The London School of Economics and Political Science

Essays in Financial Intermediation and Banking

Elizabeth Ellen Foote

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, June 2011

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (see paragraph below).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Single/Joint Work Statement

Chapter 1. : This paper is single-authored.

Chapter 2. : This paper is single-authored. However, it was begun during time spent at the Bank of England. As a result, it was motivated by discussions with and guidance from members of staff working in the area of payment systems, in particular Mark Adams. (However, the usual disclaimer applies that it does not necessarily reflect the views of the Bank of England.)

Chapter 3. : This paper is joint work with Wilko Bolt (De Nederlandsche Bank) and Heiko Schmiedel (European Central Bank). In the initial stages of the paper, I was working at the ECB as an intern (Summer 2009). I was familiar with Bolt and Schmiedel (2011) and aware of the comments about the different business models made by Rochet (see discussion in Introduction). I encouraged Bolt and Schmiedel to consider this new angle in a different paper, building on the basic framework already in place. During my internship we had multiple discussions about the appropriate set-up, while I was the principal one working on the specifics of the algebraic framework, extending the model of Bolt and Schmiedel (2011). I had written up this version of the model by the end of the internship. Since then, Bolt and I have worked closely on all aspects of the model, from analysing the algebraic results to writing and elaborating the rest of the paper. Since our paper follows closely some of the model set-up of Bolt and Schmiedel (2011), the relevant description in our work follows that paper closely. Bolt worked on a Mathematica file to check the algebraic calculations, and to do numerical analysis. I have since extended this to consider the welfare analysis. Throughout this we have been in discussion with Schmiedel on interpreting the results. (The usual disclaimer applies that it does not necessarily reflect the views of the ECB or De Nederlandsche Bank.)

Abstract

Banks' role as intermediaries between short term investors and long term borrowers has dominated the literature. Whilst this is an important feature, there are many other characteristics of banks. Each chapter in this PhD explores a different aspect of banking, from other forms of lending to banks' role in payment services. The first, and principal, chapter considers credit lines: 'commitments' to lend if required. These remain off the bank's balance sheet until drawn upon. As off-balance sheet items, unused commitments face low capital charges under existing capital regulation. I explore how this regulatory feature incentivises banks to build up exposure to these lines. This may lead to a suboptimal allocation of credit, ex post, following a market shock, as high drawdowns cause the balance sheet to balloon and the capital requirement to bind. In the second chapter, I consider banks as agents in large-value payment systems. In choosing the optimal time to settle a payment, banks trade off delay costs against the risk of having insufficient liquidity to make future payments. With banks participating in multiple systems, I show how default in one system may spill over into another, through the strategic behaviour of multi-system participants. I explore how this risk varies with the degree of information asymmetry between agents in different systems. The third chapter focuses on retail banking. In joint work, we examine how the provision of consumer credit, either through current account overdrafts, or through credit card credit lines, affects the way in which debit and credit card networks compete. We find that, even when debit and credit cards compete, there are elements of complementarity between them. Banks providing debit cards and current accounts benefit when the consumer delays withdrawal of funds from her current account by using a credit card. This leads to surprisingly high debit merchant fees.

Acknowledgements

I thank my supervisor Amil Dasgupta for numerous discussions and advice throughout my PhD. I am grateful for comments on Chapter 1 from Ulf Axelson, Mike Burkart, Maria Cecilia Bustamente, Mikhail Chernov, Alex Edmans, Sujit Kapadia, Philippe Mueller, Francesco Nava, seminar participants at LSE, the Bank of England, Federal Reserve Board of Governors, Federal Reserve Bank of Boston, Bank of Canada, Norges Bank, Bank of Finland, University of Maryland, Stockholm School of Economics and Warwick Business School. I have benefited from comments on Chapter 2 from Matthew Willison, Peter Sinclair, Mark Adams, Anne Wetherilt, Ben Norman, Jochen Schanz, James Chapman, and seminar participants at Bank of England, LSE, Norges Bank and participants at the Federal Reserve Board Conference 'Economics of Payments V' 2011. I thank my Chapter 3 co-authors, Wilko Bolt and Heiko Schmiedel, and also Stu Weiner, our discussant at the ECB-OENB Conference 'The Future of Retail Payments'. I am indebted to Charlie Kahn for introducing me to the economics of payments and for his guidance in this area. Finally, to Andrew Mell, for endless hours of discussion about my work, economics in general, for proof-reading and for unstinting encouragement, thank you.

I received funding from ESRC for two years of my PhD at LSE.

Contents

Thesis Introduction

What is a bank? That question has been answered and explored frequently in the academic literature, following Diamond and Dybvig (1983). At its simplest, banks intermediate funds from short term investors to long term borrowers. While this is true, it fails to fully capture the many different roles played, and influences exerted, by banks. For example, banks do not just offer term loans, they also offer promises, or 'commitments' to lend if required. Even this slight variation from the classic model has significant implications for thinking about the services banks provide, the way they compete and also the impact of capital regulation on banks' optimisation.

Nor does the simple answer given above say anything about how banks perform the unglamorous role of oiling the system: making payments and settling trades between financial agents. Although these payment systems receive little attention in the literature or media, the ability to actually make a payment is fundamental to any financial system. Indeed, it has been a central role played by banks for centuries, at the very least rivaling the importance of their role as financial intermediaries (see Kohn (1999)). As discussed by Kahn and Roberds (2009), this is related, but not identical, to discussions in the monetary literature surrounding the transactions velocity of money and equally the role of inside money. The focus in the payments literature is rather on the optimal design for payment systems to facilitate settlement.

Discussion of payment systems falls generally into two categories: large-value payment systems and retail payment systems. As the name suggests, the former refers to systems such as CHAPS in the UK and Fedwire in the US, in which banks make high value, wholesale payments. In most cases, these are made across the books of the central bank using accounts held by clearing banks at the central bank. However, an increasing number of systems are run by private agents, with payments settling across the books of that agent (see Chapter 2 for discussion). Banks may be making payments to settle their own trades, or the payments may be made on behalf of customers.

In some cases, banks will make payments on behalf of other banks who may not be connected to the central bank payment system. Whilst these payment systems generally operate without a hiatus, the potential cost of default or disruption is severe in a system in which trillions of dollars are transferred each day. Given the nature of such networks, a problem at one bank can have consequences for the system as a whole. For instance, banks rely in part on recycling liquidity from incoming payments in order to fund outgoing payments. If one bank receives payments, but has an operational problem and is unable to make outgoing payments, this may result in a 'liquidity' sink in which some of the liquidity in the system gets 'stuck' at the stricken bank. This impacts the ability of other banks to make payments between themselves. As a result, concern in the literature and in policy circles focuses on the stability implications of potential disruptions.

The latter category, retail payment systems, refers to consumer payment services such as debit and credit cards. Here, the focus is on the optimal pricing of these services, taking into account that these are 'two-sided' markets. I discuss this in more detail below.

In this PhD, I have explored some of these roles played by banks, through a series of papers. Chapter 1 is the principal paper, in which I consider credit lines, banks' commitments to lend. I focus on the fact that these commitments, when unused, face low capital regulations as off-balance sheet items. I show how this incentivises banks to build up high exposures to these lines, with implications for the allocation of bank funds following a market shock. In Chapter 2, I examine banks' role in large-value payment systems. Departing from most of the existing literature, I consider interactions between two systems, and explore how information asymmetries between agents may lead problems in one system to spill over into the other. Chapter 3 is a joint paper, focusing on retail payments. We explore how the pricing and provision of consumer credit has implications for competition between debit and credit card networks. In the remainder of this introduction, I discuss each paper in turn. However, since Chapter 3 is joint work, I will devote more space here to describe the paper.

Credit Lines and Capital Adequacy. This paper focuses on why banks build up exposure to credit lines and how this can affect new credit supply following a market shock. Previous work has documented that firms draw down heavily on credit lines following market turmoil, such as the collapse of Lehman Brothers. Ivashina and Scharfstein (2010) argue that such high drawdowns represent negative liquidity shocks for a bank, constraining future credit supply from the bank.

However, I argue there is an additional channel through which credit line exposure affects future bank credit supply. This operates through the role of regulatory capital requirements. I show how capital regulation, and its treatment of credit line exposures, may promote a suboptimal allocation of bank credit following a market shock. Banks provide committed credit lines to corporates; yet, when undrawn, these lines remain off the banks' balance sheets. Under Basel I and Basel II, these exposures face zero or low capital charges. However, once the line is drawn down, it becomes a balance sheet item and faces a full capital charge. A drawdown shock is therefore like a shock to regulatory capital requirements. If the bank has insufficient capital to support large drawdowns following market turmoil, firms without credit lines may be rationed by the bank. This can result in a suboptimal allocation of credit; i.e., where the marginal return on projects funded via drawdowns is lower than the marginal return on loans to non-credit line holders.

Why, ex ante, would a bank set its optimal exposure so high that this could cause a suboptimal allocation of credit ex post? I argue banks have an incentive to offer credit lines if such commitments play a signalling role in reducing information asymmetry between the firm and market lenders. If a market lender cannot directly observe the quality of a firm's collateral, it may be unwilling to lend, or only willing to lend at high interest rates. If a bank commits to lend to the firm in the future, it can provide a credible signal of firm quality to the market. Market lenders will more readily lend to firms in the presence of this signal. Indeed, Mosebach (1999) and Loukoianova et al. (2006) discuss evidence supporting this signalling role of credit lines.

A credit line, when undrawn, generates high surplus from signalling, some of which the bank can extract through fees. Yet a credit line does not generate costs; this is because the bank does not face a full capital charge on credit lines when they remain undrawn. Crucially, therefore, a bank has an incentive to offer a large amount of credit lines, knowing that, with high probability, the firms will not draw the funds.

Nevertheless, with low probability, an adverse shock causes markets to dry up; firms will then draw upon their lines. In this case, the bank will be constrained by its available regulatory capital and will be forced to ration credit to non credit line holders.

I present suggestive empirical evidence consistent with this mechanism. I also draw implications for both monetary policy and capital regulations. I show that it is the interest rate in the good state, rather than the bad, which can worsen misallocation. If agents anticipate that interest rates will be low in good times, banks will increase their exposure ex ante in order to benefit from the surplus generated by undrawn lines. Furthermore, a change in capital regulations can reduce ex post misallocation. If a higher capital charge is imposed on undrawn lines, banks may reduce exposure, leading to a reduction in misallocation following the market shock. Intuitively, a charge on undrawn lines requires the bank to pay the capital charge in good times as well as bad. This reduces the surplus the bank can obtain from offering credit lines. In the extreme case, if the full capital charge were applied to undrawn lines, the bank's optimal exposure would be set to eliminate credit misallocation in the bad state. Furthermore, I derive implications for cyclical capital requirements. If overall capital adequacy requirements are relaxed in the bad state, such that the bank has to find less capital for each new loan, this will result in less misallocation. This effect occurs, despite the fact that the bank will increase its ex ante exposure to credit lines.

As explained in the paper, this work relates to two areas of existing literature: an area which considers credit lines in the context of optimal contracting (Holmstrom and Tirole (1998) and Boot et al. (1987)) and an area which considers the liquidity risk to banks of issuing credit lines (Cornett et al. (2010), Gatev and Strahan (2006), Kashyap et al. (2002)). Neither strand considers the implications of capital regulation for credit lines. Kanatas (1987) makes a similar argument regarding the signalling role of credit lines, but does not relate this to capital regulations and the broader implications for banks' balance sheets. Cornett et al. (2010) and Ivashina and Scharfstein (2010) discuss the constraints on bank lending following market shocks, if they previously had high exposure to credit lines. However, their story is about liquidity risk as opposed to the incentives in place from capital regulation. In my story, therefore, I effectively link these two strands of the literature.

Information Asymmetries and Spillover Risk in Settlement Systems. In the second chapter, I consider large-value payment systems. I focus on Real-Time Gross Settlement Systems (RTGS), in which participants must have sufficient liquidity in order to make a payment.[1] I show how delay and default in one system may spill over into another system. I focus on the strategic interaction of players (banks), examining in particular how the strategic behaviour of a multi-system participant can provide the channel through which problems spill across systems. Whilst I follow existing literature in focusing on strategic behaviour, most papers simply focus on one system in isolation and do not consider the potential for cross-system disruption.

A key ingredient in my paper is the presence of asymmetric information between players in different systems. Specifically, I assume that banks who participate in just one system are less likely to have information about disruptions in another system, compared with banks who participate in both. For simplicity, I refer to the first system as the 'domestic' system and the second as the 'offshore' system, where there is a risk of disruption. I show that if multi-system participants can rely on receiving and recycling liquidity in the domestic system, they have an incentive to exploit the information asymmetry by continuing to make payments in the offshore system where delay costs are high. However, this risky behaviour increases the probability that the multi-system participant himself defaults, thus causing the disruption offshore to spill over into the domestic system.

I explore how the equilibrium changes as we allow for different degrees and types of information asymmetry. The best outcome, where spillover risk is eliminated, occurs when the domestic-only participants can observe both the state of the world offshore, as well as the strategy played offshore by the multi-system participant. However, this seems a lot to ask in practice of information flows between systems.

I also draw implications for Liquidity Saving Mechanisms. Such mechanisms can potentially reduce the liquidity-intensive nature of RTGS systems. RTGS systems place great demands on the liquidity holdings of participating banks. As a partial remedy, Liquidity Saving Mechanisms (LSMs) allow participants to make a payment

[1] This may be contrasted with Netting systems, in which obligations are calculated as net positions between participants at the end of the day. Due to concerns about counterparty risk in netting systems, RTGS has become the predominant system type.

contingent on receiving an incoming payment. Effectively, the bank orders a payment to go through the system, but the payment will only be released if the incoming payment is also made. This is cost-saving as it enables banks to economise on liquidity.

In my paper, I show how the introduction of an LSM in one system has a positive externality, benefiting the other system by reducing the risk of spillover. This begs the question of whether such externalities will be internalized by policymakers in considering whether to introduce an LSM in their own system.

The modelling framework of this paper draws on the liquidity management literature which followed Bech and Garrett (2003). As discussed in the paper, this literature has focused on the tradeoffs faced by banks between making a payment and therefore reducing their liquidity holding, or paying a delay cost. Most papers, however, have analysed the game within one system. In my paper I extend the framework to consider spillover risk between systems, and how this is affected by the presence of information asymmetries. In so doing, I contribute to the literature on Liquidity Saving Mechanisms (e.g. Martin and McAndrews (2008)), which hitherto has focused primarily on the benefits (or otherwise) of LSMS within a given system, rather than considering the interconnections with other systems.

Consumer Credit and Payment Cards. The third chapter combines both credit and payment issues, but explores these issues within the literature on two-sided markets and retail banking. In this joint paper, with Wilko Bolt and Heiko Schmiedel, we explore how consumer credit provision affects the nature of competition between debit and credit cards. In retail banking, consumer credit is frequently offered alongside a payment facility, so the two services are intertwined. Yet, the role of credit provision in competing payment card networks has received little attention from the literature.

Our principal contribution, therefore, is to show how different business models of card/credit combinations affect the nature of competition. Specifically, we show that there may be elements of both complementarity and competition between debit and credit cards.

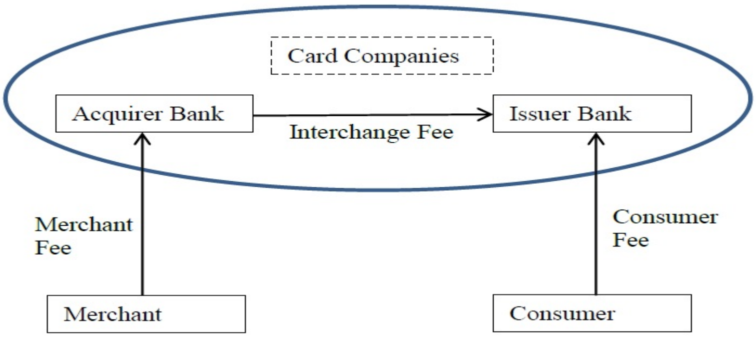
In this introduction, I will briefly overview the theory of two-sided markets, and then place our paper in the context of the existing literature.

Retail payment cards operate as two-sided markets. Card networks must attract both consumers and merchants, since the value of the card to one 'side' is in part dependent on how acceptable it is to the other. This complicates the nature of pricing, even in the case where the card network is a monopolist. The network must find the profit maximising price *structure*, the optimal combination of fees charged to each side of the market. In other words, the network must find the optimal combination of consumer fee and merchant fee. For instance, if the fee is too high for consumers, the card will be less attractive to them; if less consumers hold the card, this will make the card less attractive to merchants, and consequently this will lower the amount they are prepared to pay in merchant fee. See Rochet and Tirole (2002) and Rochet and Tirole (2006) for a good overview of two-sided markets. As they explain, this feature of two-sided markets comes from the fact that an agent from one side does not internalize the positive externality of his participation on potential participants from the other side.

Figure 1 below illustrates the nature of the payment card network. In theory, the network consists of five parties: an acquiring bank, an issuing bank, merchants and consumers, alongside a card company. The acquiring bank is the merchant's bank, to whom the merchant pays a fee, while the issuing bank is the consumer's bank, to whom the consumer pays a fee. The literature on payment cards has largely focused on characterising the private and socially optimal fees for a given network. This has been driven in part by policy concerns over the level of interchange fees (see recent debates surrounding the Durbin amendment). These are fees paid by an acquiring bank to an issuing bank; the purpose of an interchange fee is for one side to subsidise the other, thus enabling the network as a whole to maximise profits. Credit card interchange fees are typically around 1 to 2% of transaction value, while those for debit cards are typically around 0 to 1% of transaction value (Rochet and Tirole (2011)). In the US, policymakers have recently agreed to cap interchange fees for debit cards.

Since the literature focuses on the profit maximising fee structure for the network as a whole, many papers consider four-party networks, since the card company can be subsumed into either of the banks without loss of generality. Indeed, the card company

FIGURE 1. Retail Payment Network



is often an association of banks. Furthermore, a lot of papers simply consider three-party networks, in which the two banks are combined into one. This avoids the need to directly compute the interchange fee, without changing the optimality of the merchant and consumer fees. As explained in Bolt (2006), the interchange fee is directly, and positively, related to the merchant fee. It can be explicitly solved for in this context if we assume acquiring banks are perfectly competitive. In our paper, we follow the three-party network approach.

While the existing literature has explored different networks, varying aspects such as the degree of consumer and merchant heterogeneity, very few papers have considered competing networks (see Rochet (2007)). The main insights from the few papers which do introduce competition (e.g. Guthrie and Wright (2007) and Chakravorti and Rosen (2006)) is that, while competition lowers card fees, it does not lower them sufficiently to meet the socially optimal level.

None of these papers explicitly consider competition between debit and credit card models, except for Bolt and Schmiedel (2011). Equally a lot of the literature has ignored the fact that payment cards are often twinned with forms of credit. The welfare enhancing element of payment cards thus comes primarily from avoiding the cost of cash handling on both sides of the market.

The exceptions to this include Rochet and Tirole (2009), Chakravorti and To (2007), Bolt and Chakravorti (2008) and Bolt and Schmiedel (2011). (See discussion in the chapter for more details.) Yet, although these papers allow for credit in one particular network, they do not explore how different business models of credit via payment cards affect the nature of competition between different networks.

