βA_τ, where A_τ corresponds to the discounted present value of all future δ_t up to infinity. Mathematically,

$$A_\tau = \frac{\overline{\delta}_\tau}{r + \overline{m}\sigma - \mu_\delta}, \tag{4}$$

where r is the after-tax risk-free interest rate and $\overline{m}$ is the market price of risk (i.e. the amount of return demanded by investors by unit of risk). $\overline{m}$ can be interpreted as the project Sharpe ratio. The best way to understand this is to think that the firm continuously holds a project that generates δ_t up to infinity and whose value, A_t, corresponds to the perpetual sum of all future δ_t.[5] If δ_t becomes unsatisfactory the firm is closed and the project is sold to a competitor firm. The project is infinite-lived but the firm is not. The β accounts for the fact that the firm stakeholders only receive a share of the project value when this occurs. The usual pecking order implies that shareholders only receive something if that share, i.e. βA_τ, is higher than nominal debt L_τ. For simplicity, it is assumed that β is sufficiently low so that shareholders receive nothing in case of liquidation. β affects equity value through the cash inflow, d_t, the firm receives when it issues new debt. The higher is β, the more debt holders recover after default, and thus the higher is the capital inflow whenever the firm issues new debt. β is thus a relevant parameter for equity valuation in this model.

For valuing this firm's stock, it is assumed the existence of a unique probability measure by which the discounted value of δ_t becomes a martingale.[6] Equity can then be valued as the discounted sum of all future

5. By applying Itô's lemma to the asset function it is possible to derive the dynamics of this fictive security. Since the market price of risk is assumed to be constant we have that $\sigma_A = \sigma$.

6. A martingale is a stochastic process where the expected value of the next observation in the process equals the previous one. See Björk (2009) for a discussion on the technical conditions required for the existence of this unique probability measure.

after-tax free cash flows up to the moment the firm is closed plus its current after-tax cash position.[7] The tax rate $\bar{t}$ is interpreted as a weighted average of the expected dividend and capital gains tax rates.[8] Equity value in this model is obtained solving the expression in equation (A.4) in the Appendix. The reader less familiar with the idea of risk neutral pricing may find strange discounting the future FCFE at the risk-free rate. However, under this framework investors compensation for taking risk is taken into account by changing the probabilities of the different outcomes instead of demanding an higher discount rate. The two approaches are equivalent. Risk neutral pricing allows us to price all contracts that are contingent on the firm's business without having to compute their specific discount rate. This can be very appealing whenever one wants to extend the methodology to other contingent claims such as credit default swap contracts (CDS). Equation (A.10) in the Appendix explains how CDS can be priced in this model for a general level of senior liabilities X. Figure 1 (Panel D and E) illustrates Equity and CDS spreads in the context of our simulation exercise.

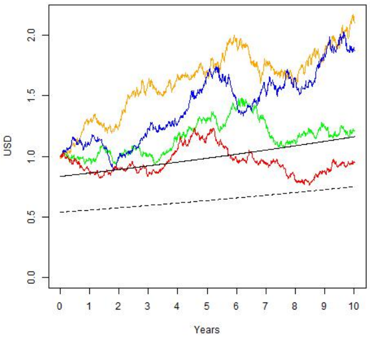
Equation (A.4) can be used for equity valuation, whenever one is able to provide estimates on all model inputs. Alternatively, as better explained in the next section, one can use observed equity prices to extract the market price of risk $\overline{m}$ implied by stock prices. This can then be used to compute the equity risk premium and the cost of equity. The latter corresponds to the drift of the equity process, which is given by

$$\mu_{E_t} = r + \overline{m}\sigma_{E_t}, \tag{5}$$

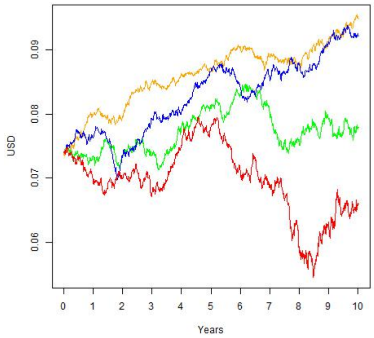
where σ_{E_t} refers to equity return volatility, whose formula is provided in equation (A.9) in the Appendix. The cost of equity in our simulation exercise is presented in Figure 1 (Panel F). In contrast to the Black-Scholes model, equity volatility is not constant in this model due to financial and operating leverage. As result, the equity risk premium and the cost of equity are also not constant.

7. Substantial cash holdings are a signal of potential dividends and stock buybacks. For this reason, cash holdings are relevant for a shareholder that takes a "control" perspective over the firm and were added in the equity valuation formula.

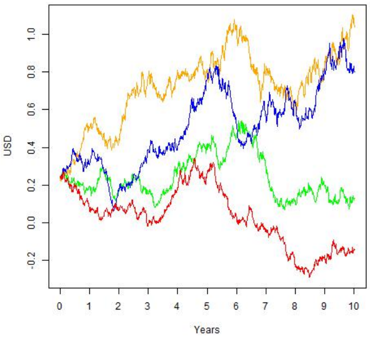
8. Notice that only these taxes need to be taken into account in equation (A.4) since the state variable already accounts for corporate level taxation.



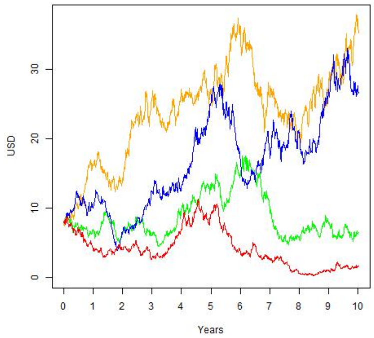
(A) Examples of different δ_t paths.



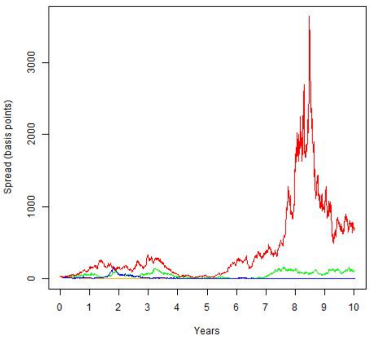
(B) Cash inflow from new debt issues.



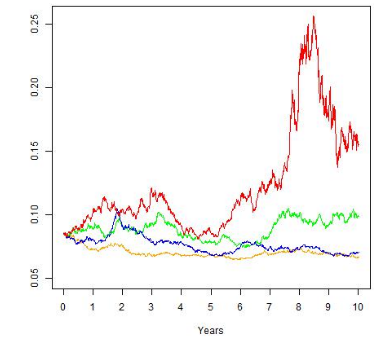
(C) FCFE



(D) Equity value



(E) CDS (5-years)

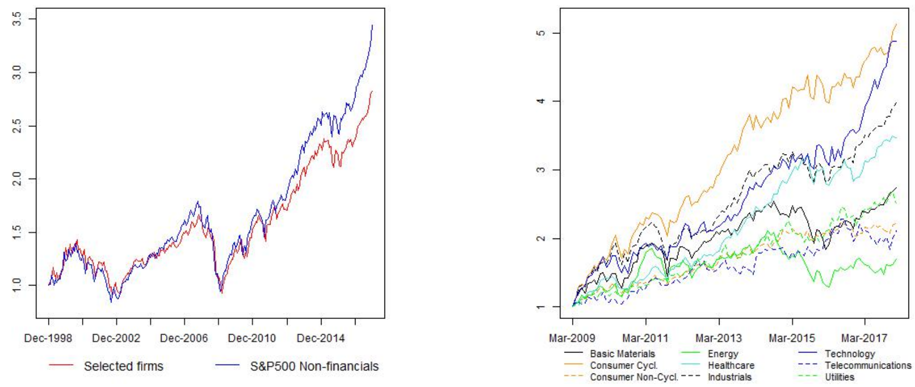


(F) Cost of equity

FIGURE 1: Simulation exercise. $\delta_0 = 1$, $r = \mu_\delta = \alpha = 0.033$, $\sigma = 0.106$, $q_0 = 0.79$, $c = 0.016$, $L_0 = 2.65$, $\overline{m} = 0.133$, $\beta = 0.049$, $\bar{t} = 0.15$, $Cash = 0.23$, $X = 1.64$ and $\varphi = 0.57$. The values used are normalized values based on December 1998 calibration. In Panel A, the continuous think black line corresponds to the sum of interest and fixed costs and the dashed black line is the default barrier.