The framework for our analysis draws heavily on the model in Bolt and Schmiedel (2011). They consider competition between two networks, one of which they call a credit card network and the other a debit card network (building on a previous model in Bolt and Chakravorti (2008)). The key feature in Bolt and Schmiedel (2011) is that credit cards provide consumers with credit as well as a payment service, while debit cards just provide a payment service with no accompanying credit. As a result, credit cards allow payment in one extra state of the world.

One criticism of that paper is that it fails to capture the empirical reality of the two business models of consumer credit in Europe, a point made by Rochet as discussant at an ECB conference in 2009. In practice, credit is provided through both credit cards and debit cards. We decided to take this observation further and see what it means for competition between the networks. We started with the observation that there are typically three principal differences in the business models (albeit with exceptions): 1) the amount of credit provided via a credit card is larger; 2) the credit via the credit card is typically interest free for the first month, a 'grace' period; 3) the credit provided via the debit card comes from an overdraft associated with a current account.

The third characteristic means that the debit card should be seen as part of a broader package involving the consumer's current account. It also means that, although credit can be used via a debit card, there is no extra credit offered through a debit card, relative to cash. The consumer could simply access his overdraft by paying with cash.

To reflect this third characteristic, therefore, we introduce an overdraft facility which can be used both with cash and debit cards. To reflect the first characteristic above, the credit line in our model is larger than the overdraft facility, allowing the consumer to purchase in one extra state of the world, the state when initial income is low. Also, to reflect the second characteristic, we introduce explicit costs of credit via interest rates: we capture the fact that the consumer must pay interest in the first month on an overdraft, but not on credit received via a credit card during the first month. We find that the relative costs of credit in one business model have an effect on the equilibrium pricing of payment card services in the other.

We draw implications for the case of a credit-card only world, as well as for the case of competing networks. In the former case, credit card merchant fees depend on the expected costs of servicing the overdraft. Since the alternative to using a credit card is to use cash, which is associated with a given overdraft, the costs of servicing the overdraft have a negative effect on credit card merchant fees. The higher the expected costs of the overdraft, the more the consumer saves, at the margin, by using a credit line. This means the network can charge a higher fee to the consumer, and a lower fee to merchants.

We also find that debit merchant fees do not depend on default risk and funding costs when they operate on their own. This is because the only extra benefit from debit cards over cash comes from extra security. However, when debit cards compete with credit cards, debit merchant fees become dependent on default risk and funding costs since credit card fees do directly depend on these characteristics. In fact, we see that debit card merchant acceptance increases with default risk, as some merchants move from accepting credit cards to debit cards.

In the case of competing debit and credit card networks, we find our key result. In addition to a downward pressure on payment fees coming from competition, we find that there are actually elements of complementarity between the business models. If the consumer uses her credit card to make a payment she will maintain a higher than otherwise balance in her current account, whilst enjoying the initial interest free 'grace' period. At the margin, therefore, the debit card network has an incentive to raise payment card fees to discourage debit card acceptance for high values of funding cost.

CHAPTER 1

Credit Lines and Capital Adequacy

1. Introduction

1.1. Overview. This paper focuses on why banks build up exposure to credit lines and how this can affect new credit supply following a market shock. Traditional models of bank lending focus on standard, term loans, in which borrowers take the funds with certainty at the beginning of the contract. In reality, however, a significant proportion of bank lending operates through credit lines. In these contracts, banks commit to lend in advance, but funds are only 'drawn down' if required by the firm at a later date. The bank faces uncertainty about whether or not it will be called upon to provide funds. Credit lines can represent various degrees of commitment. Some are simply intentions to lend which do not bind the bank, whilst others are irrevocable. In this paper, we focus entirely on irrevocable commitments.

Recent evidence suggests credit lines play a crucial role in the distribution of new credit, particularly in times of market turmoil. At these times, firms that are unable to obtain market financing draw down significantly on their credit lines; Ivashina and Scharfstein (2010) document this effect following the Lehman collapse, whilst the Bank of England documents that corporate drawdowns increased in the last quarter of 2007 (Bank of England (2008a)). Indeed, C and I loans on aggregate balance sheets of US banks actually rose by around 100 billion dollars from September to mid-October 2008 (Chari et al. (2008)), of which approximately 25% can be explained by credit line drawdowns, according to news report research by Ivashina and Scharfstein (2010).[1]

Ivashina and Scharfstein (2010) follow earlier papers to argue that credit line drawdowns are like a liquidity shock for the bank.[2] Banks with higher drawdowns face a greater drain on liquidity, thus constraining their ability to supply newly originated

[1]Although some credit lines were revoked by banks during the crisis, evidence from the Bank of England, Ivashina and Scharfstein and others indicates a substantial amount of lines were indeed irrevokable.

[2]Earlier papers highlighting this issue include Gatev et al. (2009), Gatev and Strahan (2006) and Kashyap et al. (2002).

credit. They provide evidence that banks with higher credit line exposure (more undrawn lines) had lower growth of new credit during the crisis.

However, I argue that there is an additional channel through which credit line exposure affects future bank credit origination. This operates through the role of regulatory capital requirements. Undrawn credit lines are 'off balance sheet' and so do not face a full capital charge commensurate with equivalent loans that are 'on balance sheet'. However, once the line is drawn down, it becomes a balance sheet item and faces a full capital charge. A drawdown shock is therefore like a shock to regulatory capital requirements.

Indeed, under Basel I, irrevocable commitments, of maturity less than 1 year face no capital charge. During the crisis, most US banks were still operating under Basel 1. Even under Basel II, these short term commitments face very low charges: a credit conversion factor (CCF) of 20% and principal risk factor (PRF) of 100%. Whilst this increases the capital charge, it still means the off balance sheet item faces a lower capital charge than if it were a drawn loan. For longer term irrevocable commitments, the CCF is 50% and the PRF is 100% under both Basel I and Basel II.[3]

The implications of credit line drawdowns for the regulatory position of a bank are indeed significant, as the IMF points out in 2008:

> "Using the standards of Basel I, Fitch Ratings (2007) estimated that, under a worst-case scenario, if liquidity lines were to be fully drawn down, declines in the Tier 1 capital ratio of European banks would peak at 50 percent and for U.S. banks at almost 29 percent" (International Monetary Fund (2008) p.77).

A bank which suddenly faces a large amount of drawdowns may find it approaches its regulatory capital constraint. If it is unable to readily raise new capital, it will have to cut back on the amount of newly originated loans it plans to make. This will result in an ex post suboptimal allocation of credit; i.e., the marginal return on projects funded via drawdowns will be lower than the marginal return on newly originated loans.

[3]Although we do not focus on revocable commitments, it is worth noting both the CCF and PRF for these under Basel II remain 0%. (Chateau (2007) and International Monetary Fund (2008)). There is no capital charge for these commitments, despite banks frequently honouring commitments in adverse conditions to avoid losing reputational capital (Bhalla (2008) p.407).

Why, ex ante, would a bank set its optimal exposure so high that this may cause a suboptimal allocation of credit ex post, if a market shock occurs? Why doesn't it just offer standard bank loans, thus avoiding any ex ante commitment and ex post misallocation? A key contribution of my paper is to answer this question. Suppose there is information asymmetry between market lenders and the firm, such that the former cannot observe firm quality. Suppose also that the bank has a superior monitoring capability, allowing it to observe the quality of a firm (see e.g. Diamond and Rajan (2001)). If the bank commits to lend to the firm in the future, it can provide a signal of firm quality to the market. Market lenders will more readily lend to firms, in the presence of this signal.

In the absence of capital regulations, this signal would not increase surplus; without a credit line, the firm could still get financing from the bank in the form of standard bank loan as long as the bank can observe firm quality. However, in the presence of capital regulations, the bank is constrained in the total amount of credit it can offer even in the good state when there is no market shock. By providing a signal to the market, the bank can ensure that more firms get financing in the good state than the number it would optimally wish to support on its own balance sheet.

This has an important implication: the bank would prefer to commit itself via credit lines than to simply provide loans when firms are rationed by the market. It has an incentive to do this because, when they remain undrawn, credit lines increase surplus via signalling but do not generate costs: this is because the bank does not face a full capital charge on credit lines when they remain undrawn. The bank can extract some of this surplus from the firm ex ante via an upfront fixed fee. This offsets the expected cost of a credit line in the bad state, which is the cost to the bank of having to ration credit to other types of borrower, who do not hold credit lines.[4].

This motivation of credit lines is consistent with empirical evidence. Indeed, there is significant evidence of this signalling effect (see Mosebach (1999), Loukoianova et al. (2006)). As argued by Loukoianova et al. (2006), 'opening a credit line with a highly reputable bank usually sends a positive signal to other financial market participants'. Mosebach (1999) finds evidence of a positive market reaction to a firm's stock on

[4]In practice, a high proportion of bank revenue from credit lines comes from fees which are paid on the 'undrawn' portion of the credit line (see Sufi (2009) and Loukoianova et al. (2006))

the news that a credit line has been granted.[5] The credit line is a credible signal because the bank commits to lend to the firm, regardless of aggregate conditions, but contingent on firm quality; as part of the contract, the bank is able to observe firm quality. In practice, credit line contracts require the firm to file reports to the bank on a regular basis, even when the line is undrawn, and are contingent on the firm satisfying certain characteristics, such as solvency ratios.

However, not all firms can obtain credit lines. Typically, firms who hold credit lines must have at least a certain credit-rating level and visibility in the market (see Sufi (2009) and Loukoianova et al. (2006)). Credit lines play a key role as bargaining chips for banks in competing for such firms' custom. By contrast, for those firms which are dependent on loans from a specific bank, banks do not need to compete. As a result, the bank has no need to provide contractual commitments to this latter type of firm. It is these firms who will be rationed when drawdowns are high.

By exploring this mechanism in a simple model, I highlight key policy implications. I show how changes in capital regulation may affect the degree of misallocation. If undrawn lines face a higher capital charge, relative to drawn lines, the bank is forced to pay the cost of credit lines even when they are not drawn. This can lead the bank to decrease its exposure, and therefore decrease misallocation in the bad state.

Furthermore, I derive implications for countercyclical capital requirements. If overall capital adequacy requirements are relaxed when aggregate credit is scarce (the bad state), this will reduce misallocation ex post. This is despite the fact that the bank will anticipate this relaxation and will increase its ex ante exposure to credit lines.

I also show that the bank's funding cost in the good state has an effect on the degree of misallocation in the bad state. This is because it is the good state in which the credit lines generate surplus, but no cost to the bank. A lower funding cost in the good state increases the surplus in that state, and increases the bank's incentive to build up exposure. This has implications for monetary policy, in so much as monetary authorities can affect lenders' funding costs.

In the rest of the introduction, I present a summary of the model, followed by an explanation of how policy can affect this ex post misallocation. I then briefly outline

[5]Mosebach (1999) also finds evidence that a bank's stock responds positively to such news.

my empirical work before concluding with a discussion of how this paper relates to existing research on credit lines.

1.2. Summary of Model, Results and Empirical Motivation. The model operates over three periods, 0, 1 and 2. The agents in the model are firms, market lenders and a bank.[6] At period 1, each firm needs to borrow a fixed amount for a given investment project, which will succeed or fail at period 2. I consider two types of firm; type 1 and type 2, where the latter are bank-dependent. Within each type, there is a continuum of qualities; firms vary in the amount of residual assets that would be left if the project failed at period 2.

For type 1 firms, there are two frictions in obtaining financing from the market at period 1. The first is asymmetric information; the market cannot observe the firm's quality (its potential residual assets) and is therefore unwilling to lend. (The firm's type, however, is public information.) The second comes from the lender's own ability to extract value from a given amount of residual assets. At date 1, there is an aggregate shock and the state is revealed as either good or bad; the lender can either extract a high or a low value from a given firm's residual assets.[7] This second friction means the market would be unwilling to lend in the bad state, irrespective of whether it had full information about a firm's quality.

At period 0, the bank chooses the number of firms to which it will offer credit lines. The bank will effectively maximise social surplus ex ante at period 0. The credit line increases social surplus since it enables the firm to borrow at period 1, in both states, when otherwise it would have been unable.[8] However, part of the surplus comes from a signalling role of credit lines. By committing to lend to the firm if necessary, the bank can learn the firm's true quality and so provide a signal of firm quality to the market. This overcomes the asymmetric information mentioned above, enabling the firm to obtain market financing in the good state. As a result, the bank will offer

[6]Since the model focuses on a situation where banks behave perfectly competitively in offering credit lines, I analyse the optimal problem of just one bank in the model. I also abstract from any inter-linkages between banks, since this is not the focus of the paper. Such systemic effects would however be interesting to consider in further work.

[7]The amount of residual assets varies across firms, but the amount of value that can be extracted from a given set of residual assets varies across the state.

[8]The firm requires funds at period 1 for the start of a new project. This could alternatively be modeled as an interim liquidity shock for an existing project, as in Holmstrom and Tirole (1998) (albeit with a different motivation for the credit line).

more credit lines than its balance sheet could support in the good state. In so doing, it increases social surplus at period 0.

In the bad state, however, all firms with credit lines will draw on them; the high drawdowns cause the bank's capital adequacy constraint to bind. This causes a misallocation of credit as the bank is constrained in its ability to lend to firms without credit lines (i.e. type 2 firms). In the bad state, the marginal return to type 2 loans is greater than the marginal return to loans via drawdowns. This will mean that the amount of drawdowns, relative to foregone loans, is suboptimally high in the bad state.[9]

Ex post misallocation in the bad state is not surprising when we consider that, ex ante, the bank maximises a weighted average of surplus in the good state and net cost in the bad state. Misallocation in the bad state is given by the wedge between the marginal benefit of drawdowns in the bad state and the marginal benefit of foregone loans in the bad state. Ex ante, the bank only sets a wedge between these because there is positive expected surplus (from signalling) from credit lines in the good state. Ex ante the bank solves for the constrained first best allocation. Effectively, it maximises surplus, *taking capital regulations and funding costs as given.*

A key implication of the model, therefore, is that a change in policy can change the degree of ex post misallocation in the bad state. It can do this by changing the ex ante optimal allocation chosen by the bank. I find that a significant redistribution of capital charges from drawn to undrawn lines would reduce misallocation in the bad state. This is equivalent to imposing a high Credit Conversion Factor on undrawn lines (see footnote in the introduction section). If the bank is forced to pay a high cost on undrawn lines, this will reduce the expected surplus in the good state; the bank will have an incentive to reduce its exposure ex ante. This supports the Basel Committee's suggestion to impose higher CCFs on off balance sheet items, such as credit lines (see Basel Committee on Banking Supervision (2009)). However, this result should not, however, be taken as a direct policy prescription, since I do not take into account the positive social benefits of credit lines in the good state.

[9]Of course, there will also be suboptimal allocation in the good state, just the other way around. However, in this paper I focus on the credit misallocation in the bad state. It seems reasonable to suppose credit misallocation and rationing in a downturn may have more severe immediate macroeconomic consequences. I leave the long term consequences of credit misallocation in the good state as a topic for further research.

Furthermore, I show that a loosening of capital requirements in the bad state will reduce misallocation in that state. This speaks to recent discussions about time varying capital regulations. Ex post, banks will be less constrained by capital requirements for a given exposure to credit lines. Although banks anticipate this loosening of policy, and thereby increase their exposure ex ante, this is not sufficient to outweigh the beneficial effect.[10]

I also explore the effect of the lenders' anticipated funding costs on the degree of misallocation. A lower interest rate on borrowed funds for lenders in the good state actually *increases* the degree of ex post misallocation in the bad state. A lower cost of funding in the good state (for all lenders) means there will be a a higher surplus per loan in the good state. This increases the surplus from credit lines, without increasing the cost, since the cost is only experienced in the bad state when credit lines are drawn down. This result resonates with arguments claiming low interest rates in the pre 2007 period increased vulnerability to a crisis (Bank For International Settlements (2009) Part III). [11] Along the same lines, if agents place low weight on the likelihood of market turmoil, banks will build up exposure and thus there will be high misallocation in the bad state.

In the final section of the paper, I explore whether there is any suggestive evidence in the data that is consistent with the model. This is a first pass at exploring the empirical interaction between capital and credit line exposure.

Previous work has just looked at the liquidity aspect of credit line exposure (Ivashina and Scharfstein (2010)), showing that the level of undrawn credit lines has a negative and significant effect on newly originated credit following a drawdown shock. However, I show preliminary evidence that the extra capital charge faced by a bank, should credit lines be drawn, is significant, over and above standard controls for the regulatory capital buffer. Following Ivashina and Scharfstein (2010), I use the collapse of Lehman as a key shock affecting market credit supply. Although we do not observe drawdowns

[10]I do not consider broader questions of optimal capital regulation in this framework. This is because I simply treat capital regulations as an exogenous constraint and do not consider the beneficial effects of capital regulation. It would be interesting to consider this in further work.

[11]By contrast a lower interest rate in the bad state has a symmetric effect on the marginal return to drawdowns in the bad state, and the marginal return to other loans in the same state. A lower interest rate in the bad state therefore has an offsetting effect on the benefit and cost of credit lines.

in the dataset, there is substantial evidence that this shock led to a reduction in available market financing, and thus affected credit line usage. As documented by Ivashina and Scharfstein (2010) and others (e.g. Bank of England (2008b)), this shock to the market led corporates to increase their drawdowns on existing credit lines.

1.3. Related Literature. This paper is related to two strands of the literature on credit lines. The first explores the motivation behind credit lines, focusing largely on the contract between the firm and the bank. Holmstrom and Tirole (1997, 1998) and Boot et al. (1987) use agency problems to motivate why there may be demand for a credit line. Sufi (2009) and Yun (2009) empirically examine the role played by credit lines in firm financing. In a similar spirit to my paper, Kanatas (1987) argues that, if firms' credit risk is unobservable, loan commitments (i.e. credit lines) can be purchased by firms to reveal their quality, and thus can be used as a signal in the sale of commercial paper. However, the author largely focuses on optimal contracting in the presence of adverse selection, and does not examine the wider impact of drawdowns on the bank's balance sheet and capital adequacy.

On the whole, these papers focus almost exclusively on the firm's demand, and do not link the bank's supply of a credit line with other elements of a bank's lending portfolio, or capital requirements. Blavy (2005), in justifying loan commitments as a monitoring tool for banks, explores how in a multiperiod model such contracts may incentivise banks to continue with non-performing loans and therefore ration credit to more promising borrowers. However, he does not explore the impact of capital regulation. With the exception of Kanatas (1987), most papers do not model credit lines as an outside option to market borrowing. Explaining the coexistence of credit lines and market lending is important when considering macroeconomic questions of credit allocation.

The second strand in the literature explores the interplay of credit lines and banks' deposits. Kashyap et al. (2002) focus on the liquidity risk to which credit lines expose a bank. They argue there are synergies between deposit-taking and the provision of loan commitments, as long as withdrawals and drawdowns are not perfectly positively correlated. Gatev and Strahan (2006) extend this point, by arguing there is in fact a negative correlation between withdrawals and drawdowns. More recently, Ivashina and Scharfstein (2010) and Cornett et al. (2010) show how banks responded to such liquidity risk in the context of the recent crisis. This strand of the literature does not focus on the demand for the credit line or the role of capital requirements.