Data and calibration

This section presents the data and calibration procedure used in this study. All data is collected from Thomson Reuters for the period between December 1998 and December 2017. Accounting data is collected with annual frequency, while market data is collected with monthly frequency. The initial dataset corresponds to 406 non-financial firms composing the S&P500 in December 2017. This was subsequently restricted to 205 firms in order to include only those firms for which all the required data is avaliable for the whole period. The large majority of the firms excluded did not exist or were not listed in December 1998. Except for technology, basic materials and telecommunications, sampled firms represent more than 60% of each sector market capitalization. This figure falls to approximately 40% for the technology and basic materials sectors. The telecommunications sector is not represented in the sample. Figure 2 (Panel A) compares the evolution of the market capitalization for these firms with an index based on the initial sample of firms controlling for entrances and exits. Figure 2 (Panel B) shows similar indices per sector of activity, but starting in March 2009, when market indices reached their bottom. Despite the two series following a similar trajectory, it is clear that firms on our sample have had an increase in market capitalization below others. Rather than a sector underrepresentation problem, this seems related with the predominance of mature firms in the sample. A point can obviously be made that the selected sample of firms does not totally capture the recent increase in the $S\&P500$. Though true, the fact that our sample of firms is constant across time better allow us to study what is going on.



(A) Market capitalization (1998-2017) normalized in December 1998. Comparison between sample and initial aggregate.

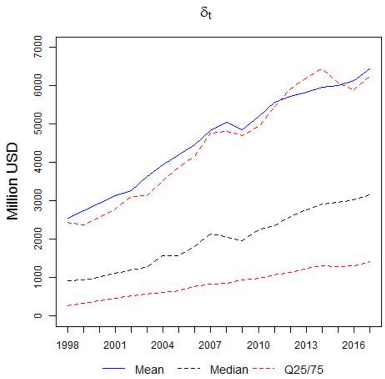
(B) Market capitalization per sector of activity (2009-2017) normalized in March 2009.

FIGURE 2: Market capitalization.

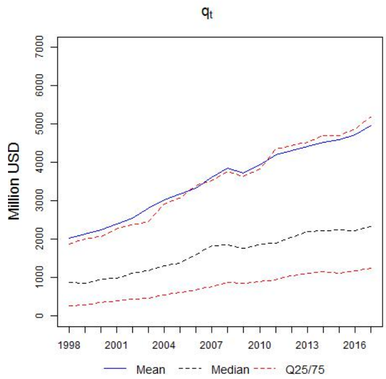
The model presented in the previous section has 14 inputs, notably, the sum of the cash flow from operations, the cash flow from investment activities, fixed costs and after-tax interest expenses (δ_0), fixed costs (q_0), short term financial assets ($Cash_0$), total liabilities (L_0), senior liabilities (X), the share of non-interest-bearing liabilities (φ), after-tax coupon rate on total liabilities (c), dividend and capital gains tax rate ($\bar{t}$), the after-tax risk free rate (r), expected growth rate of debt (α), expected growth rate of the state variable (μ_δ), business risk (σ), the amount of return demanded by investors by unit of risk ($\overline{m}$) and a recovery rate-related parameter (β). δ_0, q_0, $Cash_0$, L_0, X_0, φ_0 and c are readily available from financial documentation and presented in Figure 3. δ_0 was computed summing cash flow from operations, cash flow from investment activities (smoothed), SG&A and after tax interest expense. SG&A, which includes all costs that cannot be tied directly to the firm's output, is thus used as proxy for firms' fixed costs, q_0. SG&A represents on average 76% of our state variable. $Cash_0$ corresponds to the cash account plus other short term financial assets. L_0 corresponds to total non-equity liabilities excluding minority interests. X_0 equals L_0 minus long-term debt. φ was set as 57%, which corresponds to 1 minus the ratio of total debt outstanding to total liabilities in Reuters. Finally, c was computed as interest expense divided by total nominal liabilities and multiplied by 1 minus the corporate tax rate, which was assumed to be 20%.[9] δ_0, q_0, $Cash_0$, L_0 and X_0 correspond to the sum of all individual firms observations. c is the weighted average based on each firm end-of-month market capitalization. r was obtained multiplying the yield on 30-year U.S. Treasury bonds by 1 minus the interest income tax rate, which was assumed to equal 35%. $\bar{t}$ was set at 15%. α was assumed to be equal to μ_δ in order to keep the expected value of the leverage ratio constant across the firm's life.