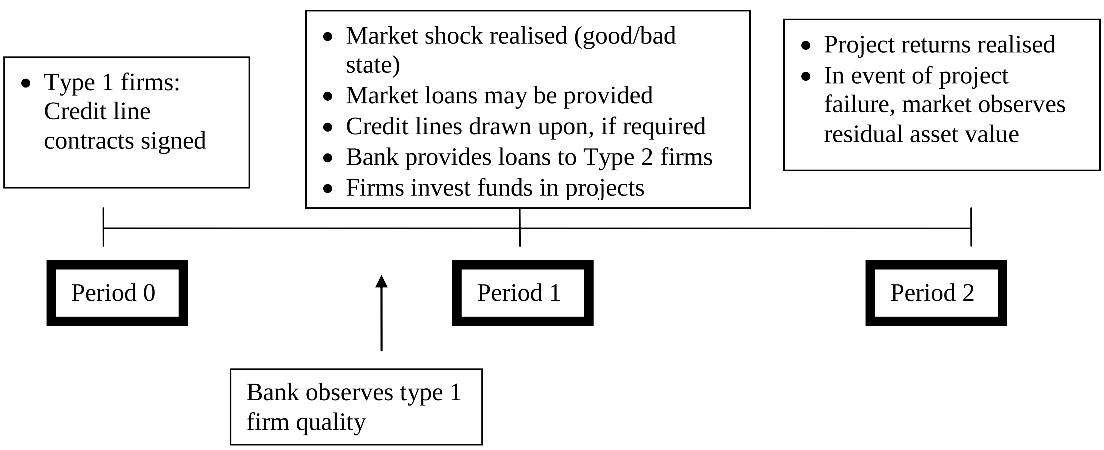
2. The Model

The model operates over three periods, periods 0, 1 and 2. The agents consist of banks, perfectly competitive market lenders, and also entrepreneurial firms. There are two types of firm, type 1 and type 2. Both types of firm will need to borrow in order to invest in a project at period 1. Loans may potentially be provided through 1) a direct market loan, 2) a draw down of a pre-arranged bank credit line or 3) a direct bank loan. However, not all options will be available to all firm types. In particular, type 2 firms will be bank-dependent so will only be able to access a direct bank loan.[12]

In providing the model set-up, we will focus primarily on type 1 firms. First, we will examine the project's parameters and the characteristics of the residual assets that remain in the event of project failure. Second, we will focus on the role played by credit lines. Third, we will consider the bank's loans to type 2 firms, and the constraint faced by the bank in the form of regulatory capital requirements.

To aid description of the model, here is the timeline.

Figure 1. Timeline



[12]The existence of bank-dependent firms is a standard assumption in the literature: it can be motivated by assuming project payoffs for type 2 firms are only observable by a bank with whom they have a close relationship.

2.1. Project parameters and residual assets. As a convenient reference whilst we describe the model, please find below a table of parameters. Each one will be described in due course.

TABLE 1. Parameter Reference

Parameter	Meaning
i_s	per unit cost of financing in state s
p	probability of project success
Y	project payoff if successful
γ_j	firm j's residual assets
β_s	liquid fraction of residual assets in state s
γ_c	lowest quality type 1 firm to be offered a credit line
γ_{bd}^s	lowest quality type 2 firm to be offered a bank loan in state s
L_s	loans to type 2 firms in state s
Q_s	loans via drawn credit lines in state s
ϵ	deadweight cost of bank financing relative to market financing
F	credit line upfront fixed fee
Y_c	face value payable on drawn credit line
ω	available capital
k	regulatory minimum capital ratio

At period 1, type 1 firms will require project financing. The gross per unit cost of financing is i_s, where s denotes the state. This is the direct cost of funds faced by any lender. We take i_s to be exogenous, influenced by monetary authorities (see later policy discussion), but known by agents with perfect foresight. Firms require a unit investment for a project which will succeed with probability p and give payoff Y at period 2. With probability $(1-p)$ the project will fail. In this case, some residual assets remain. We assume residual value varies across firms, γ_j, and is privately observed by the firm at the beginning of period 0. The net present value (NPV) of the project for firm j in state s is given by[13]

$$pY + (1-p)\gamma_j - i_s. \tag{2.1}$$

Residual assets play an important role in the model for two reasons. First, their value is unobservable to the public, i.e. to market lenders. Second, residual assets are

[13]Without loss of generality, I normalise the discount rate to 0. The project payoff is received and interest on the loan repaid in the same future period. As a result, the discount rate has no substantive effect on the analysis below.

partially illiquid, such that lenders cannot extract their full value. We now consider each of these characteristics in turn.

a) Residual asset value is publicly unobservable.

I assume the market cannot observe the residual asset value of any given firm at period 1.[14] This asymmetric information leads to market rationing. I suppose type 1 firms can be of 'high' or 'low' quality. If the firm is a low quality, its residual assets are worthless. If the firm is of high quality, it has positive residual assets. Residual asset value for high quality firm j will be denoted as γ_j. These high quality firms are uniformly distributed along a continuum in γ_j, over the interval $[\underline{\gamma}, \bar{\gamma}]$, where $\bar{\gamma} - \underline{\gamma} = 1$. The table below summarises this point:

Table 2. Type 1 firms

	Residual Assets
Low quality firm	$\gamma_j = 0$
High quality firm	$\gamma_j \sim U[\underline{\gamma}, \bar{\gamma}]$

Throughout the paper I focus on high quality firms. These high quality firms have projects with positive NPV since I assume that:

$$pY + (1-p)\underline{\gamma_j} > i_s.$$

However, the presence of low quality firms means the market's expectation of a high quality firm's residual assets will be significantly lower than the true value.[15]
I assume the proportion of low quality firms is sufficiently large such that the loan would have negative NPV. So, in the absence of any signalling mechanism, market lenders will not provide loans to any firm, even though high quality firms have positive NPV projects.

[14]For simplicity, I assume that the true residual asset value is observed at period 2.

[15]To formalise the market's expectation at period 1, suppose that proportion v of type 1 firms are of high quality, where $v < 1$. The market's expectation of residual assets for any given firm is therefore $v[\frac{\overline{\gamma}+\underline{\gamma}}{2}]$. For v sufficiently low, the loan will have negative net present value (NPV) for the market, even in the good state, and even if residual assets are fully pledgeable to outside investors. In other words,

$$pY + (1-p)v[\frac{\overline{\gamma}+\underline{\gamma}}{2}] < i_s \qquad \text{for } s = G, B.$$

b) Residual asset value is relatively more pledgeable in the good state, versus the bad state.

Of a firm's residual assets, some fraction will be liquid and thus pledgeable to outside lenders. The remaining fraction will be illiquid and thus unpledgeable. At period 1, there is an aggregate shock to the liquidity of those assets. This shock will produce either a good state or a bad state ($s = G, B$). In state s, a fraction β_s of residual assets will be sufficiently liquid such that they are pledgeable to outside lenders, in the event of project failure. The remaining fraction $(1 - \beta_s)$ will be illiquid and nonpledgeable. The good state represents a situation of greater liquidity and thus greater ease of credit: a higher fraction of residual assets are pledgeable ($\beta_G \gg \beta_B$). The bad state occurs with probability q and the good state with probability $(1 - q)$.

Since we assume $\beta_G \gg \beta_B$, this means credit rationing is significantly worse in the bad state. For simplicity, we assume β_B is so low that loans to all firms will be negative NPV. In other words, we assume:

$$pY + (1-p)\beta_B\bar{\gamma} - i_B < 0, \tag{2.2}$$

In the bad state, therefore, the market will not lend to any firm (even if there were no problem of observing the value of residual assets).

In the good state, we assume β_G is sufficiently high such that loans to all high quality firms will be positive NPV. A sufficient condition for this is

$$pY + (1-p)\beta_G\underline{\gamma} - i_G > 0. \tag{2.3}$$

To summarise, there are two key characteristics of residual assets. The first, the lack of observability, requires firms to find some means of signalling their type. The second, the variation in pledgeability, provides the shock: in the bad state, positive NPV firms will be unable to tap market financing, even without problems of observability.

2.2. Credit lines as a signalling device. Credit lines can be used as a means of overcoming the problem of observability. Indirectly, they also overcome the problem of market financing in the bad state. At period 0, in advance of investment requirements, a firm can sign a credit line with a bank, upon paying an upfront fixed fee F. This is a commitment to lend to the firm, at a given face value Y_c, should the firm require bank financing. As discussed in the introduction, the bank has a screening technology which enables it to observe the true value γ_j after it has signed the credit line, at the end of period 0.[16] I assume the bank can make the credit line contingent on the value of γ_j, and will revoke it if the firm turns out to have misrepresented its γ_j. As a result, if a firm has a credit line in place by period 1, the market can correctly infer the value of γ_j from the specifics of the contract; the presence of the credit line is a credible signal since the bank has committed to lend to the firm.[17]

Whilst agents can contract on the basis of observed payoffs and value, we assume they cannot contract on the state of the world, i.e. the state is non-verifiable (see Hart and Moore (1998)).[18] As a result, the credit line contract cannot be made contingent on the state. This is consistent with empirical observation. While credit line covenants tend to be contingent on firm specific characteristics, they are not contingent on aggregate macroeconomic/financial activity or states.[19]

We assume that the bank behaves perfectly competitively in offering credit lines.[20] In expectation at period 0, therefore, the bank will break even on each credit line. At this point, the bank will also choose its optimal exposure to credit lines; it will choose the lowest quality firm γ_c with which it will sign a credit line.

[16]In practice, banks do obtain significant information about the firm during periods when a credit line is undrawn. Typical covenants in credit line contracts require firms to submit quarterly or monthly financial statements to the bank, as well as preventing the firm from making changes to management without the lender's permission (see Financial Leadership Exchange (2008)). Banks are also able to observe small but frequent drawdowns for inventory purposes, Whilst distinct from the size of drawdowns motivating this model, they do enable the bank to obtain information about the firm's activities.

[17]Of course, the bank could deliberately fail to revoke a credit line where the firm has zero residual asset value, thereby misleading the market into lending; this might enable the bank to extract fees from that firm. However, as we will see, the bank will set the credit line contingent on an observed value γ_c where $\gamma_c > \underline{\gamma} > 0$. As long as γ is observed before the state of the world is revealed, the bank has no incentive to *fail to revoke* a credit line where $\gamma < \gamma_c$.

[18]This is a common assumption in the literature (see Houston and Venkataraman (1996)). The exact nature of the state may be observed with delay - at least, verifiable characteristics of the state may be observed only after credit lines are drawn down (see Jovanovic and Ueda (1997) and Meh et al. (2010)). Indeed, in the recent crisis, the exact extent of credit supply and credit rationing has only been observed with any accuracy with some delay (via surveys etc). Of course, since we allow for different costs of funding across the two states (i_s), this means we assume a court could not immediately observe and verify the state by i_s; indeed, the cost of funding faced by a lender is observed by himself, but not immediately by outsiders, in particular the courts.

[19]On a related note, we assume the bank is unable to hedge against credit line drawdowns.

[20]This does not mean the bank behaves perfectly competitively in all markets. In particular, as we will see, it has a monopoly over type 2 firms which are bank-dependent.

2.3. Bank's objectives. What does the bank's choice look like? To understand this, we must consider two things: a) what constraints does the bank face in offering lines and b) when will lines be drawn upon?

a) Type 2 loans and the capital adequacy requirement

In addition to providing credit lines to type 1 firms at period 0, the bank provides loans to type 2 firms at period 1. We assume type 2 firms are each dependent on loans from a specific bank. Cash flows or realised project returns of the firm are unobservable to the market, and only observed by the bank with whom the firm has a long-standing relationship.[21] Since they are bank-dependent, the bank does not need to provide them with pre-committed credit lines, in order to compete for their custom. For simplicity we model type 2 firm project payoffs as identical to that of type 1 projects, with type 2 firms also uniformly distributed on a continuum in γ, on the interval $[\underline{\gamma}, \overline{\gamma}]$. Whilst this is not necessary for the results, it helps with tractability.

Given the bank has monopoly power over type 2 firms, we assume for simplicity it extracts the whole surplus of each project. Unconstrained loans to type 2 firms are therefore given by

$$\bar{\gamma} - \underline{\gamma}.$$

The bank also faces a capital adequacy requirement (CAR). In particular, we assume the bank must hold capital w at least equal to fraction k of its risky assets. For simplicity, we take this constraint as given and do not endogenously model the regulators' decision.[22] If we denote the quantity of drawn credit lines as Q_s and the quantity of bank dependent loans as L_s, then the CAR is:

$$k[Q_s + L_s] \leq \omega.$$

We assume capital is fixed. This captures the idea that, at least over short run periods, a bank may face costs to increasing capital.

[21]This is particularly the case for small or medium sized enterprises, whose public presence is minimal.

[22]There is no bank default in this model. We are implicitly assuming the bank has a high cost of default. This cost must be sufficiently high such that the bank would rather constrain its lending if the CAR binds, than choose to have insufficient capital and risk regulatory sanctions and possible default. Note also we implicitly assume the need for capital regulation. In other words, we assume sufficient correlation between individual firm profitability so the firm shocks do not cancel out, leaving the bank with no risk. The specific nature of the correlation does not change our results for a given level of k.

We make two further assumptions about available capital. The first is that the CAR is non-trivial: the bank cannot support loans to all type 2 firms, and all high quality type 1 firms, without the constraint binding:[23] Recall that the total number of high quality firms type 1 firms is the same as that of type 2 firms, and given by $\bar{\gamma} - \underline{\gamma}$. So this assumption implies:

$$
\begin{aligned}
(\bar{\gamma} - \underline{\gamma}) + (\bar{\gamma} - \underline{\gamma}) &> \frac{\omega}{k}, \qquad (2.4)\\
2(\bar{\gamma} - \underline{\gamma}) &> \frac{\omega}{k}. \qquad (2.5)
\end{aligned}
$$

The second assumption is that the bank is unconstrained by its CAR if it only makes loans to type 2 firms. This implies:

$$
\bar{\gamma} - \underline{\gamma} < \frac{\omega}{k}. \qquad (2.6)
$$

We also assume that the pool of bank dependent firms is not affected by the β shock. Given these firms are bank dependent, it seems reasonable to suppose the bank can extract at least some of their illiquid residual assets (see Diamond and Rajan (2001)), and thus these loans are not so directly affected by the state. To simplify the model, therefore, we assume the bank can extract all residual assets from type 2 firms, regardless of the state. Whilst the strength of this assumption is not vital to the model, it is necessary that the profitability of bank dependent loans does not decrease by the same extent as market loans in the bad state. Otherwise, the increase in drawdowns of credit lines will be completely offset by a reduction in optimal lending to bank dependent firms, and the bank will not face a binding capital constraint in the bad state.

b)Credit Line drawdowns vs Market loans

When will firms draw upon their credit lines? It is clear that all firms holding a line will draw upon it in the bad state, since there is no possibility of market financing.

[23] At this point, we can see why it is important that the marginal benefit of lending to non-credit line holders does not decrease 1:1 with the increase in drawdowns. Indeed, if optimal unconstrained loans were to decrease in the bad state, relative to the good, by the same extent as the increase in drawdowns, the CAR would not bind in the bad state.

The key point however is that, in the good state, the presence of a credit line signals firm quality to the market. This ensures all credit line holders can obtain market financing in the good state. Indeed, since market lenders are perfectly competitive, the firm can always extract full project surplus by going to the market. Since there is no reason in the model why the bank would wish to set a lower face value than the market, firms will be at least indifferent between line drawdowns and market loans at period 1 in the good state.

This feature is crucial in incentivising the bank to offer more credit lines than it would optimally want to support fully drawn on its balance sheet. By signalling a type 1 firm's quality via the presence of the credit line, the bank anticipates that its CAR will not bind in that state and it will not have to forgo loans to type 2 firms. Given the bank receives a fixed fee from the firm regardless of whether the firm draws down, the bank nevertheless has an incentive to offer a large number of credit lines. So the bank can earn income from credit lines in the good state, but face no cost of foregone type 2 loans, as long as sufficient number of type 1 firms use market financing.

For the purposes of solving for the optimal supply, therefore, it is not important to specify the exact amount of firms that draw down in the good state. For ease of exposition, however, we assume there is an $\varepsilon > 0$ deadweight cost of bank financing, relative to market financing. This ensures that, ex ante, the terms of the credit line will be set such that no firm will draw down unless unable to obtain market financing; in other words, the optimal contract will minimise deadweight costs. This is consistent with empirical evidence, which suggests firms would rather seek market financing than bank financing.[24]

3. Solving for the optimal supply of credit lines

In order to solve for the optimal supply of credit lines, we first need to consider how the supply of lines impacts the supply of loans to type 2 firms. In the bad state, a large number of drawdowns would constrain the bank's available lending to type 2

[24] Nevertheless, given the argument above, this assumption does not affect the results; it simply ensures a clearer discussion of the model. In practice, the bank could simply set the face value of the credit line loan such that no firm prefers to draw down instead of accessing market financing. However, we do not solve for the face value: it does not affect optimal supply of lines and cannot be uniquely determined in this model setup.

firms if the CAR binds: these foregone type 2 loans represent a cost of offering credit lines for the bank.

3.0.1. *Supply of type 2 loans.* The supply of type 2 loans will be determined by the CAR. In the good state, as discussed above, there will be no CAR constraint and so type 2 loans will be offered unconstrained, equal to

$$L_G = \bar{\gamma} - \underline{\gamma}.$$

Whether or not the CAR binds in the bad state will depend on the minimum level of γ_j against which the bank commits to provide a credit line γ_c. Recall that all firms will draw down in the bad state, since they cannot borrow from the market.

It is helpful to consider the optimal value of γ_c in the absence of the CAR constraint. This will simply be given by $\underline{\gamma}$, since we assume all projects have positive NPV for firms $\gamma_j > \underline{\gamma}$. The bank can extract some of this surplus ex ante through the upfront fixed fee at period 0.