Two different assumptions are considered regarding μ_δ. First, it was assumed that corporate long-term growth rate equals the risk free rate (i.e. $\mu_\delta = r$). This assumption is very common in equity valuation. The idea behind is that one day the firm will stop over or underperforming the economy and converge to its long-term nominal rate of growth. The results obtained were compared with the ones that come out from assuming that μ_δ is a multiple of analysts 3 to 5-years earnings forecasts (compounded growth rate). These were taken from Thomson Reuters I/B/E/S database and are presented in Figure 4 (Panel A). Studies on analysts' capacity to correctly forecast corporate growth have generated mixed results. For the sample of firms considered, a moderate correlation (42%) is found between the compounded annual average growth rate of analysts' forecasts and the compounded annual average growth rate of our state variable between 1999 and 2017. More

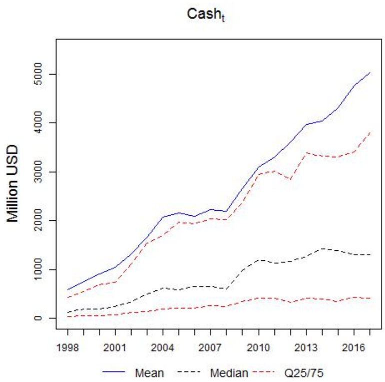
9. The corporate tax rate is not very important in this model because the CFO is computed after tax. Changing the corporate tax rate assumption will only slightly affect the firm's financial leverage and thus the optimal default barrier.



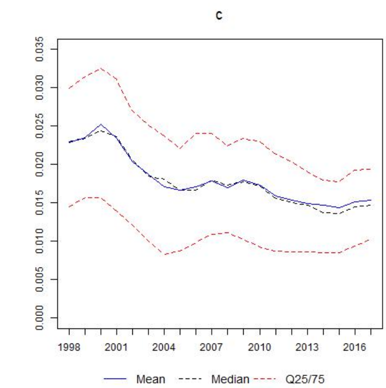
(A) State variable



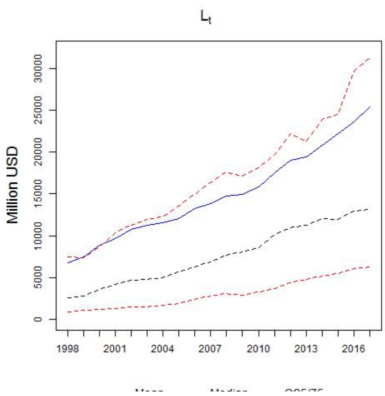
(B) Fixed costs



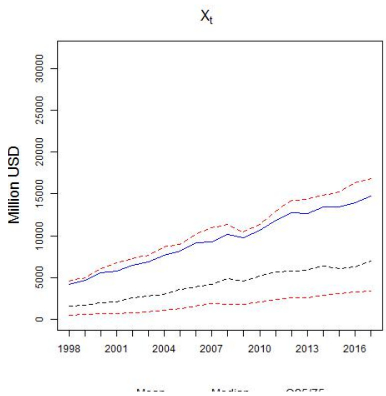
(C) Cash and equivalents



(D) Coupon rate



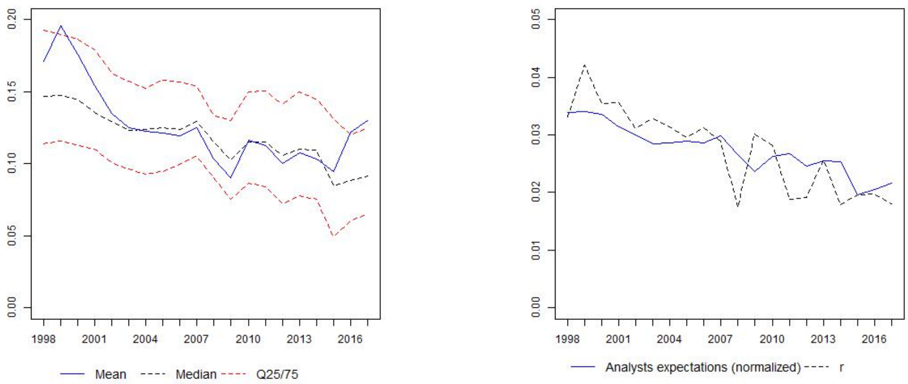
(E) Total debt



(F) Senior debt

FIGURE 3: Firm fundamentals.

interestingly, a correlation of 89% is found between median analysts' forecasts and 30-year U.S. Treasury bonds during the same period (Figure 4 Panel B). This suggests that analysts' forecasts can be used as an alternative to long-term nominal rates. The fact that these forecasts reflect analysts' momentum on firm fundamentals is useful to understand what is leading stock markets. In line with the literature that points out that analysts' forecasts are generally too optimistic, the average annual growth rate of analysts' forecasts is found to be approximately 6 percentage points above the annual growth rate of our state variable. Analysts' forecasts are also very high to be thought as sustainable long-term growth rates. For these reasons the obtained figures were scaled down by multiplying by the mean ratio between r and analysts' growth forecasts.[10] The median value was then chosen as proxy for long term growth expectations. The median value was preferred to the weighted mean because it is less sensitive to abrupt changes in analysts' forecasts on some very large firms. This is particulary relevant given the high sensitivity of equity value to this parameter in this model.



(A) Analysts' growth forecasts (end-of-year).

(B) Normalized median analysts' forecasts and 30-year U.S. bond yields (end-of-year).

FIGURE 4: Long-term growth expectations.

In line with the model assumptions, σ, which captures business risk, was considered to be constant across the whole estimation period. As it is clear from Figure A.1 in the Appendix, this does not imply constant equity volatility. Each firm σ was estimated through a robust linear regression of the logs difference of the state variable δ_t on a constant. Figure 5 shows an histogram of these estimates. Approximately 40% of our σ estimates lie between 8% and 15%. The 10th and 90th percentile of the distribution are 5.2% and 25.1%, respectively. Since the exercise in this article was carried at

10. The use of a multiple of analysts' forecasts is also done in the well-known Yardeni model (Yardeni (2003)). This multiple is not computed in the same way, though.

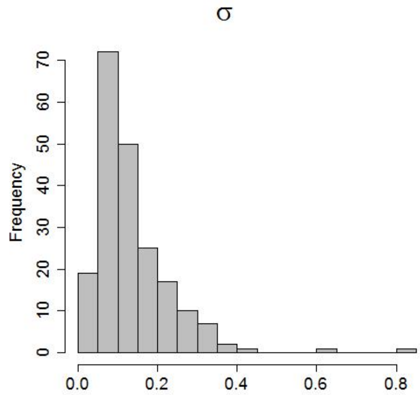


FIGURE 5: Histogram of σ estimates.

the aggregate level, σ was set as the median of individual volatility estimates (i.e. 0.106). Finally, $\overline{m}$ and β are estimated by solving a system of equations where $\overline{m}$ and β are chosen so that equity value in the model matches the observed market capitalization and CDS spreads. A weighted average of the CDS spreads (5-years) of 62 firms is used (Figure 6).[11] Given the lack of CDS data of good quality for the period before 2009, in this period $\overline{m}$ and β were estimated assuming a recovery rate of 0.23. This corresponds to the average recovery rate obtained during our exercise for the period after 2009.

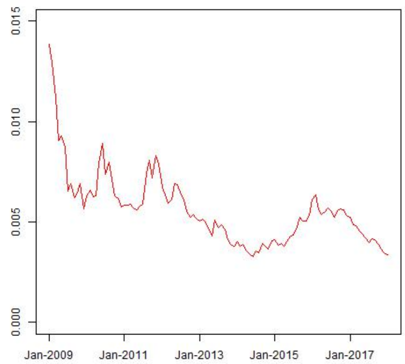
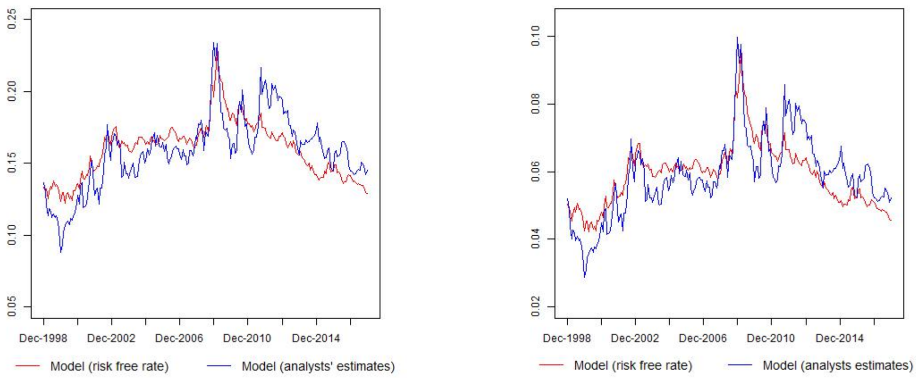


FIGURE 6: Credit default swap spreads (5-years).

11. This procedure was carried with monthly frequency between December 1998 and December 2017. Monthly accounting figures were linearly interpolated from annual figures.