Given condition 2.4, which states the bank cannot support loans to all types 1 and 2, the CAR must therefore bind in the bad state:

$$Q_B(\gamma_c) + L(\gamma_c) = \frac{\omega}{k},$$

where

$$Q_B(\gamma_c) = \bar{\gamma} - \gamma_c,$$

such that constrained loans to bank dependent firms (as a function of γ_c) are

$$L(\gamma_c) = \bar{\gamma} - \gamma^B_{bd},$$

where the marginal bank dependent loan γ^B_{bd} is determined by the binding CAR constraint:[25]

[25]Note that I have set up the problem so the CAR does not bind in the good state. However, even if it did bind in the good state, it would bind less tightly than in the bad given $\beta_B < \beta_G$; i.e. it would bind less tightly as long as drawdowns were greater in the bad state compared with the good state. In this case, there would still be a misallocation result in the bad state because the marginal benefit of drawdowns in the bad state would still be less than the marginal benefit of bank dependent loans.

$$\begin{aligned}\gamma_{bd}^{B} &= \bar{\gamma}-\gamma_c+\bar{\gamma}-\frac{\omega}{k}\\ &= 2\bar{\gamma}-\gamma_c-\frac{\omega}{k}.\end{aligned}$$

3.1. Contract fees. Since type 1 loans in the bad state will be negative NPV for the bank, the bank will have to charge a fixed fee F in order to break even in expectation[26]. The optimal contract will also specify Y_c, the face value of the loan in the event the line is drawn.[27] Once again, this reflects empirical reality. Typically, a credit line contract will specify an interest rate to be paid on any drawdown, *plus* a commitment fee paid for the portion of the line that remains undrawn, e.g. in the good state, and/or an upfront fixed fee (see Loukoianova et al. (2006)).[28]

3.2. Optimal credit lines. The optimal credit line contract at period 0 will maximise expected firm surplus from credit lines, subject to the bank's participation constraint. In choosing γ_c, the bank faces a tradeoff. A lower γ_c (i.e. higher exposure) means that overall returns from credit lines are higher in the good state. However, a lower γ_c also means that lending to bank dependent firms is further away from the optimal level when the CAR binds in the bad state.

Recalling condition (2.1), individual firm j's objective function is given by:

[26]We can assume that the fixed fee is small relative to the size of the investment requirement at period 1. So the firm is able to fund the fixed fee out of liquid assets at period 0 but still needs to borrow to invest at period 1. Alternatively, we could assume the firm has no liquid assets at period 0, so must borrow to pay this fixed fee. It will then have to roll over this loan at period 1, affecting the total borrowing requirement in that period. As long as F is relatively small, either assumption is valid.

[27]For completeness, it is worth mentioning the following constraints which will apply to F and Y_c. Clearly, Y_c must be feasible, so $Y_c \leq Y$; the firm must have sufficient funds to pay the face value of debt in the event of project success. Regarding F; if the bank wants to ensure the firm is able to obtain market financing in the good state, it will not set F so high that rolling over any loan for F at period 1 causes the market to ration a firm. In other words, F will be set so as to avoid violating the following condition:

$$pY+(1-p)\beta_G\gamma_j-i_G(1+F)\geq 0.$$

In what follows below, we focus on cases in which these constraints do not bind.

[28]However, although we include Y_c and F in our discussion below, we do not need to solve for them, nor do they affect our results. In fact, as we will see below, we cannot determine each of them uniquely. Yet this is not a problem: as we will see, since Y_c and F are just transfers between the bank and the firm, they do not affect the optimal condition determining the amount of credit line exposure γ_c.

$$\pi_j = (1-q)[pY + (1-p)\gamma_j - i_G] + q[p(Y - Y_c) + (1-p)(1-\beta_B)\gamma_j] - F. \tag{3.1}$$

The first term reflects the full surplus net of funding cost obtained in the good state from the project, given the firm obtains financing from perfectly competitive market lenders. In the bad state, the firm holding a credit line will still be rationed by the market.[29] The firm will however be able to draw down its credit line. So the second term reflects the expected payoff from the project in the bad state, given it is financed by a credit line. If the project succeeds, the firm must pay the face value of the credit line drawdown, Y_c. Given the setup of the model the bank can only extract pledgeable assets in the event of project failure. Therefore, the firm will always obtain $(1-\beta_B)\gamma$ in the bad state, in the event of project failure. We can think of this as nonpledgeable income. In both states of the world, the firm must pay the credit line fixed fee F, the third and final term above. Note that in the absence of a credit line, the firm would have zero payoff in both states, because the firm would be unable to obtain financing; no one would observe its true γ_j.[30]

The bank will make zero profits in expectation on each firm's credit line: it's participation constraint for each firm will therefore bind. Given γ_c, each individual participation constraint will jointly determine Y_c and F. Since we are only interested in determining γ_c, we consider the sum of all these participation constraints. Likewise, we consider the sum of all firms' objective functions. As a result, Y_c and F will drop out of the optimal problem determining γ_c, since they are just transfers between the firm and the bank.[31] In what follows, we focus on these aggregated functions.

[29]This follows given condition (2.2).

[30]I assume that if the bank has not signed a credit line with the firm, it does not embark on a relationship with that firm and therefore cannot observe the firm's γ_j at period 1. This assumption is not critical. Suppose the bank could observe γ_j at period 1, without signing a credit line. Given banks are identical, they would all be able to observe the γ_j, so any loan at period 1 to the firm in the good state would be provided perfectly competitively. As a result, the bank would still want to lend first to type 2 firms in the good state, from whom it can extract the full surplus. Since the bank has no commitment with the type 1 firm, the CAR will not bind. As I show in the appendix, type 1 firms who *do* hold credit lines have no incentive to deviate and rely on a spot loan from the bank in the good state only.

[31]These cannot be pinned down uniquely since there is no incentive compatibility constraint in this model.

The aggregate objective function for the firms contemplating signing a credit line is therefore the sum of each firm's objective function (i.e. the sum of 3.1 for each j):

$$\begin{aligned}
&\{(1-q)\int_{\gamma_c}^{\bar{\gamma}}[pY+(1-p)\gamma-i_G]d\gamma\} \\
&+\{q\int_{\gamma_c}^{\bar{\gamma}}[p(Y-Y_c)+(1-p)(1-\beta_B)\gamma]d\gamma\}-F[\overline{\gamma}-\gamma_c].
\end{aligned} \tag{3.2}$$

In the good state, the firm receives funding as long as $\gamma > \gamma_c$; i.e. as long as he holds a credit line.

The bank's (aggregate) participation constraint is given by

$$\begin{aligned}
& F[\overline{\gamma}-\gamma_c]+q\int_{\gamma_c}^{\bar{\gamma}}[pY_c+(1-p)\,\beta_B\gamma-i_B-\epsilon]d\gamma \\
\geq\ & q\int_{\underline{\gamma}}^{\gamma_{bd}^B(\gamma_c)}(pY+(1-p)\,\gamma-i_B)\,d\gamma
\end{aligned} \tag{3.3}$$

Condition 3.3 is the bank's participation constraint. The first term on the left hand side is fee income from credit lines, while the second term is the revenue from credit lines in the bad state. The term on the right hand side is the expected foregone profits on type 2 loans due to credit line drawdowns. Given banks behave competitively in offering credit lines, the bank's participation constraint will bind in equilibrium. Since the foregone profits from type 2 loans are positive, it must be that the left hand side is positive. In other words, the bank must make *positive* expected profits on credit lines, equal to expected foregone profits on bank dependent loans.[32].

Substituting the bank's aggregate participation constraint into the aggregate objective function allows us to solve for the optimal γ_c

$$\begin{aligned}
\max_{\gamma_c}\quad & (1-q)\int_{\gamma_c}^{\bar{\gamma}}[pY+(1-p)\gamma-i_G]d\gamma+q\int_{\gamma_c}^{\bar{\gamma}}[pY+(1-p)\gamma-i_B-\varepsilon]d\gamma \\
& -\{q\int_{\underline{\gamma}}^{\gamma_{bd}^B(\gamma_c)}(pY+(1-p)\gamma-i_B)d\gamma
\end{aligned}$$

This shows we can characterise the optimal exposure to credit lines, γ_c, independently of the optimal values of Y_c and F. Indeed, the optimal exposure is found by maximising

[32]The fixed fee F is effectively written in future value terms, since all terms in the optimal problem are written in terms of value at period 2.

overall net surplus of credit lines. The function (3.4) is simply net surplus from credit lines.

It is important to note that the bank will strictly prefer to issue credit lines to *all* firms in the range $[\gamma_c, \bar{\gamma}]$ rather than simply learning firm quality during period 0 and then offering uncommitted bank loans at period 1. In other words, it would not choose to offer uncommitted bank loans for some range $(\gamma^*, \bar{\gamma}]$, and committed credit lines to the remainder $[\gamma_c, \gamma^*]$. I show this in the appendix. Intuitively, any firm in this range is prepared to pay a higher fixed fee than the cost the bank would face from having committed itself via a credit line. Given this, credit lines are superior to uncommitted bank loans: the bank will choose to commit itself at period 1 for all type 1 firms in the range $[\gamma_c, \bar{\gamma}]$.

4. Results

The first order condition for γ_c is given by

$$\begin{aligned} 0 &= -(1-q)[pY + (1-p)\gamma_c - i_G] - q[pY + (1-p)\gamma_c - i_B - \varepsilon] \\ &\quad + q(pY + (1-p)[2\bar{\gamma} - \gamma_c - \frac{\omega}{k}] - i_B. \end{aligned}$$

So

$$\gamma_c = \frac{[qi_B + (1-q)i_G] - pY + q\varepsilon + q[pY + (1-p)(2\bar{\gamma} - \frac{\omega}{k}) - i_B]}{(q+1)(1-p)}.$$

4.1. Credit Misallocation. What does a given level of exposure γ_c mean for the degree of credit misallocation in the bad state? Before we consider comparative statics, it is helpful to recall what we mean by misallocation. Suppose there was an optimal allocation of credit in the bad state; then the marginal benefit of drawn credit lines would equal the marginal cost of foregone bank dependent loans. Using M_A to denote misallocation in the bad state, we can write:[33]

$$\begin{aligned} M_A &\equiv MB^B(\text{bank dep. loans})\text{-}MB^B(\text{credit lines}) \geq 0 \\ &= [pY + (1-p)(2\bar{\gamma} - \gamma_c - \frac{\omega}{k}) - i_B] - [pY + (1-p)\gamma_c - i_B - \varepsilon], \end{aligned}$$

[33]Strictly speaking, M_A is the marginal cost of misallocation. However, the total cost of misallocation is strictly increasing in M_A, and M_A is easier to analyse, so we focus on the latter. If we denote total cost of misallocation as TM_A, then M_A=$2(1-p)^{1/2}(TM_A)^{1/2} + \varepsilon$.

where $MB^s(n)$ represents marginal benefit of n in state s.

We can also show that the degree of misallocation in the bad state depends on the surplus in the good state. Note that the first order condition for γ_c can be rearranged and written as

$$MB^B(\text{credit lines}) + \frac{(1-q)}{q} MB^G(\text{credit lines}) - MB^B(\text{bank dep. loans}){=}0.$$

We therefore have the following lemma:

LEMMA 1. *Misallocation in the bad state is a function of the marginal benefit of credit lines in the good state:*

$$\begin{aligned} M_A &\equiv MB^B(\textit{bank dep. loans})\text{-}MB^B(\textit{credit lines}) \\ &= \frac{(1-q)}{q} MB^G(\textit{credit lines}) > 0 \end{aligned}$$

Note that this is an equilibrium result. As long as there is positive surplus generated by the signal from credit lines in the good state, the optimal exposure to credit lines will be set so that there is misallocation of credit in the bad state. In other words, too few loans will go to bank dependent firms, and too many to credit line holders, relative to the first-best in the bad state. As we will see below, changes in certain parameters, such as q and i_G, can exacerbate this misallocation.

4.2. Change in q, the probability of the bad state.

LEMMA 2. *As the probability of the bad state decreases, banks have greater incentive to increase their exposure to credit lines (i.e. to reduce γ_c):*

$$\frac{\partial \gamma_c}{\partial q} = \frac{[pY + (1-p)(2\bar{\gamma} - \gamma_c - \frac{\omega}{k}) - i_B] - (i_G - i_B) + \varepsilon}{[(q+1)(1-p)]} > 0.$$

Note this condition is positive as long as $(i_G - i_B)$ is not too large relative to the profits made on bank dependent loans. This is reasonable; it simply means there is some de facto limit to the range of interest rates which can be set by the monetary authorities.[34]

[34]This would be influenced by other factors in the economy, external to the model.

The key question, however, is whether this increased exposure translates into greater credit misallocation in the bad state of the world? It turns out that it does.

Proposition 1. *As the probability of the bad state decreases, there is an increase in misallocation of credit in the bad state of the world:*

$$\begin{aligned}\frac{\partial M_A}{\partial q} &= \frac{MB^B(bank\ dep.\ loans)}{\partial q}\text{-}\frac{MB^B(credit\ lines)}{\partial q}\\ &= -(1-p)\frac{\partial \gamma_c}{\partial q}-(1-p)\frac{\partial \gamma_c}{\partial q}\\ &= -2(1-p)\frac{\partial \gamma_c}{\partial q}<0.\end{aligned}$$

As the probability of the bad state decreases, this will lead banks to build up greater exposure to credit lines. This makes intuitive sense: the bank will put low weight on the cost of credit lines when q is low. This is because the cost of drawdowns, in terms of foregone other lending, only occurs in the bad state.

4.3. Change in funding costs. In the model we have taken the cost of funding to be exogenous, yet allowed it to vary across the states. Whilst we will not impose any further structure on these rates, it seems reasonable to suppose they are a function of monetary policy in each state.[35] By exploring comparative statics for i_G and i_B, we can therefore make qualitative statements about the effect of monetary policy in the two states. Given the set up of the model, the interest rates in each state are fully anticipated at period 0. In what follows, we use the phrase interest rates and funding costs interchangeably; the 'interest rate' does not refer to any aspect of the credit line contract but rather refers to the lender's funding cost.

Perhaps surprisingly, the anticipated funding cost in the good state has an effect on credit misallocation in the bad state.

Lemma 3. *If agents anticipate a relatively low funding cost in the good state, this will lead banks to increase their exposure to credit lines.*

[35] Costs of funding in each state will also depend on liquidity risk, counterparty risk etc. We do not focus on these issues here; we only examine the *marginal* effect of policy across the two states.

To understand this, consider the effect of a lower i_G on surplus in the good state. Since i_G is the per unit cost of funds in the good state, a lower i_G means higher surplus for a given level of γ_c in the good state. As a result, the marginal benefit of credit lines increases, giving banks the incentive to increase exposure (i.e. decrease γ_c):

$$\frac{\partial \gamma_c}{\partial i_G} = \frac{1-q}{[(q+1)(1-p)]} > 0.$$

Again, we have to consider the effect of this on misallocation.

PROPOSITION 2. *If agents anticipate a relatively low funding cost in the good state, this will lead to greater misallocation of credit in the bad state:*

$$\begin{aligned}
\frac{\partial M_A}{\partial i_G} &= \frac{\partial MB^B(\textit{bank dep. loans})}{\partial i_G} - \frac{\partial MB^B(\textit{credit lines})}{\partial i_G} \\
&= -(1-p)\frac{\partial \gamma_c}{\partial i_G} - (1-p)\frac{\partial \gamma_c}{\partial i_G} \\
&= -2(1-p)\frac{\partial \gamma_c}{\partial i_G} < 0.
\end{aligned}$$

As this shows, a decrease in i_G will lead to an increase in M_A: there will be greater misallocation of credit. So a credible commitment by the authorities to keep interest rates low in the good state will actually lead banks to build up high exposure to credit lines, resulting in greater misallocation of credit in the bad state. Again this is because a decrease in the funding cost in the good state increases the surplus from credit lines in the good state. This drives a wedge between the marginal benefit of credit lines in the bad state and the marginal benefit of foregone loans in the bad state.

PROPOSITION 3. *The expected funding cost in the bad state has no effect on credit line exposure:*

$$\begin{aligned}
\frac{\partial \gamma_c}{\partial i_B} &= \frac{q}{(q+1)(1-p)} - \frac{q}{(q+1)(1-p)} \\
&= 0.
\end{aligned}$$

To see this, note that a higher i_B will decrease the marginal benefit of credit lines in the bad state, but will also decrease the marginal benefit of bank dependent loans by the same amount. So i_B has no effect on ex ante exposure or ex post misallocation.[36]

4.4. Effect of policy: increase in capital requirements. We now turn to consider the implications for capital regulation. This can be broken down into two aspects. First, what is the effect of changing the way in which credit lines are regulated? If banks had to put aside a larger amount of capital to support *undrawn* lines, would this decrease credit misallocation in the bad state? Second, what happens to misallocation if the total capital charge faced by the bank is reduced in the bad state? We consider these two questions below.[37]

1: Increase in ex ante capital charges on credit lines. What if undrawn credit lines faced a positive capital charge - such that drawdowns did not result in such a large shock to capital requirements? Consider $\alpha \in [0, 1]$: if fraction $\alpha \times k$ of undrawn lines must be met in capital at the point of signing the line (period 0), then a subsequent drawdown of this line only requires the bank to find an extra fraction, $(1 - \alpha) \times k$, in capital. In this context, α may be interpreted as the Credit Conversion Factor (CCF) (see footnote in the introduction). In practice, α ranges from 0 to 50% for credit lines under Basel I and Basel II.

To analyse a change in α, we need to consider how available capital at period 0 relates to that at period 1. It seems reasonable, given the static nature of the model, to continue with the assumption that capital is fixed for the duration of the game. I leave consideration of varying capital to further work.

Given ω is fixed throughout, a proportional capital charge at period 0, amounting to $\alpha k Q_B(\gamma_c)$, simply decreases the available capital at period 1. This means the CAR constraint in the bad state is the same as before. Although the bank only has to find

[36]If we had modeled the bad state slightly differently, such that only some, but not all, credit line holders were rationed by the market, then i_B would affect γ_c. This is because the degree of market rationing would affect the number of bank dependent loans that could be offered, thus affecting the CAR constraint. If we denote by γ_m^B the marginal firm facing market rationing in the bad state, then γ_m^B must be a function of i_B. Furthermore, the marginal bank dependent loan would be a function of γ_m^B, and thus of i_B. However, a decrease in i_B, although causing an increase in exposure, would actually *decrease* misallocation in the bad state. This is because it would decrease market rationing in the bad state, and so relax the CAR constraint at the margin: this would decrease the marginal cost of credit lines in the bad state.

[37]We are ignoring issues of risk-weighting. Since all the assets are risky in this model, we focus, for simplicity, on a simple fraction of these assets in defining capital charges.

$(1-\alpha)kQ_B(\gamma_c)$ capital at period 1, it has less available capital with which to meet this charge. So

$$k[(1-\alpha)Q_B(\gamma_c)+L(i_B,\gamma_c)] \leq \omega - \alpha k Q_B(\gamma_c),$$

reduces to

$$k[Q_B(\gamma_c)+L(i_B,\gamma_c)] \leq \omega.$$

For low α, the bank's exposure to credit lines remains unchanged, as does the extent of misallocation in the bad state. However, if α is sufficiently high, the CAR will begin to bind in the good state, as well as the bad. As a result, the bank will change its optimal exposure. To see this, consider the CAR in the good state:

$$kL^{unc}(i_G) \leq \omega - \alpha k Q_B(\gamma_c).$$

Although there are no drawdowns in this state, the available capital is reduced by the charge on *undrawn* lines. For α sufficiently large, this will bind. Indeed, if $\alpha = 1$, then this constraint will be identical to that in the bad state.