Results

Figure 7 shows the market price of risk and the equity risk premium obtained assuming growth expectations based on the risk-free rate and on long-term analysts' forecasts, respectively.[12] A mean equity risk premium of approximately 5.9% is observed in both cases. The two series also have a similar pattern, marked by very low values in the beginning and in the end of the estimation interval and very high values during the financial crisis. Currently, the equity risk premium is in a downward trend reaching 4.6% in the end of 2017 when the risk free rate is used and 5.2% when analysts' forecasts are used. It is interesting to note that while in the first case the equity risk premium is very close to the minimum of the series, in the second case it is somewhat closer to the average. The equity risk premium is nevertheless significantly more volatile in this second case.[13]



(A) Market price of risk.

(B) Equity risk premium.

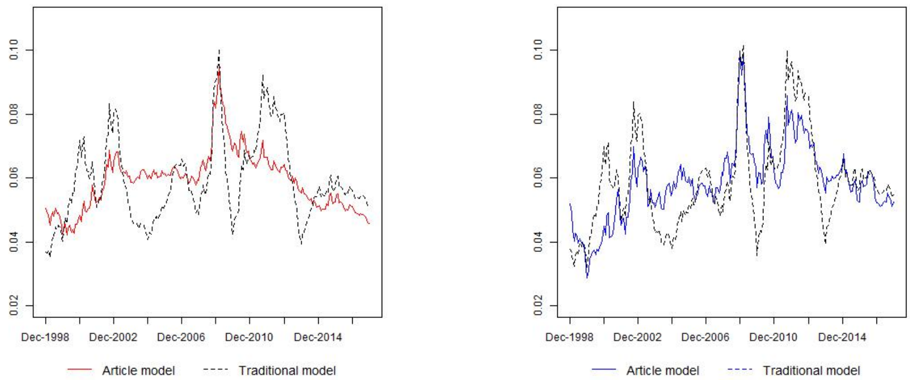
FIGURE 7: Model implied market price of risk and equity risk premium.

The results obtained with the model presented in this article are not materially different from those that come out from the traditional single-stage FCFE model (Figure 8). Adjusting for taxes and cash holdings an

12. The equity risk premium and the market price of risk are not a multiple of each other because model implied equity volatility is not constant across time (see Figure A.1 (Panel A) in Appendix). Though business risk measured by the state variable volatility is constant, financial and operating leverage lead to stochastic volatility. The latter is nevertheless far from reproducing the observed equity volatility (see Figure A.1 (Panel B) in Appendix), which is computed as the annualized standard deviation of daily equity returns.

13. In this regard, it is interesting to note that when growth expectations are equal to the risk free rate, despite some small spikes being observed during the European sovereign debt crisis, the implied equity risk premium is very far from the levels observed during the peak of the financial crisis. In contrast, when growth expectations are based on financial analysts forecasts, the equity risk premium jumps significantly in the second half of 2010 and 2011.

implied equity risk premium of 5.9% is also found in this case. The two series have nevertheless a correlation that is far from perfect (56% when growth expectations equal the risk-free rate and 74% when growth expectations are proxied by analysts' forecasts). This is largely the result of the series that comes out from the traditional single-stage FCFE model being significantly more volatile under both expected growth assumptions. The very significant increases in the equity risk premium observed in March 2001, September 2002 and September 2011 are good examples of this. These spikes are observed under both growth rate assumptions in the case of the traditional FCFE model. However, when the structural model is applied these spikes are very contained, especially when growth expectations equal the risk-free rate.



(A) μ_δ based on the long-term risk-free rate.

(B) μ_δ based on analysts' forecasts.

FIGURE 8: Equity risk premium. Comparison with the single-stage traditional FCFE model.

Concluding remarks

This article derives the equity risk premium implicit in S&P500 stock prices using a single-stage FCFE-based structural model. An aggregate perpective was followed. In line with literature and historical observation, a mean equity risk premium of approximately 5.9% is found for the period between 1999 and 2017. Independently of using the risk free rate or a multiple of analysts' forecasts the equity risk premium is found to be currently on a downward trend. The level observed in December 2017 is nevertheless different depending on how growth expectations are set. While in the first case, the equity risk premium is found to be 4.6%, very close to the minimum of the series, in the second case it is found to be 5.2%, somewhat closer to average. This difference is justified by the recent decoupling of normalized analysts' forecasts from 30-year U.S. bond yields. This decoupling can be interpreted as a signal of analysts optimism on firms future performance. However, it can also be related with the current abnormally low level of long-term interest rates given the U.S. economy fundamentals.

References

Björk, Tomas (2009). *Arbitrage theory in continuous time*. Oxford university press.

Black, F. and M. Scholes (1973). "The pricing of options and corporate liabilities." *Journal of Political Economy, 81(3)*, pp. 637–654.

Easton, Peter *et al.* (2009). "Estimating the cost of capital implied by market prices and accounting data." *Foundations and Trends in accounting*, 2(4), 241–364.

Ericsson, Jan and Joel Reneby (2003). "The valuation of corporate liabilities: theory and tests. SSE." Tech. rep., EFI Working Paper Series in Economics and Finance.

Gebhardt, William R, Charles MC Lee, and Bhaskaran Swaminathan (2001). "Toward an implied cost of capital." *Journal of accounting research*, 39(1), 135–176.

Goldstein, Robert, Nengjiu Ju, and Hayne Leland (2001). "An EBIT-based model of dynamic capital structure." *The Journal of Business*, 74(4), 483–512.

He, Zhiguo and Wei Xiong (2012). "Rollover risk and credit risk." *The Journal of Finance*, 67(2), 391–430.

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

Ohlson, James A and Beate E Juettner-Nauroth (2005). "Expected EPS and EPS growth as determinants of value." *Review of accounting studies*, 10(2-3), 349–365.

Yardeni, Edward (2003). "Stock valuation models (4.1)." *Prudential Financial Research*.

Appendix

δ_t, q_t and L_t dynamics are assumed to be given by the following differential equations:

$$\frac{d\delta_t}{\delta_t} = \mu_\delta dt + \sigma dW_t^{\mathbb{P}}, \tag{A.1}$$

$$dq_t = \alpha q_t dt \tag{A.2}$$

and

$$dL_t = \alpha L_t dt. \tag{A.3}$$

It is considered that nominal debt L_t is composed by a non-interest bearing component, L_t^{NonInt}, and an interest bearing component, L_t^{Int}. Each of these components follows an ordinary differential equation similar to the one given in equation (A.3). As a result, both components are a constant fraction of L_t. It is considered that $L_t^{NonInt} = \varphi L_t$ and $L_t^{Int} = (1-\varphi) L_t$. The owner of the interest-bearing component earns a coupon payment equal to $c^{Int} L_t^{Int}$. Since both components are a constant fraction of L_t, we have that $c^{Int} = \frac{c}{1-\varphi}$.

Equity value is obtained solving the below expression

$$E_0 = \left(1-\bar{t}\right)\left(Cash_0 + E^{\mathbb{Q}}\left[\int\limits_0^{+\infty} e^{-rs}(\delta_s - q_s - cL_s + d_s)\, 1_{\{\tau>s\}} ds|\mathcal{F}_0\right]\right), \tag{A.4}$$

where the term within the integral corresponds to the sum of all future FCFE until firm liquidation. The expected value of the discounted sum of all future $\delta_s - q_s - cL_s$ is standard in the contingent claims pricing literature. For the sum of all future d_s we have to decompose it between cash inflow from non-interest-bearing and interest-bearing debt:

$$\begin{aligned} E^{\mathbb{Q}}\left[\int_0^{+\infty} e^{-rs} d_s 1_{\{\tau>s\}}|\mathcal{F}_0\right] ds = {} & E^{\mathbb{Q}}\left[\int_0^{+\infty} e^{-rs} d_s^{NonInt} 1_{\{\tau>s\}} ds|\mathcal{F}_0\right] \\ & + E^{\mathbb{Q}}\left[\int_0^{+\infty} e^{-rs} d_s^{Int} 1_{\{\tau>s\}} ds|\mathcal{F}_0\right]. \end{aligned} \tag{A.5}$$

Since non-interest-bearing debt is sold at nominal value we have that

$$E^{\mathbb{Q}}\left[\int_0^{+\infty} e^{-rs} d_s^{NonInt} 1_{\{\tau>s\}} ds|\mathcal{F}_0\right] = E^{\mathbb{Q}}\left[\int_0^{+\infty} e^{-rs} \mu_\delta \varphi L_s 1_{\{\tau>s\}} ds|\mathcal{F}_0\right]. \tag{A.6}$$