At a point where the CAR binds in the good state, the first order condition for γ_c will be given as follows:

$$\begin{aligned} 0 &= -(1-q)[pY+(1-p)\gamma_c - i_G] - q[pY+(1-p)\gamma_c - i_B - \varepsilon] \\ &\quad -q[pY+(1-p)\gamma_{bd}^B(\gamma_c) - i_B]\frac{\partial \gamma_{bd}^B(\gamma_c)}{\partial \gamma_c} \\ &\quad -(1-q)[pY+(1-p)\gamma_{bd}^G(\gamma_c,\alpha) - i_B]\frac{\partial \gamma_{bd}^G(\gamma_c,\alpha)}{\partial \gamma_c}, \end{aligned}$$

where $\gamma_{bd}^B(\gamma_c)$ is given as before and now $\gamma_{bd}^G(\gamma_c,\alpha)$ is determined by the binding CAR in the good state:

$$\gamma_{bd}^G(\gamma_c,\alpha) = \bar{\gamma} + (\bar{\gamma}-\gamma_c)\alpha - \frac{\omega}{k}.$$

Differentiating this function with respect to α, we obtain

$$\begin{aligned} \frac{\partial \gamma_c}{\partial \alpha} &= \frac{(1-q)\{pY+(1-p)[\bar{\gamma}+2(\bar{\gamma}-\gamma_c)\alpha - \frac{\omega}{k}] - i_B\}}{[(1-p)(1+q)+(1-q)(1-p)\alpha^2]} \\ &> 0. \end{aligned}$$

An increase in α will decrease exposure (i.e. increase γ_c). Therefore an increase in α will decrease misallocation through its effect on γ_c.

Proposition 4. *For sufficiently high capital charges on undrawn lines, such that the CAR binds in the good state as well as the bad, an increase in the proportion of capital charge to be paid on undrawn lines will lead to a reduction in misallocation in the bad state:*

$$\begin{aligned}\frac{\partial M_A}{\partial \alpha} &= \frac{MB^B(\textit{bank dep. loans})}{\partial \alpha} - \frac{MB^B(\textit{credit lines})}{\partial \alpha} \\ &= -2(1-p)\frac{\partial \gamma_c}{\partial \alpha} < 0.\end{aligned}$$

For given exposure γ_c, and overall capital requirement k, an increase in α will increase the proportion of capital charges paid on *undrawn* lines. For large α, the bank effectively has to pay the cost in the good state as well as the bad. As a result, the bank will reduce its exposure to credit lines, thus reducing misallocation in the bad state. This supports the proposal made by the BCBS, in December 2009, to consider using a 100% Credit Conversion Factor on off balance sheet items for calculating the new leverage ratio in Basel III (Basel Committee on Banking Supervision (2009)).

Proposition 5. *If the bank had to pay full capital charges on undrawn lines, there would be no misallocation of credit in the bad state.*

This is equivalent to the case where $\alpha = 1$. In this case, the bank's CAR will bind equally in both states, regardless of the actual amount of drawdowns. The bank will therefore set its optimal exposure such that the marginal benefit of credit line drawdowns in the bad state is equal to the marginal benefit of foregone loans in the bad state. Algebraically, this equates to[38]

$$\gamma_c = \gamma_{bd}(\gamma_c).$$

It should be noted however that this is not a direct policy prescription. In the analysis above, we have ignored any potential social benefits of credit lines. In reality, policy tools such as capital regulation should take this into account, and would likely lead to a CCF of less than 100%. We have simply highlighted the implications of capital regulation for misallocation in the bad state.

[38]Strictly speaking, this holds as ε tends to 0.

2: A decrease in overall capital charge in the bad state. Recent policy discussions have focused on counter-cyclical capital requirements. This policy would increase the capital requirements for banks in states where aggregate credit is plentiful, but allow banks to relax the constraint to some degree in times of credit scarcity. Whilst proper consideration of this policy requires a dynamic model, in which banks optimally adjusted their capital, we can, as a first pass, consider how such a policy might affect the mechanism in this paper.

If capital requirements are looser in states of aggregate credit scarcity, this would imply a decrease in overall capital requirements in the bad state, equivalent to a lowering of k in the bad state. Recall the CAR constraint:

$$k[Q_s + L(i_s)] \leq \omega.$$

Here, k is the fraction of assets against which capital must be held. As k decreases, the capital charge decreases. If agents expect capital requirements to be eased following a bad shock, this is equivalent to a lowering of k in the present model, in the bad state.[39]

On the one hand, an decrease in k decreases misallocation in the bad state for a given level of exposure (γ_c). This is because a decrease in k makes the CAR bind less tightly for a given level of credit line drawdowns. On the other hand, at period 0, the bank anticipates a lower k and therefore chooses higher optimal exposure to credit lines (higher γ_c) since

$$\frac{\partial \gamma_c}{\partial k} = [\frac{q}{(q+1)}]\frac{\omega}{k^2} > 0.$$

Proposition 6. *A reduction in capital requirements in the bad state will reduce misallocation in the bad state. This is despite the fact that the bank will increase its exposure in response to an expected loosening of capital requirements in the bad state:*

[39]Note that a decrease in k in the bad state, as modeled here, has the same effect as a decrease in k in both states. This is because the CAR never binds for the bank in the good state.

$$
\begin{aligned}
\frac{\partial M_A}{\partial k} &= \frac{\partial MB^B(bank\ dep.\ loans)}{\partial k} - \frac{\partial MB^B(credit\ lines)}{\partial k} \\
&= (1-p)[\frac{\omega}{k^2} - \frac{\partial \gamma_c}{\partial k}] - (1-p)\frac{\partial \gamma_c}{\partial k} \\
&= (1-p)[\frac{\omega}{k^2} - 2\frac{\partial \gamma_c}{\partial k}] \\
&= (1-p)[1 - 2\frac{q}{(q+1)}]\frac{\omega}{k^2} > 0.
\end{aligned}
$$

A decrease in k will lead banks to increase their exposure ex ante, since they will anticipate lower misallocation in the bad state. However, this increase in exposure will not be sufficient to outweigh the beneficial effects of relaxing the constraint ex post (by decreasing k). Indeed, the first effect always dominates, since $q < 1$. Intuitively, the bank's weight on the bad state, including any relative gain in the bad state, is always less than 1. Overall, there will still be lower misallocation in the bad state, if k is expected to be lower in the bad state.[40] This lends support to the arguments in favour of cyclically adjusted capital requirements.[41]

5. Empirical Methodology

In this theoretical model, I have highlighted the way in which low capital requirements on undrawn credit lines incentivise banks to build up high exposure to these lines. This has implications for the allocation of bank credit following a market shock, when firms draw heavily on their lines. Previous empirical research has examined the liquidity effect of credit line drawdowns (see Gatev et al. (2009), Gatev and Strahan (2006), Cornett et al. (2010), Ivashina and Scharfstein (2010)). These authors argue that sudden drawdowns represent a liquidity shock for the bank, constraining future credit origination. I argue that an extra effect is present in the ratio of credit lines to available regulatory capital. My theoretical model suggests that banks with greater risk weighted undrawn lines *relative* to available regulatory capital will face greater

[40] Since we are assuming the contract cannot be state contingent, we must also assume that capital regulations specific to the state are not observable or verifiable to the courts immediately at the beginning of period 1. This would be the case if the relaxation of the constraint in the bad state were only observed informally in the immediate aftermath of the shock, in effect with the authorities confirming the relaxation only after some delay.

[41] Suppose the bank were to adjust its ex ante capital holding (ω) in the light of lower capital requirements in the bad state: in this case, the results described above would hold as long as the reduction in ω did not completely offset the reduction in k, so as long as the ratio ($\frac{\omega}{k}$) were to increase.

constraints in issuing new credit following a shock. In particular, this effect would be pronounced following a severe shock to markets.

I therefore construct a cross sectional dataset of bank balance sheet data, using the collapse of Lehman Brothers as the shock. Previous work, such as Ivashina and Scharfstein (2010), documents evidence that many corporates drew down on existing credit lines following the collapse of Lehman (see introduction). Using this as a shock is therefore a proxy for a drawdown shock.

I examine the effect of banks' exposure to undrawn credit lines on their credit growth following the shock. I show preliminary evidence which is consistent with my theory. However, I cannot say much about causality, since I cannot properly control for selection issues, given the available data, as discussed below. Moreover, I do not observe the flow of drawdowns or even the stock of drawn lines. This section should therefore be read as a preliminary exploration of the data, rather than a test of the model. I end this section highlighting the need for further work.

5.1. Call Report data. I use quarterly bank level data from Federal Reserve Reports of Income and Condition (Call Reports) to construct a cross-sectional dataset.[42] Each observation is reported at the end of the quarter; so assets for quarter 4 of a given year reflects the amount of assets held by the bank going into quarter 1 of the following year.

Following standard practice in empirical banking research, I have aggregated at the high holder (HH) level so that each bank-quarter observation reflects aggregate information for a banking group. For the purposes of Call Report data, this means aggregating by the identifier 'rssd9348'. After aggregating, this leaves us with 5676 banks. However, I then drop observations where values for relevant variables and controls are zero or missing. After this, and winsorizing at the 1st and 99th percentiles, we are left with 2433 observations for the cross section around the Lehman collapse.

5.1.1. *Data on credit lines.* In addition to 'balance sheet' data, the Call Reports give data on 'off balance sheet' items. This includes undrawn credit lines.

There is no data on the flow of credit line draw downs or on the stock of drawn credit lines. This produces two limitations. First, it prevents us from observing those

[42]Data is available on the website of the Chicago Federal Reserve.

banks for whom the drawdown shock was highest; we will discuss this in more detail below. Second, capturing new loan growth is complicated by the fact that the stock of observed total loans for a given quarter include newly originated-and-drawn loans, as well as drawn credit lines, with no separate entry to identify the latter. Given I want to explore the effect of credit line exposure on newly originated loans, I need to exclude the effect of drawn credit lines on new credit. Cornett et al. (2010) have a nice way of doing this. They sum total loans and total undrawn lines for a given quarter, and define this new variable as 'Credit'. Taking differences, this gives us:

$$\Delta \text{Credit}_t = \Delta(\text{Total Loans})_t + \Delta(\text{Total Undrawn Lines})_t.$$

Any increase in total loans due to drawdowns of lines will be offset by a decrease in total undrawn lines. To see this, suppose that there were no newly originated loans or credit lines: an increase in total loans from $t-1$ to t came entirely from drawn lines. In this case the positive term $\Delta(\text{Total Loans})_t$ will be exactly offset by the negative term $\Delta(\text{Total Undrawn Lines})_t$. Of course, if there were no newly originated-and-drawn loans, but the bank originated new, but undrawn, credit lines, this would show up as a positive value for ΔCredit_t. The growth in credit variable, therefore, captures newly originated supply of credit, whether drawn or undrawn. The key is that it does not include credit that is associated with loans originated before period t.

5.2. Credit line exposure. I am interested in highlighting the role played by the sudden drawing of credit lines, which hitherto have faced zero or low capital charges, since they are off-balance sheet items. The extra capital charge they face once drawn is the relevant shock for the bank. As already highlighted, we cannot observe the flow of drawdowns from the data. We simply observe the stock of undrawn lines for each bank for each quarter; and therefore the stock just before the Lehman collapse. This is broken down into an entry for total credit lines, and also one for credit lines with maturity greater than one year.

In the absence of data on drawdowns, we would like to proxy the potential shock to regulatory risk-weighted assets , should the undrawn lines actually be drawn upon. It is impossible to identify this potential shock for lines of less than one year maturity, since we have no measure of their risk weighting. However, we can capture the exposure of those with longer maturity. We can do this by using information about the capital charge on these lines when they are *undrawn*, as follows.

Unlike short maturity lines, those with maturity greater than one year face a capital charge when undrawn; from the call reports, we can observe the risk-weighted value of these lines. However, they only face a 50% credit conversion factor (CCF), so only half of their risk weighted measure is actually reflected in the bank's overall risk-weighted assets (RWA).[43] Whilst 50% of their risk weighted value is already accounted for, this means the potential shock if all these were to be drawn upon would be equivalent to the remaining 50% of their risk weighted value; i.e. since the CCF is 50%, the potential shock is equal to the existing amount reflected in the bank's RWA. In what follows, we call this variable 'r.w. lines', where:

$$\text{`r.w. lines'} \quad = \quad 50\%(\text{`risk weighted undrawn lines} > 1\text{yr'}).$$

In the regressions below, we will control for the measure of RWAs in the lead up to the Lehman collapse. Given the value of 'r.w. lines' will already be captured in RWA, any significant coefficient on 'r.w. lines' should capture the remaining 50% of risk weighted lines, *not* captured in RWA. This reflects the potential regulatory shock from a drawdown of all lines with maturity greater than one year.

5.3. Constructing variables. The timing of the shock lies towards the end of 2008 (although markets were increasingly disrupted throughout that whole quarter). I construct a variable 'credit growth' which reflects the growth in credit from 2008 quarter 2 to 2009 quarter 4.

$$\text{credit growth} = \frac{\text{credit}_{09q4} - \text{credit}_{08q2}}{\text{credit}_{08q2}}.$$

As discussed above, the variable 'credit' is defined as in Cornett et al. (2010)

I am interested in banks' positions going into the shock. In order to avoid outliers, I take averages of each independent variable for the three quarters preceding the shock.

[43]This is meant to capture the fact they will only be drawn upon with some probability.

For instance, 'total undrawn lines' is calculated as

$$\text{`total undrawn lines'} = \frac{1}{3}(\text{`total undrawn lines}_{07q4}\text{'} + \sum_{i=1}^{2}\text{`total undrawn lines}_{08q_i}\text{'}).$$

It is common in empirical banking research to normalize variables by assets. Previous work has indicated that a bank's size affects its supply of credit; weighting by assets is therefore designed to help compare banks on an equal footing (see Kashyap and Stein (2000)). Moreover, when the dependent variable is a growth variable, it makes sense to normalize independent variables by size to create similar units. As a result, I will use this normalisation in my main results.

6. Empirical Results

Please see the appendix for summary statistics (Table 3).

6.1. Ratio of risk-weighted lines to tier 1 capital. In this section, we show that risk-weighted lines, as a proportion of available regulatory tier 1 capital, are significant in regressions of newly originated credit growth, following the Lehman collapse. We include two credit line variables. In the table below, we show the results from typical regressions with these two variables.

[Table 6 here.]

The first is simply the total stock of undrawn lines in the lead-up to the collapse. This has a significant negative coefficient in most of the regressions, and is consistent with both the liquidity story from Ivashina and Scharfstein (2010) as well as my story about regulatory capital. The second variable, 'r.w. lines', also has a negative and significant coefficient, even when we control for the bank's overall measure of RWA. When using this control, we can interpret 'r.w. lines' as capturing the potential regulatory capital shock, should lines of maturity greater than a year be drawn down. The negative coefficient therefore implies that the capital exposure from these drawdowns had a significant negative effect on new credit growth, over and above the liquidity effect of drawdowns captured by the first variable.

6.2. Regulatory capital buffers. In this section, I examine the significance of capital ratios, insofar as these may become binding with sudden high drawdowns.

The standard regulatory capital ratio under Basel I and II requires banks to hold tier 1 capital equal to at least 4% of RWA. In my regressions, this ratio is denoted by 'cap/RWA', so the degree of constraint for a bank is denoted by 'cap/(RWA) $-$ k' where $k = 4\%$ (this follows a similar empirical specification in Jokipii and Milne (2011)). However, as we have discussed, the measure of RWA does not capture most of the exposure from undrawn lines.

If the theory in this paper is correct, we should see the banks with low available regulatory capital, relative to their *potential* drawdowns, as the ones more likely to restrict credit growth following a drawdown shock. In other words, not only should 'cap/(RWA) $-$ k' be significant, but so should 'cap/(RWA + r.w.lines) $-$ k', where 'r.w. lines' reflects the additional charge should lines of maturity greater than one year be drawn upon. We shall refer to these two variables as the capital buffers.[44]

Unfortunately, these two variables have a high correlation (0.99), so regressions with both ratios will give spurious results (as shown in the final column of the table below). However, we can consider non-linear terms in these variables.

[Table 7 here.]

Table 7 presents these results. The first variable in each regression is simply the total undrawn lines ('lines'), normalised by assets, as discussed above. The significant negative coefficient on this term supports both my argument, and that of Ivashina and Scharfstein and others, but cannot disentangle the two effects. The second variable 'cap/(RWA + r.w.lines) $-$ k' is the attempt at focusing entirely on the capital impact from drawdowns. The coefficient on this term is positive and significant, even when we introduce additional nonlinear functions of 'cap/(RWA) $-$ k', in the 4th column onwards. This suggests that the pre-shock exposure to undrawn lines has an extra impact on credit growth, over and above any information and exposure already reflected in the calculation 'risk-weighted assets'. If banks have high tier 1 capital, relative to the sum of *both* risk-weighted assets and long-maturity undrawn lines, with sufficient buffer above the regulatory minimum (k%), these banks are more likely to have higher credit growth following the drawdown shock.

[44] Thanks to A. Milne for suggesting this empirical specification.

6.3. Demand effects. In the first regressions of each table, we cannot say anything concrete about causality. For instance, it could be that banks with high relative exposure are also the banks who faced lower demand for credit following the Lehman collapse. Whilst this does not seem particularly plausible, it is important to try to control for such demand effects. Following Cornett et al. (2010), I use the following variables to control for demand variation across banks; the amount of commercial and industrial loans, and real estate loans, in the bank's loan portfolio. I show results where these variables are calculated as proportions of the bank's loan portfolio. As we can see from both tables, the causal variable of interest is still significant when we include these demand controls.

However, this is still a fairly crude and imperfect control for demand variation, since we cannot directly observe the demand side of loans in this dataset. As further work, it would be good to find data in which one could observe the demand side of loan originations. In an ideal world, firm level data would enable one to more fully disentangle the degree to which a reduction in credit growth represents misallocation of credit across borrowers.

6.4. Discussion: Empirical Section. The main results in this section showed the significant negative effect of risk weighted long-maturity lines, as a proportion of capital, on new credit growth following the Lehman collapse. When controlling for RWAs, this captures the effect from the remaining (50%) risk weighted exposure of these lines, not captured in RWAs. This supports the argument that undrawn credit lines represent a regulatory capital exposure for the bank, restricting new credit during a period of high drawdowns.

The second set of regressions lent further support to this argument, by focussing on the regulatory capital buffer. This analysis suffered from the close correlation between the two regulatory buffers (with and without 'non r.w. lines'). However, we showed results obtained using non-linear terms, in which the regulatory buffer including lines was significant, even when controlling for terms involving the buffer without lines.

However, we have been unable to capture a full measure of the risk weighted exposure from off balance sheet lines, since we do not observe that for lines with maturity less than one year. Nevertheless, this does not preclude us from focussing just on the exposure of long-maturity lines. Indeed, drawdowns of these lines should have a greater effect in constraining credit; their maturity means their presence on the balance sheet would last longer, arguably having a greater impact on new credit provision.[45]

[45] A proxy for the value of short-maturity lines was analysed, but given the absence of this measure in call reports, it must be calculated using the difference between total undrawn lines and long-maturity lines. This has two limitations. First, it does not capture the risk-weighted value of these lines. Second, it results in high correlation (0.91) with the first variable in the regressions (total undrawn lines), leading to inconclusive results.

7. Conclusion

The key contribution of this paper has been to highlight the way capital charges on credit lines can cause misallocation of bank credit following a market shock. Since undrawn lines face much lower capital charges than drawn lines, the bank will face a sudden increase in capital requirements when a large amount of credit is drawn from existing committed lines. If regulatory capital is scare, the bank may have to cut back on lending to other borrowers - those to whom it did not have a pre-commitment to lend. This can result in misallocation of bank credit across different types of borrowers, despite the bank setting optimal credit line exposure ex ante. Empirical evidence is at least consistent with such an effect.