The solution of equation (A.6) is standard in the literature. For interest-bearing debt, which is sold at market value, it is assumed that the value of all future cash inflows must equal the value of all coupons that accrue to the new debt issues plus their share on the recovered value after firm liquidation. Mathematically,

$$E^{\mathbb{Q}}\left[\int_0^{+\infty} e^{-rs} d_s^{Int} 1_{\{\tau>s\}} ds|\mathcal{F}_0\right] = E^{\mathbb{Q}}\left[\int\limits_0^{+\infty} e^{-rs}\left(cL_s - cL_0\right) 1_{\{\tau>s\}} ds|\mathcal{F}_0\right] + \left(1-\varphi\right)\beta E^{\mathbb{Q}}\left[e^{-r\tau}\left(\overline{v}_s - \overline{v}_0\right)|\mathcal{F}_0\right], \tag{A.7}$$

where $\overline{v}_0$ is the project value that leads the firm to default at time 0. The solution to equation (A.7) is standard in the literature.

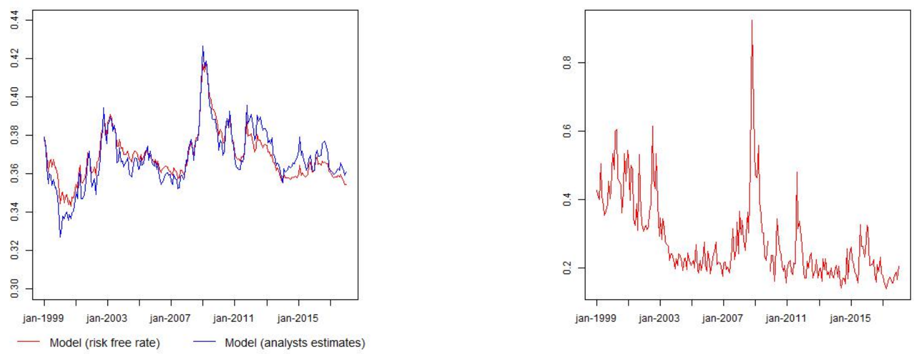
The smooth pasting condition is given by

$$\left.\frac{\partial E}{\partial \delta}\right|_{\delta=\overline{\delta}} = 0. \tag{A.8}$$

Applying Itô's lemma to the equity function one obtains the equity process dynamics, whose volatility is given by

$$\sigma_{E_t} = \frac{\partial E}{\partial \delta_t}\frac{\delta_t}{E_t}\sigma. \tag{A.9}$$

Figure A.1 compares model implied and empirical volatility.



(A) Model implied volatility.

(B) Empirical volatility.

FIGURE A.1: Equity volatility.

A CDS is a contract by which its seller agrees to compensate the buyer in case of a credit event. In return, as long as the underlying entity does not default, the CDS buyer makes a series of payments to the seller, the CDS spread. This is the coupon value that turns both legs of the contract equal. Mathematically,

$$E^{\mathbb{Q}}\left[cds\int_{0}^{t^{cds}} e^{-rs}1_{\{\tau>s\}}ds\,|\mathcal{F}_0\right] = E^{\mathbb{Q}}\left[e^{-r\tau}1_{\tau<t^{cds}}|\mathcal{F}_0\right] - E^{\mathbb{Q}}\left[e^{-r\tau}Rec_{\tau}\right], \tag{A.10}$$

where t^{cds} is the CDS maturity and $E^{\mathbb{Q}}\left[e^{-r\tau}Rec_{\tau}\right]$ stands for the discounted expected recovery rate. The latter is given by

$$E^{\mathbb{Q}}\left[e^{-r\tau}Rec_{\tau}\right] = \begin{cases} 0, \beta\overline{v}_0 \leq X \\ \left(\frac{\beta\overline{v}_0 - X}{L^*}\right)E^{\mathbb{Q}}\left[e^{-r\tau}1_{\tau<t^{cds}}|\mathcal{F}_0\right], X < \beta\overline{v}_0 \leq X + L^*, \\ E^{\mathbb{Q}}\left[e^{-r\tau}1_{\tau<t^{cds}}|\mathcal{F}_0\right], \beta\overline{v}_0 > X + L^* \end{cases} \tag{A.11}$$

where $L^{\star}$ is the nominal value of the debt class insured, X is the amount of liabilities senior to the debt class insured, which is assumed to grow at the same rate as L, and $E^{\mathbb{Q}}\left[e^{-r\tau}\right]$ is the value of a claim that pays unity whenever the firm is liquidated.

www.bportugal.pt