The drivers of this mechanism have been explored in the theoretical section. The bank has an incentive to offer a large amount of credit lines. Credit lines produce surplus in the good state by signalling firm quality to the market, but are not associated with a significant cost to the bank in the good state as long as they remain undrawn. Since credit lines provide a signal of firm quality to the market in the good state, the bank will always have an incentive to set optimal exposure such that misallocation of credit occurs in the bad state. This is accentuated if banks place low probability on the bad state. Conversely, if there were no requirement for signalling at period 1, there would be no need for credit lines and no ex post misallocation.

This has implications for policy. An increase in the percentage capital requirement for *un*drawn lines (and thus a decrease in the *extra* percentage requirement when they are drawn down) can decrease misallocation in the bad state; this is because it forces the bank to incur the cost of credit lines in the good state. This should not, however, be taken as a clear policy prescription: in this model, we do not model optimal policy, which would additionally take into account the social benefits of credit lines in the good state.

We also find implications for countercyclical capital requirements: the model predicts that lower capital requirements in the bad state will reduce misallocation in that same state.

Such misallocation is made worse if banks anticipate funding costs will be low in the good state. A low funding cost in the good state increases the surplus obtained by credit lines in the good state. Anticipated policy interest rates in the good state, which impact funding costs in that state, therefore play a role in affecting the degree of misallocation following market shocks.

In future research, it would be helpful to incorporate this model into a dynamic framework. This would enable one to draw further implications for both counter-cyclical capital regulations and monetary policy. A richer model would also enable a consideration of optimal policy, taking into account the benefits from credit lines in the good state. Complementing the theoretical work, it would also be useful to explore these effects in a more comprehensive dataset, in which one could observe both the demand and the supply side of credit provision. One could then examine the extent of misallocation in more detail. It would also open up further possibilities of research, in particular to consider how exposure to credit lines affects monetary policy transmission.

Appendix 1: Credit Lines vs Bank Loans

In the text, I show the bank will optimally choose the marginal credit line such that there is misallocation in the bad state. However, is there a γ^* such that the bank would want to issue standard bank loans (non committed) to type 1 firms in the range $(\gamma^*, \overline{\gamma}]$ but credit lines to a lower range $[\gamma_c, \gamma^*]$? This would involve choosing γ^* and γ_c. Below I show that the bank would always want to issue credit lines rather than bank loans to the full range $[\gamma_c, \overline{\gamma}]$.

To see that the CAR will always bind in the bad state in equilibrium (at the optimal *gamma** and γ_c), consider the following argument. Suppose it were not to bind; then the bank would just issue credit lines to all firms in the range $[\underline{\gamma}, \overline{\gamma}]$. However, this immediately contradicts the premise that the CAR does not bind.

In equilibrium, therefore, the CAR must bind. In this case, the bank's problem, corresponding to the function (3.4), will be given as follows

$$\begin{aligned} S_1 \equiv\ & (1-q)\int_{\gamma^*}^{\overline{\gamma}}[pR+(1-p)\gamma - i_G]d\gamma + (1-q)\int_{\gamma_c}^{\gamma^*}[pR+(1-p)\gamma - i_G]d\gamma \\ & +q\int_{\gamma_c}^{\gamma^*}[pR+(1-p)\gamma - i_B]d\gamma - q\int_{\underline{\gamma}}^{\gamma_{bd}(\gamma^*)}[pR+(1-p)\gamma - i_B]d\gamma \end{aligned}$$

The first term reflects the surplus from bank loans, whilst the second term reflects surplus from credit lines, both in the good state. The third term reflects the (social) surplus from drawdowns in the bad state, whilst the final term captures the foregone loans to type 2 firms in the bad state. As before, $\gamma_{bd}(\gamma^*)$ is given by the binding CAR in the bad state

$$\gamma_{bd}(\gamma^*) = \overline{\gamma} - \gamma_c + \gamma^* - \frac{\omega}{k}$$

Loans to the type 1 firms are negative NPV in the bad state for any lender, given non pledgeable assets (see condition 2.2). The bank will therefore choose not to lend to the type 1 firms to which it does not have a commitment, i.e. those type 1 firms in the range $[\gamma^*, \overline{\gamma}]$.

The first order condition for γ_c is exactly the same as that in the main text. The first order condition for γ^* is given by

$$\begin{aligned} \frac{\partial S}{\partial \gamma^*} &= -(1-p)(\overline{\gamma} - \gamma_c - \frac{\omega}{k}) > 0 \\ \text{since } \overline{\gamma} - \gamma_c - \frac{\omega}{k} &< 0 \end{aligned}$$

To verify that $\overline{\gamma} - \gamma_c - \frac{\omega}{k} < 0$, consider the following: the optimal γ_c is the same as in the main text, and the CAR in the main text is given by

$$\overline{\gamma} - \gamma_c + \overline{\gamma} - \gamma_{bd}(\gamma_c) = \frac{\omega}{k}$$

Rearranging, this gives us

$$\begin{aligned} \overline{\gamma} - \gamma_c - \frac{\omega}{k} &= -[\overline{\gamma} - \gamma_{bd}(\gamma_c)] \\ &< 0 \end{aligned}$$

As a result, the optimal problem is identical to that in the text and the bank will set

$$\gamma^* = \overline{\gamma}$$

Intuitively, by offering bank loans, rather than credit lines, to firms with high γ, the surplus is unchanged in the good state, but worse in the bad state. This is because the loan to these firms is negative NPV to the bank in the bad state, due to low pledgeability β_B. As a result, in the bad state the bank ends up rationing these firms, instead of type 2 firms with lower γ. In equilibrium, the bank will prefer to commit itself via a credit line to all type 1 firms, because the firm is willing to pay a higher ex ante fee than the expected cost to the bank of such a commitment.

Appendix 2: Empirical Results

Table 3. Summary Statistics

	mean	standard deviation	minimum value	maximum value	10th percentile	90th percentile
credit growth	0.01	0.16	-0.89	1.82	-0.17	0.17
lines/assets	0.12	0.08	0.00	1.55	0.05	0.21
r.w. lines/capital	0.24	0.22	0.00	1.89	0.03	0.52
cap/(RWA + r.w.lines)$-k$	0.09	0.05	0.02	0.60	0.05	0.14
(cap/(RWA + r.w.lines)$-k)^2$	0.01	0.02	0.00	0.36	0.00	0.02
(cap/(RWA + r.w.lines)$-k)^3$	0.00	0.01	0.00	0.21	0.00	0.00
cap/(RWA)$-k$	0.09	0.05	0.02	0.65	0.05	0.14
(cap/(RWA)$-k)^2$	0.01	0.02	0.00	0.42	0.00	0.02
(cap/(RWA)$-k)^3$	0.00	0.01	0.00	0.27	0.00	0.00
liq. assets/assets	0.19	0.11	0.01	0.89	0.07	0.35
core deposits/assets	0.63	0.17	0.10	1.21	0.41	0.85
log assets	12.39	1.20	10.13	16.96	10.95	13.90
C and I / tot. loans	0.15	0.09	0.00	0.71	0.06	0.26
Real Est. / tot. loans	0.71	0.16	0.00	1.00	0.48	0.88
Observations	2433					

This table shows summary statistics for the variables used in the regressions below. All variables are shown as fractions (e.g. 0.12 is 12%). The variables are constructed from quarterly US commercial bank balance sheet data, aggregated at the high holding company, obtained from the Federal Reserve Call Reports of Condition and Income. The data is available at www.chicagofed.org. All variables, except for credit growth, are taken as the average for the three quarters prior to 2008q3 (in which the Lehman collapse occurred). The variable 'credit growth' reflects the growth in newly originated credit from 2008 quarter 2 to 2009 quarter 4. This captures newly originated loans as well as newly originated credit lines. The variable 'lines' refers to total undrawn lines. The variable 'r.w. lines' refers to the weighted portion of undrawn lines, with maturity greater than one year, that are included in the measure of risk weighted assets. Given the credit conversion factor is 50%, this variable also captures the amount by which risk weighted assets would increase if all such long maturity lines were to be fully drawn.

Table 4. Correlations 1

	lines /assets	cap /(RWA + r.w.lines)$-k$	(cap /(RWA + r.w.lines)$-k)^2$	(cap /(RWA + r.w.lines)$-k)^3$	cap /(RWA)$-k$	(cap /(RWA)$-k)^2$	(cap /(RWA)$-k)^3$	C and I /tot. loans	Real Est. /tot. loans
lines/assets	1								
cap/(RWA + r.w.lines)$-k$	-0.364***	1							
(cap/(RWA + r.w.lines)$-k)^2$	-0.257***	0.913***	1						
(cap/(RWA + r.w.lines)$-k)^3$	-0.157***	0.728***	0.938***	1					
cap/(RWA)$-k$	-0.340***	0.998***	0.916***	0.735***	1				
(cap/(RWA)$-k)^2$	-0.240***	0.900***	0.994***	0.942***	0.908***	1			
(cap/(RWA)$-k)^3$	-0.141***	0.695***	0.908***	0.982***	0.709***	0.932***	1		
C and I / tot. loans	0.254***	-0.124***	-0.0841***	-0.0537**	-0.120***	-0.0831***	-0.0534**	1	
Real Est. / tot. loans	-0.0513*	-0.136***	-0.108***	-0.0689***	-0.132***	-0.103***	-0.0610**	-0.579***	1

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

This table shows correlations between total undrawn credit lines, regulatory buffer variables and demand variables. The variables are constructed from quarterly US commercial bank balance sheet data, aggregated at the high holding company, obtained from the Federal Reserve Call Reports of Condition and Income. The data is available at www.chicagofed.org. All variables are taken as the average for the three quarters prior to 2008q3 (in which the Lehman collapse occurred). The variable 'lines' refers to total undrawn lines. The variable 'r.w. lines' refers to the weighted portion of undrawn lines, with maturity greater than one year, that are included in the measure of risk weighted assets. Given the credit conversion factor is 50%, this variable also captures the amount by which risk weighted assets would increase if all such long maturity lines were to be fully drawn.

Table 5. Correlations 2

	lines /assets	lines /liq. assets	r.w.lines /cap	cap /assets	cap /RWA	liq. assets /assets	core deposits /assets	log assets	C and I /tot. loans	Real Est. /tot. loans
lines/assets	1									
lines/liq. assets	0.579***	1								
r.w. lines/capital	0.584***	0.409***	1							
cap/assets	-0.153***	-0.107***	-0.210***	1						
cap/RWA	-0.340***	-0.292***	-0.335***	0.820***	1					
liq. assets/assets	-0.331***	-0.522***	-0.290***	0.355***	0.623***	1				
log assets	0.425***	0.347***	0.397***	-0.293***	-0.329***	-0.372***	1			
core deposits/assets	-0.378***	-0.374***	-0.368***	0.175***	0.306***	0.416***	-0.645***	1		
C and I / tot. loans	0.254***	0.141***	0.163***	-0.0640**	-0.120***	-0.0100	0.0505*	-0.117***	1	
Real Est. / tot. loans	-0.0513*	0.0979***	0.0889***	-0.0762***	-0.132***	-0.298***	0.262***	-0.262***	-0.579***	1

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

This table shows correlations between total undrawn credit lines (lines), risk-weighted lines (r.w. lines) as a proportion of capital, and control variables. The variables are constructed from quarterly US commercial bank balance sheet data, aggregated at the high holding company, obtained from the Federal Reserve Call Reports of Condition and Income. The data is available at www.chicagofed.org. All variables are taken as the average for the three quarters prior to 2008q3 (in which the Lehman collapse occurred). The variable 'lines' refers to total undrawn lines. The variable 'r.w. lines' refers to the weighted portion of undrawn lines, with maturity greater than one year, that are included in the measure of risk weighted assets. Given the credit conversion factor is 50%, this variable also captures the amount by which risk weighted assets would increase if all such long maturity lines were to be fully drawn.

Table 6. Regressions: Ratio of risk-weighted lines to capital

VARIABLES	(1) credit growth	(2) credit growth	(3) credit growth	(4) credit growth	(5) credit growth
lines/assets	-0.126**	-0.101*	-0.0723	-0.0663	-0.166***
	(0.0540)	(0.0552)	(0.0603)	(0.0598)	(0.0599)
r.w. lines/capital	-0.0751***	-0.0731***	-0.0652***	-0.0609***	-0.0550***
	(0.0179)	(0.0179)	(0.0179)	(0.0179)	(0.0179)
cap/RWA	0.335***	0.157*	0.144	0.158*	0.181*
	(0.0775)	(0.0943)	(0.0939)	(0.0936)	(0.0954)
liq. assets/assets		0.135***	0.118***	0.0935**	0.0515
		(0.0397)	(0.0386)	(0.0388)	(0.0420)
log assets			-0.00793**		
			(0.00398)		
core deposits/assets				0.0906***	
				(0.0253)	
Real Est. / tot. loans					-0.172***
					(0.0254)
C and I / tot. loans					-0.0702
					(0.0559)
Constant	-0.00489	-0.0107	0.0869*	-0.0678***	0.138***
	(0.0138)	(0.0139)	(0.0482)	(0.0230)	(0.0264)
Observations	2,433	2,433	2,433	2,433	2,433
R-squared	0.042	0.047	0.050	0.053	0.066

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

This table shows regressions of credit growth on both total undrawn lines ('lines') and the ratio of risk-weighted undrawn lines to tier 1 capital ('r.w.lines/cap'). The regressions also include controls, as indicated. The variables are constructed from quarterly US commercial bank balance sheet data, aggregated at the high holding company, obtained from the Federal Reserve Call Reports of Condition and Income. The data is available at www.chicagofed.org. All variables are taken as the average for the three quarters prior to 2008q3 (in which the Lehman collapse occurred). The variable 'credit growth' reflects the growth in newly originated credit from 2008 quarter 2 to 2009 quarter 4. This captures newly originated loans as well as newly originated credit lines. The variable 'lines' refers to total undrawn lines. The variable 'r.w. lines' refers to the weighted portion of undrawn lines, with maturity greater than one year, that are included in the measure of risk weighted assets. Given the credit conversion factor is 50%, this variable also captures the amount by which risk weighted assets would increase if all such long maturity lines were to be fully drawn.

Table 7. Regressions: Regulatory Ratios

VARIABLES	(1) credit growth	(2) credit growth	(3) credit growth	(4) credit growth	(5) credit growth
lines/assets	-0.128***	-0.112**	-0.0823	-0.154**	-0.101*
	(0.0487)	(0.0496)	(0.0562)	(0.0657)	(0.0539)
cap/(RWA + r.w.lines)$-k$	2.220***	1.898***	1.725***	1.871***	6.813**
	(0.372)	(0.400)	(0.420)	(0.422)	(3.239)
(cap/(RWA + r.w.lines)$-k)^2$	11.08	13.15*	13.33*	6.203	-10.03
	(7.715)	(7.861)	(7.857)	(8.070)	(16.73)
(cap/(RWA + r.w.lines)$-k)^3$	-11.86	-13.86	-14.07	-6.039	7.432
	(8.713)	(8.858)	(8.859)	(9.060)	(15.99)
(cap/(RWA)$-k)^2$	-18.45**	-19.70**	-19.19**	-12.93*	2.564
	(7.599)	(7.668)	(7.660)	(7.819)	(16.74)
(cap/(RWA)$-k)^3$	18.96**	20.15**	19.64**	12.87	0.0562
	(8.326)	(8.395)	(8.387)	(8.543)	(15.73)
liq. assets/assets		0.106**	0.0943**	0.00487	
		(0.0422)	(0.0410)	(0.0441)	
log assets			-0.00695*	0.00336	
			(0.00406)	(0.00445)	
core deposits/assets				0.0609**	
				(0.0263)	
C and I / tot. loans				-0.0404	
				(0.0540)	
Real Est. / tot. loans				-0.159***	
				(0.0263)	
cap/(r.w.ass)$-k$					-4.673
					(3.293)
Constant	-0.102***	-0.103***	-0.00872	-0.0402	-0.0938***
	(0.0217)	(0.0217)	(0.0572)	(0.0733)	(0.0226)
Observations	2,433	2,433	2,433	2,433	2,433
R-squared	0.050	0.053	0.055	0.075	0.051

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

This table shows regressions of credit growth on regulatory buffers, accounting for the extra exposure to undrawn lines. The regressions also include controls, as indicated. The variables are constructed from quarterly US commercial bank balance sheet data, aggregated at the high holding company, obtained from the Federal Reserve Call Reports of Condition and Income. The data is available at www.chicagofed.org. All variables are taken as the average for the three quarters prior to 2008q3 (in which the Lehman collapse occurred). The variable 'credit growth' reflects the growth in newly originated credit from 2008 quarter 2 to 2009 quarter 4. This captures newly originated loans as well as newly originated credit lines. The variable 'lines' refers to total undrawn lines. The variable 'r.w. lines' refers to the weighted portion of undrawn lines, with maturity greater than one year, that are included in the measure of risk weighted assets. Given the credit conversion factor is 50%, this variable also captures the amount by which risk weighted assets would increase if all such long maturity lines were to be fully drawn.

CHAPTER 2

Information Asymmetries and Spillover Risk in Settlement Systems

1. Introduction

It goes without saying that financial markets in different locations and jurisdictions cannot operate in isolation; a disruption in one may well impact another. A similar argument can be applied to large-value payment and settlement systems, the platforms on which banks settle the underlying claims between two financial transactors. Whilst existing research has focused on the risks inherent *within* a given system, there is very little academic research on the way in which one system can destabilize another.

In this paper, I focus on information asymmetries between different settlement systems. I show how the presence of these asymmetries can increase the probability that a disruption in one system may spill over into another. The channel of transmission comes from the strategic behaviour of banks which participate in both systems. I assume only those banks which participate in the problem system are aware of the disruption, including the multi-system participants. These banks may continue to make payments to the affected bank, thus risking a shortage of liquidity, if they believe they can rely on incoming liquidity from other agents in other systems. Intuitively, the multi-system participants may have an incentive to take advantage of the ignorance of agents in other systems. In so doing, they run the risk of themselves defaulting in hitherto unaffected systems.

To understand the motivation and intuition for this, it is important to consider some key issues relevant to payment systems. Banks operate within payment systems as facilitators; this is unlike standard banking models which focus on the role of banks as liquidity transformers and their relationship with depositors. In payment systems, banks receive exogenous payment instructions and must choose, intraday, whether to settle these payments early or to bear a cost of delay. Typically these payments are time-specific, and banks are legally obliged to either pay them by a given time intraday, or by the close of day. Since I focus on Real-Time Gross Settlement (RTGS) systems, banks must have sufficient liquidity to fund each gross payment. Banks therefore trade off delay costs against the risk that by making an 'early' payment they lose liquidity and may have insufficient liquidity later on in the game (or have to bear a high cost to find extra liquidity). Of course, if they receive incoming payments this provides them with liquidity which they can recycle.

Operational problems and the risk of payment failure are potential dangers within a payment system. An operational problem may result in a liquidity sink, in which a bank which has received incoming payments is unable to make outgoing payments. Sometimes it may not be clear whether a payment problem at a given bank represents merely a transitory technical issue, or may actually lead that bank to default on payments which are due by the end of the day. In other words, this means other banks must operate under uncertainty as to whether they will receive their expected incoming payments. Policymakers often encourage banks to make payments early in the day in order to reduce the potential disruption within the system of any later operational problem.

I build a model in which a bank operates in both systems, and recycles liquidity intraday across the two. With some probability, there is a disruption in one system (call it the 'offshore' system), reducing the incoming liquidity received by the multi-system bank. Specifically, this reduces the available liquidity for the multi-system bank to use for making payments in the other system (call this the 'domestic' system). However, with high delay costs, this bank may continue making outgoing payments offshore if it believes it can recycle incoming liquidity from the domestic system. In so doing, it may end up with a liquidity shortage in the domestic location later in the day, if the problem offshore is never resolved. In this event, it will default in the domestic location.

The key driver comes from the fact that domestic only participants do not know of the problem offshore, and so may continue making early payments to the multi-system bank, believing they will receive back liquidity from their counterparty later in the day. Crucially, if the multi-system bank believes these domestic participants will pay early, it has a greater incentive to keep making payments in the offshore system, and thereby increase the risk of spillover to the domestic system.

This has serious implications for policy. First, it implies that information flows from one system to another are crucial in reducing risks of spillover. Second, and perhaps more fundamentally, it casts doubt on the common perception of policymakers that encouraging banks to settle payments early in the day is always beneficial for the system. In fact, in the model here, if domestic participants are encouraged to make early payments, this increases the risk of spillover to and default within that system.

In an extension to the model, I ask the question of whether a Liquidity Saving Mechanism is a solution to the problem of information asymmetry and spillover. A liquidity saving mechanism (LSM) is a queuing arrangement sometimes put in place in a real-time gross settlement system (RTGS), in order to reduce the liquidity needs of agents. It allows banks to make a payment contingent on an incoming payment from another agent (or indeed on some other event); in this sense, by queuing the payment, the bank can ensure it does not fall short of liquidity, but reduces the probability of paying a delay cost.

Specifically, I show that the introduction of an LSM in one system, in which outgoing payments may be made contingent on in-coming payments, has a positive externality; other systems may benefit from the increased stability and reduced risk of spillover the LSM provides. This contributes to an existing literature on LSMs which analyses only the effects on the system where the LSM is implemented.

However, by the nature of this externality, it suggests that policymakers who are concerned about spillover from outside to their own system cannot simply reduce spillover by introducing an LSM. In fact, it requires the implementation of an LSM within the *other* systems for the risk of spillover to be reduced.

Why is it important to consider the potential for problems within, and spillover between, payment systems? Although they receive less public attention, large-value payment and settlement systems have significant systemic importance, as highlighted in a recent consultative report published by the Committee on Payment and Settlement Systems (Committee on Payment and Settlement Systems (2011)). Whilst they mostly operate well, the potential for severe disruption is huge given the network effects. In his autobiography, Greenspan documents how there were ‘half a dozen near disasters, mostly involving the payment system’ in the days following Black Monday in 1987 (Greenspan (2008) p.108).

This systemic importance is only going to increase as more settlement systems develop, outside of the direct control of central banks and also across different jurisdictions. At the same time, it seems likely that information asymmetries will be worse between systems that operate in different locations.

In fact, there is an increasing trend for banks to operate in multiple systems across different jurisdictions, and to recycle liquidity across those systems. For instance, the growth of offshore same-currency payment systems seems likely to encourage the use, not just of dual-system membership, but also of liquidity recycling across systems. As an example, HSBC has operated a US-dollar RTGS system in Hong Kong (USD CHATS) since 2000; this was followed by the introduction of a Euro RTGS system in 2003 (EURO CHATS), with Standard Chartered Bank as the settlement operator (Institute for International Monetary Affairs (2004)).

Large multinational banks have also shown increasing tendencies to recycle liquidity across systems. They have access to global liquidity pools, or alternative means of transferring liquidity cross-border intraday (Committee on Payment and Settlement Systems (2006)). These large banks can play significant roles in payment systems, accounting for a large proportion of payment values (Committee on Payment and Settlement Systems (2008) p.23).

Moreover, over the past decade, new forms of settlement arrangements have arisen that explicitly connect different systems (Committee on Payment and Settlement Systems (2008)). A good example is CLS Bank, which settles foreign exchange trades. Participants must ‘pay in’ liquidity in the relevant currency in advance of a transaction, via the relevant payment system for that currency (e.g. Fedwire for USD and CHAPS for Sterling). Although CLS only settles trades when both sides have ‘paid in’, the system is a good example of how banks may recycle liquidity across jurisdictions and across trades. In this way, any operational problem in one jurisdiction could have consequences for banks recycling liquidity across CLS and other payment systems, causing further problems for their CLS counterparties in other jurisdictions. As the

Committee on Payment and Settlement Systems (2008) argues, if a participant in a domestic system has an operational problem and fails to make payments, this leads to other banks having lower balances than anticipated; this can lead to the failure of settlement in CLS which can potentially spill over into other systems connected with CLS.

Finally, it is worth mentioning that as payment systems grow and the connections between banks within a given system increase, it may be relevant to think of spillover within a single system. Specifically, if banks know of potential problems with their counterparties but not their counterparties' counterparties, the mechanism in this model would apply. The key ingredient comes from the nature of the information asymmetry: the fact that some banks are aware of a given problem but others are not.

1.1. Literature. As highlighted above, the key tradeoffs facing banks in payment systems center on the timing of payments. Banks have to choose when, intraday, to settle a payment, often in the context of uncertainty about incoming liquidity flows. This reflects the role of banks as facilitators, but contrasts with traditional banking literature.

Various papers have examined banks' liquidity management, in a game theoretic context. However, most models focus on just one payment system, ignoring interactions between systems. Bech and Garrett (2003) focus on the strategic behaviour of banks who choose whether to settle a payment early or incur a delay cost. If liquidity costs are high, banks will wish to wait until receiving incoming payments, the liquidity of which they can recycle to make their own payment later on. In a similar vein, Angelini (1988) explores the role played by liquidity costs and costs of delay.

Mills and Nesmith (2008) extend the framework, in the case of priced credit, to examine the role of settlement risk.[1] The risk that a bank will not get paid in a certain period, due to settlement delay, becomes an added incentive to delay payment. Kahn et al. (2003) also explore the role of settlement risk, whilst Merrouche and Schanz (2010) extend this idea to examine how the time of day affects whether banks will delay, given a counterparty has an operational problem. Afonso and Shin (2009)

[1]Settlement risk is the risk that a payment, which has been instructed by the bank's customer, is not finally settled.

document how banks rely heavily on recycling liquidity intraday, and provide a model to show the negative consequences of this for payment systems if banks start to hoard liquidity, as they did during September 11, 2011.

Schanz (2009) considers the effects of global liquidity pools on the optimal timing of payments in a foreign exchange settlement system. Although information asymmetries play a role in that paper, the focus is rather on the degree of information asymmetry between counterparties, relative to that between subsidiaries, which affects the optimal amount of intra-day lending to a given bank.

Following empirical evidence, there are various approaches to modeling the costs of liquidity. In practice, some RTGS systems allow banks to obtain uncollateralised intraday liquidity from the settlement bank for a fee, such as Fedwire (Federal Reserve). Yet others require banks to have in place sufficient collateral in order to obtain liquidity, such as Target 2 (ECB) and CHAPS (Bank of England).

In Bech and Garrett (2003), collateral is posted intraday, and also retrieved intraday; the cost of liquidity therefore depends on the duration for which it is required. However, in practice, banks tend to post collateral before the day has started; in CHAPS in the UK, collateral is generally posted for the entire day (see discussion in Becher et al. (2007)) In particular, collateral is not often retrieved intraday and used for an alternative purpose and banks very rarely meet the maximum debit position for which they are covered by collateral. For this reason, some papers, such as Merrouche and Schanz (2010) and Manning and Willison (2006), model banks' main collateral decisions as occurring ex ante.

In the following paper, therefore, I do not attempt to provide a comprehensive model of liquidity costs. I assume two things: banks do not have unlimited collateral available, and collateral must be posted ex ante, before the game starts. Unlike Merrouche and Schanz (2010), I do not allow for the possibility of obtaining extra collateral intraday. Indeed, as Becher et al. (2007) point out, it seems likely there are substantial frictions in obtaining extra collateral intraday, at least at short notice.[2] I do not explicitly model the ex ante collateral posting decision, although I do discuss necessary parameter conditions on collateral for the equilibria to exist.

[2]This may well be one reason why CHAPS participants tend to post collateral only at the beginning of the day.

I depart from this existing literature by examining how problems in one system can spill over and affect another, via strategic behaviour and information asymmetry. There have only been a few papers considering interdependencies between systems. Manning and Willison (2006) conduct simulation exercises exploring optimal collateral choice in the context of multiple systems, but do not model spillover between banks who have more than one counterparty. Nor do they consider the role of informational asymmetries. Renault et al. (2009) develop a numerical model to analyse interdependences between two systems with common participants: however they do not analyse strategic behaviour and the role of information asymmetries. Those papers that do consider strategic behavior tend to focus on a single system.

Furthermore, this paper contributes to the literature on liquidity saving mechanisms. Various papers have analyzed LSMs (e.g. Martin and McAndrews (2008), Jurgilas and Martin (2010)). However, these papers do not focus on how the introduction of an LSM in one system provides a positive externality to another system. This is what I explore using the model below.

In the following section, I discuss the model. After introducing the framework, I examine the role played by information asymmetries, and explore how the equilibrium changes as these asymmetries decrease. I then discuss policy implications and extend the model to consider Liquidity Saving Mechanisms.

2. The Model

2.1. Players. The model consists of a game between two strategic players, Bank A and Bank D. Each player will receive exogenous instructions to settle payments with other agents at various points in the game. I assume each system is a Real-Time Gross Settlement system (RTGS); sufficient liquidity must be held by the relevant bank before it can settle a payment. In order to settle payments with a given counterparty, the player must operate within the same payment system as that counterparty.

I assume there are two separate payment systems: I shall refer to these as the 'domestic' system and the 'offshore' system. Bank D operates solely in the domestic system. However, Bank A operates in both the domestic and the offshore systems. As a result, we can think of Bank A as a multi-system participant.

2.2. Payment Instructions and Shocks. There are four periods in the game, corresponding to the Morning and Afternoon of the two payment systems. In this way, the framework of the model builds upon that of Bech and Garrett (2003), but introduces a second system to operate in parallel with the first. Across the four periods, there will be a combination of payment instructions and shocks.

Table 1. Time Periods

Time period (t)	t=1	t=2	t=3	t=4
Offshore	Morning		Afternoon	
Domestic		Morning		Afternoon

Payment instructions are received in the morning of each system (periods 1 and 2), and also in the afternoon of the domestic system (period 4). Since Bank A operates in both systems, it receives a payment instruction in all three of those periods. By contrast, Bank D just receives an instruction in period 2, the domestic morning.

In Bech and Garrett (2003), there is ex ante uncertainty as to whether each bank will receive a payment instruction in each period. In this model, however, the framework is more complicated due to the coexistence of two systems. For parsimony, I limit such uncertainty to the final period, the domestic afternoon. I assume the payment instruction in that period for Bank A to pay Bank D can either be high (with probability p_A) or low.

In a later section, I discuss necessary conditions on collateral values and payment instructions. For the purposes of introducing the equilibrium, however, I focus on a specific numerical example at this stage.

Table 2 summarizes per period payment instructions, where m_{ij} denotes a morning payment from bank i to bank j and s_{ij} denotes afternoon payments in the same way.

Table 2. Payment Instructions

Time period (t)	t=1	t=2	t=3	t=4
Bank A	$m_{AO} = 2$	$m_{AD} = 1$		$s_{AD} \in \{1, 2\}$
Bank D		$m_{DA} = 1$		$s_{DA} = 1$

Note: Morning instructions from bank i to j are denoted by m_{ij} while afternoon payments are denoted equivalently by s_{ij}. The letter O refers to Bank A's (nonstrategic) counterparty offshore.

In the domestic system, Bank A and Bank D are the counterparties to each other's payments. So, for instance, any payment sent by D will be received by A. In the offshore system, however, I do not explicitly model the counterparty to A's payments.

Instead, we assume that the offshore system is a source of shocks to Bank A's liquidity management and payment flows. The idea of these shocks can be motivated by considering a non strategic agent who is A's counterparty in the offshore system. Before period 1, nature determines whether the state of the world is good, or bad. In the good state, Bank A will receive payment inflows from its counterparty at period 1 in the offshore system to exactly offset the payment outflows which itself owes the offshore counterparty; ceteris paribus this increases its available liquidity for future payment outflows. In the bad state, however, Bank A will not receive any payment inflow in period 1. The bad state occurs with probability q. Conditional on the bad state, A may instead receive the payment inflow in the offshore afternoon (period 3). However, this only occurs with probability γ.

In a reduced form way, these shocks capture the idea of a liquidity sink in the offshore system. So I can motivate the shocks as follows. In the bad state, an offshore participant has an operational problem at the beginning of period 1. This prevents it initially from making outgoing payments; however it can still receive and process incoming payments. The problem may or may not be resolved by the end of the day. If the problem is resolved, the payment to A will be made at period 3: if not, the non strategic participant technically defaults and A does not receive the inflow.

Before moving on to the timeline and a discussion of strategies, please find below a table referencing the notation that will be introduced:

Table 3. Parameter List

Parameters	Meaning
q	probability of bad state
γ	probability offshore problem persists, conditional on bad state
p_A	probability A receives large payment instruction at period 4
β	cost of default
d^h	delay cost in domestic (home) system
d^f	delay cost in offshore (foreign) system
π	D's cost if A defaults

2.3. Timeline, Strategies and Information Asymmetry. The table below summarises the discussion so far, and also indicates where agents make strategic decisions:

FIGURE 1. Timing Description

Ex ante	Period 1: Offshore morning	Period 2: Domestic morning	Period 3: Offshore afternoon
Nature:	**A**:	**D:**	**Nature:**
chooses state	chooses early or delay	chooses early or delay	In bad state, chooses whether offshore problem persists
(Probability of bad state: q)		(A's move is trivial)	(Conditional probability problem persists: γ)

Bank A moves at period 1 and Bank D moves at period 2. In each case, the bank chooses whether to act immediately upon the payment instruction (to pay early) or to delay until the afternoon of the respective system (to delay). Recall the requirement of an RTGS system, that a bank must have sufficient liquidity upfront in order to settle a payment. The choice over paying early or delaying therefore determines when that bank requires liquidity, and also when its counterparty receives the liquidity inflow.

In the bad state nature moves again at period 3, and there is technical default in the offshore location with probability γ; in other words, A does not receive the offshore incoming payment.

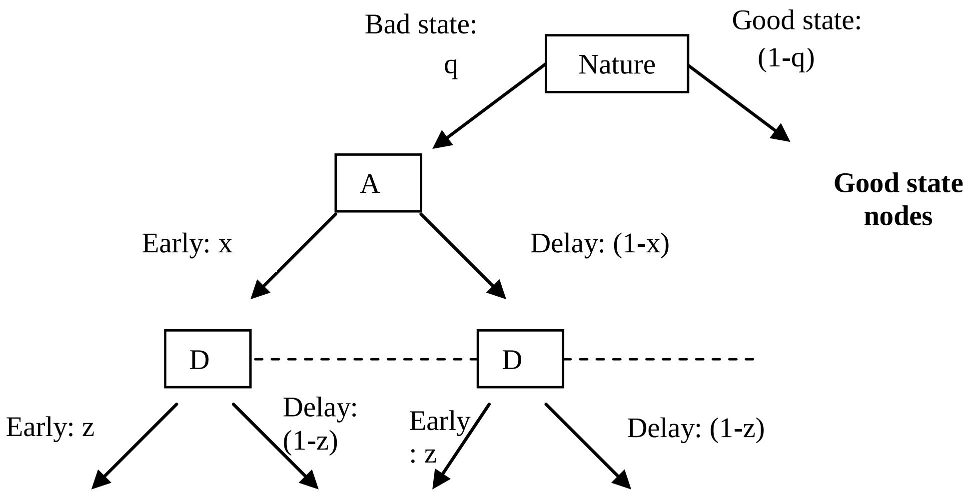
The strict deadline for all payments is the end of the day (period 3 for offshore payments and period 4 for domestic payments); afternoon behavior is therefore constrained by the strategy chosen earlier. So the strategic moves for each player are limited to the morning periods.

In addition, there is no strategic move for Bank A at period 2, despite the fact that A receives a morning instruction in this period. As we shall see, Bank A's move at period 2 makes no difference to its expected payoffs; for this reason, we can ignore its move at this period.

The strategic decisions of the two banks are complicated by an information asymmetry. Since the offshore system is in a different jurisdictional location, as well as a different time zone, I assume there is an informational asymmetry between the locations. The domestic-only agent, Bank D, will not know whether the bad state occurred in the offshore location. D will also be unaware of the strategy played by A at period 1.

If we focus on the bad state game, the strategic moves and information asymmetry can be represented in the following diagram: (with x the probability that A plays early at period 1 and z the probability that D plays early at period 2).[3]

FIGURE 2. Information Asymmetry Game Tree



Note: The probability of the bad state is denoted q, the probability A pays early is denoted x and the probability D pays early is denoted z.

[3]Note that x and z are determined in equilibrium.

2.4. Costs and Liquidity Management. If a bank chooses to delay settlement of a payment until the afternoon, it incurs a cost of delay. This reflects potential costs such as legal and financial penalties for failing to meet time-critical payments, customer dissatisfaction and loss of future business. In particular, delay costs may capture the penalties for delaying a time-critical payment to another system, like CLS, the multi-currency settlement system. (For more discussion of delay costs see Manning and Willison (2006) and Merrouche and Schanz (2010)). Each payment instruction is an independent event; each one sent by a different customer of each bank. Whilst we can think of settlement between banks as providing offsetting liquidity, the payment instructions themselves are independent so do not offset. As such, I assume delay costs of one bank are not reduced if the bank's counterparty has difficulties meeting its own payment deadlines. (An alternative version of the model, with this assumption relaxed, is discussed in the appendix.)

Delay costs are exogenous to the game and not determined by overall demand and supply in the payment network. I assume delay costs are determined simply by location. Delay costs in the offshore location are denoted by d_f, and in the domestic location by d_h.

I assume that each participant has a limited amount of collateral available throughout the game, against which liquidity can be obtained from the settlement agent.[4] For the multi-system participant (Bank A) this collateral can be recycled across locations, if unused. A bank must have sufficient collateral at the beginning of a period in order to be able to make a payment. I also assume that the opportunity costs of holding collateral are sunk costs; I assume no intraday inter-bank market for liquidity. Given that collateral costs are sunk, these costs do not affect the choice over whether to pay early or late.

I do not model the ex ante choice of how much collateral to set aside. Instead, I assume each bank has 3 units of collateral available at the beginning of the game. Effectively, I am assuming only Bank A, the multi-system participant, has a probability

[4]For simplicity, one unit of collateral enables the bank to obtain one unit of liquidity, such that there is no haircut. Incorporating a haircut would just require a higher amount of ex ante collateral; I discuss this in the later section on collateral values.

of becoming constrained by its available liquidity.[5] In the appendix, I discuss in more detail the conditions under which collateral choice of 3 units would be optimal.[6]

Bank A has limited available collateral; if its decisions lead it to technical default on a payment instruction, failing to settle by the afternoon, it will face a cost. This is denoted by β.

Settlement failure also has an impact on the counterparty to that payment. We therefore assume Bank D bears a cost of lost liquidity in the event of Bank A's default. The magnitude of this cost is denoted by π. This is a reduced form way to model the inconvenience Bank D will face in having less liquidity than expected for making payments to other agents, and in other systems, beyond the end and scope of the game.

Having set out the different costs, we can now consider the tradeoffs faced by each bank in deciding whether to pay early or delay payment.

2.5. Tradeoffs. As a baseline case, consider the good-state game, in which there is no operational problem in the offshore location. This occurs with probability $(1-q)$.

In this case, both banks would play the early strategy, in each location. Given collateral costs are sunk, at the point of choosing strategy, there is no expected cost from paying in the morning. If each bank's counterparty pays early, there is no risk that any bank will reach the limit of its available collateral. There is therefore no expected cost from having insufficient liquidity (and no expectation of technical default). By contrast, each bank faces a positive cost of delay. Paying early is a dominant strategy for both banks.

Consider now the bad-state game, occurring with probability q, in which A does not receive the incoming payment in the offshore location at period 1. Bank A knows what state has occurred, before it chooses its own strategy. Bank A knows that the operational problem will only be resolved with probability γ; only in that instance will it receive the offshore incoming payment in period 3. In this section, I assume

[5] This serves our purpose, since I am primarily interested in the actions of A as the main strategic player. Allowing Bank D to become constrained would only complicate the game without changing the underlying mechanism.

[6] Notice that the game would be trivial under collateral values of less than 3 for Bank A. If Bank A had less than 3 units of collateral, it would default with positive probability regardless of the state if Bank D delayed payments until the afternoon.

uncertainty regarding the offshore location is resolved only at the beginning of period 3.[7]

What are the tradeoffs faced by Bank A in this state? If it pays early, it may never receive an incoming liquidity from the offshore bank, if the operational problem continues throughout the day. This reduces the available liquidity for Bank A to recycle in the domestic system. Absent any other incoming liquidity, A would then be unable to complete domestic payments at period 4, as it would have had received a morning instruction to pay D $m_{AD} = 1$ at period 2, and would receive a minimum instruction of $s_{AD} = 1$ at period 4; in this case, A would technically default. A negative payoff from default would then be incurred.

However, if delay costs in the offshore location are high, A may be reluctant to wait until the afternoon to make a payment. Indeed, A may well be reluctant to delay if the operational problem is unlikely to persist (γ is low).

In fact, the incentive to keep sending payments to a bank with an operational problem was evident in the payment system directly after the attacks on 9/11. Whilst some New York banks faced problems making outgoing payments, other banks in the system were keen to keep making payments to those banks, in full knowledge of the situation. The motivation came from the fact they had legal obligations to make payments by a specific time, and would face consequences if they failed to meet that deadline.

In considering its tradeoffs, however, Bank A will consider the strategy that might be chosen by Bank D. This is important. The domestic agent who knows nothing about the operational problems offshore may actually enable Bank A to avoid delay costs offshore AND reduce the risk of default. If Bank D makes a payment to A in the domestic morning, this provides A with 'free' liquidity which it can recycle. It is this possibility which may incentivise Bank A to play the 'risky' strategy offshore (i.e. pay early even in a bad state). If D pays early at period 2, then A will only default if it receives a large payment instruction at period 4 $s_{AD} = 2$, with probability p_A. If instead it receives a low payment instruction at period 4, $s_{AD} = 1$, it will have

[7] Whether uncertainty is resolved at the end of period 1 or the beginning of period 3 is not critical to game. It is however important that A has to decide whether to play early or delay at period 1 before uncertainty is resolved. Equally, if D knows the state, as I discuss in comparative statics, it follows that its decision at period 2 is required before uncertainty about O is resolved.

sufficient liquidity to make that payment, by recycling D's incoming payment at period 4: in this case, Bank A will avoid default.[8]

In summary, if Bank A pays early offshore at period 1 (playing a 'risky strategy'), it will then default at period 4 with probability γ; that probability is reduced if D makes an incoming payment at period 2, to γp_A. However, the probability does not go to zero: as we shall see, this means Bank D still faces a risk itself from paying early, since, by playing early, it cannot *guarantee* that A avoids default. Bank A's trade-offs at period 1 in the bad state are therefore as follows.

Table 4. Bank A's expected costs in Period 1

Pay early	Delay
$\gamma[(1-z)+zp_A]\beta$	d^f

Note: The probability the offshore problem persists, conditional on a bad state, is given by γ; z is the probability D pays early, p_A the probability A receives a large payment instruction at period 4 and β the cost to A of default. Delay costs in the offshore location are denoted by d^f.

Bank A's incentives to play this strategy are affected by its beliefs about the strategy to be chosen by D. If A believes D will play a morning strategy ($z = 1$), A has a chance to recycle liquidity, as long as it is unlikely to receive a high payment instruction at period 4 (p_A is low). If, by contrast, D delays payment, then A has no chance to recycle liquidity and therefore has no chance to avoid its own technical default, should the offshore problem persist. Clearly, D's strategy affects A′s incentive to take the risk.

What are Bank D's tradeoffs at period 2? If it delays, it will incur a delay cost, d^h. If it pays early, and A does not default at period 4, D will bear no (additional) costs; all collateral costs are sunk. However, if it pays early, and then A defaults at period 4, D will bear a cost of lost, or at least temporarily lost, liquidity. Bank D's costs and tradeoffs can therefore be modeled in the following table:

3. Equilibria

3.1. Baseline Case: Spillover Equilibrium. I am interested in describing a pure (type-contingent) strategy Bayesian-Nash equilibrium in which Bank A, the

[8]Notice that the 'risky' strategy in this model refers to A playing early at period 1 in the bad state. This stands in contrast to many policy discussions in which playing late is considered a relatively risky strategy (see discussion under 'Policy Implications').

TABLE 5. Bank D's expected costs in Period 2

Pay early (morning)	Delay
$\gamma q x p_A \pi$	d^h

Note: The probability the offshore problem persists, conditional on a bad state, is given by γ; q is the probability of a bad state, x the probability A pays early, p_A the probability A receives a large payment instruction at period 4 and π the cost to D of lost liquidity. Delay costs in the domestic location are denoted by d^h.

multinational bank, exploits the informational asymmetry. In particular, A plays the risky strategy offshore, *only because* Bank D is unaware of the risk, and so continues to play an early strategy. (In a later section, this will be compared with a game of full information, in which A is no longer able to extract this rent, and so does not play the risky strategy.)

I can model this as a simultaneous Bayesian game between A and D, in which A's strategy is its move at period 1, and D's strategy is its move at period 2.[9] Whereas A knows the true state, D does not; it believes the bad state has occurred with probability q. Both players choose between playing early (E) or late (L).[10] Payoffs (as the negative of costs) are represented in the matrices below, with each one corresponding to a particular state. Bank A is the column player, and Bank D is the row player. The top right payoff in each box corresponds to the payoff for A, whilst the bottom left payoff is that of D.

In presenting the results, strategies are ordered as follows: first, A's bad-state period 1 move; second, A's good-state period 1 move; third, D's period 2 move.

PROPOSITION 7. *In the incomplete information game, the pure strategies {early, early; early} are a Bayesian-Nash equilibrium if the following conditions are satisfied:*

Condition 1) $\gamma p_A \beta < d^f$,

Condition 2) $\gamma q p_A \pi < d^h$.

The proof is in the appendix.

In addition, I assume two further conditions hold:

[9]The game can be simplified in this way since a bank's period 2 move does not affect its opponent's period 2 cost.

[10]In the discussion, I use the words play and pay interchangeably to refer to the strategy chosen (early or late).

TABLE 6. Payoff Matrices

Bad State: q

	E	L
E	$-\gamma p_A \beta$ $-\gamma p_A \pi$	$-d^f$ 0
L	$-\gamma \beta$ $-d^h$	$-d^f$ $-d^h$

Good state: $(1-q)$

	E	L
E	0 0	$-d^f$ 0
L	0 $-d^h$	$-d^f$ $-d^h$

Note: Bank A is the column player and Bank D the row player. An early strategy is denoted by E and a late strategy by L. The probability of a bad state is q. The probability the offshore problem persists, conditional on a bad state, is given by γ; p_A is the probability A receives a large payment instruction at period 4, β the cost to A of default and π the cost to D of lost liquidity. Delay costs in the offshore location are denoted by d^f and in the domestic location by d^h.

$$\text{Condition 3) } \gamma\beta > d^f,$$

$$\text{Condition 4) } \gamma p_A \pi > d^h.$$

These ensure that the equilibrium is non-trivial (see discussion of these conditions in the appendix). They ensure that Bank A is exploiting the informational asymmetry. Condition 3 states that paying early is not a dominant strategy for A; if D were to pay late, A would prefer to pay late. In other words, A's strategy in equilibrium is determined by the strategy played by D, and thus by the informational asymmetry. Condition 4 states that Bank D's cost from playing early in a bad state, given A has also played a risky strategy, is sufficient to cause him to optimally delay - *if it knew the true state*. In this equilibrium, therefore, the domestic system, and Bank D in particular, are exposed to risk of spillover, *due* to the informational asymmetry. If conditions 3) and 4) hold, therefore, I refer to this equilibrium as the Spillover Equilibrium.

It is useful to explore Conditions 1) to 4). The details of the parameter restrictions are discussed at length in the appendix. However it is worth noting here an implication from the parameter restrictions:

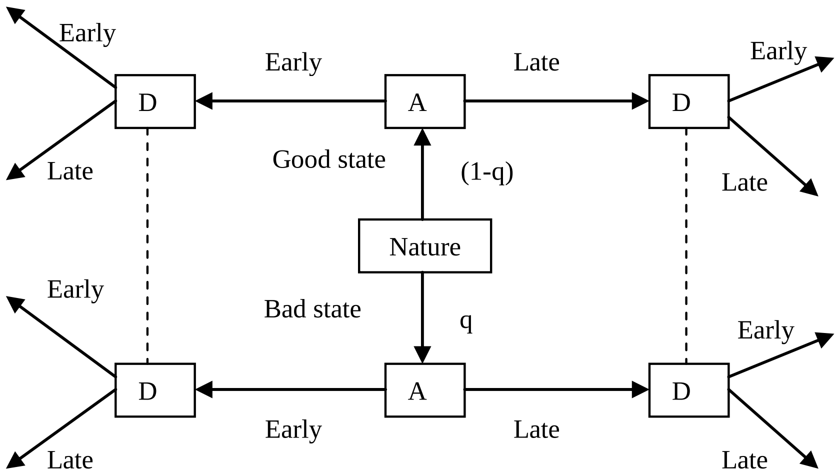
$$\frac{d^h}{\pi} \ll \frac{d^f}{\beta}$$

If default costs (β) are at least as large as costs of lost liquidity (π), then for this condition to hold, delay costs in the domestic location must be low relative to the offshore system. In other words, the risk of spillover is more likely to occur if delays costs in the unaffected system are low relative to those in the affected system. In terms of economic meaning, this could reflect a variety of things: if the offshore system has more linkages with other time-sensitive systems, or legal obligations are more tightly enforced, this would result in higher costs for banks who delay time-specific payments. Equally, the offshore delay cost may in fact be bank-specific: as a multi-system bank, A may face larger costs of delay than a smaller, more locally focused participant such as Bank D. As noted in the appendix, this could apply if A has customers who are highly connected within the financial system and face greater pressures to make timely payments.

3.2. Relaxing the information asymmetry: Comparative Statics. In the Spillover Equilibrium, there are two elements of imperfect information; D does not know the state of the world, nor does it know the strategy played by A in period 1. In effect, this makes the game equivalent to a standard Bayesian game (as if period 1 and period 2 are played simultaneously). In this section, I consider more versions of the game, with different degrees of information asymmetry. I still assume the 4 conditions above are satisfied.

I can consider a situation in which D will observe A's strategy but cannot observe the state of the world. This can be represented as a perfect Bayesian game. I construct the diagram below to mirror the standard way in which these games are displayed (see 'Beer-Quiche' game in Cho and Kreps (1987)).

Figure 3. Information Asymmetry: D knows A's strategy, but not the state



Note: The probability of a bad state is given by q.

Proposition 8. *If D knows A's strategy, but not the state, there is a pooling equilibrium in which A plays early in both states, and D will also play early.*[11]

The proof is in the appendix.

Proposition 9. *There is no separating equilibrium.*

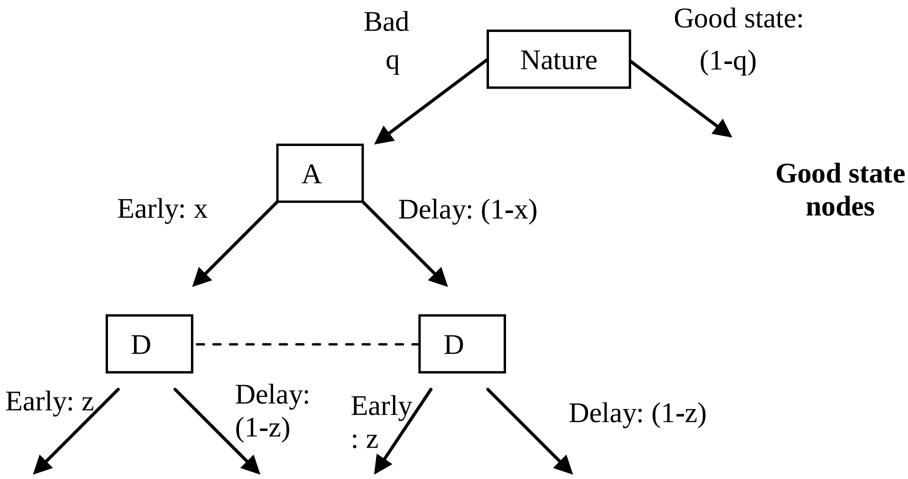
The proof is in the appendix.

In this version of the model, there is no reduction in the ex ante probability of spillover (given by $q\gamma p_A$), relative to the baseline case. It seems therefore that knowledge of the underlying operational problem offshore is crucial.

Alternatively, I could suppose D knows what state has occurred, but is unable to observe the move played by A at period 1. This is equivalent to a simultaneous, normal form game in each state. In this case, D's tradeoffs are the same as in the spillover equilibrium, but simply evaluated at either $q = 0$ or $q = 1$, for the good state and bad state respectively. In the good state, it is clear that paying early is a dominant strategy for D; there is no cost involved. However, in the bad state, the best strategy depends on D's belief as to A's move at period 1.

[11]As noted before, O will play early as a dominant strategy in the good state (and will have no strategy choice in the bad state). So we continue to consider a two player game.

FIGURE 4. Information Asymmetry: D knows the state but not A's strategy



Note: The probability of the bad state is denoted q, the probability A pays early is denoted x and the probability D pays early is denoted z.

PROPOSITION 10. *If D knows the state, but not A's move at period 1, there is no pure strategy Nash Equilibrium of the bad state game, if conditions 1) to 4) are satisfied.*

The proof is in the appendix.

Intuitively, if A plays the 'risky' strategy, by paying early, then D's best response is to delay, to avoid the expected cost of spillover. (Note I am assuming condition 4) holds). However, if D delays, A's best response is to delay (since I am assuming condition 2) holds). There is therefore no equilibrium in pure strategies. There will, of course, be a mixed strategy equilibrium in the bad state.

Proposition 11. *If D knows the state, but not A's move at period 1, the mixed strategy equilibrium of the bad state game is as follows (assuming conditions 1) to 4) are satisfied):*

A plays early with probability x and D plays early with probability z, where

$$x \equiv \frac{d^h}{\gamma p_A \pi}, \quad z \equiv \frac{\gamma\beta - d^f}{\gamma\beta(1 - p_A)}$$

.

The proof is in the appendix.

The probability that A plays a risky strategy is x. This is decreasing in the domestic delay costs, and increasing in the negative payoff D obtains from lost liquidity. A mixed-strategy equilibrium does not necessarily guarantee that the risk of default spillover to the domestic location is reduced. The ex ante probability of A defaulting in the domestic location is given by

$$q\gamma x[(1 - z) + z p_A] = \frac{q d^h d^f}{\gamma p_A \pi \beta}.$$

This is less than the probability in version 1 ($q\gamma p_A$) only if

$$(\gamma p_A)^2 > (\frac{d^h}{\pi})(\frac{d^f}{\beta}). \tag{3.1}$$

Whilst conditions 1) and 4) require

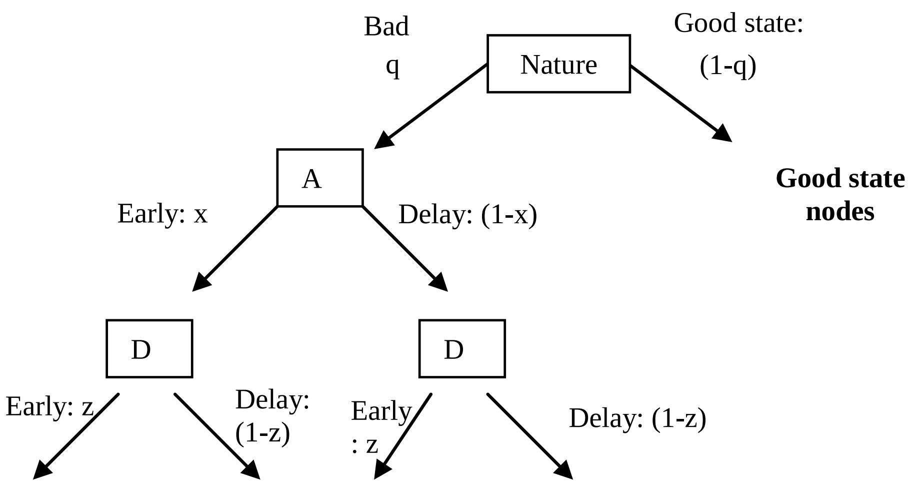
$$\frac{d^h}{\pi} < \gamma p_A < \frac{d^f}{\beta}, \tag{3.2}$$

there is nothing that requires condition (3.1) to hold. Notice that for sufficiently low domestic delay costs (d^h), or sufficiently high spillover costs in the domestic location (π), A will play the risky strategy with low probability and the spillover risk is reduced relative to version 1. Since relatively low domestic delay costs are necessary for the underlying spillover equilibrium (i.e. with condition (3.2)), it seems plausible that this mixed strategy equilibrium may reflect a lower probability of spillover.

I now consider the final version of the game, with no information asymmetry. In other words, D knows what state has occurred, *and* can observe A's move at period

1. This is now a standard sequential game, which can be solved through backwards induction.

FIGURE 5. Full Information Game



Note: The probability of the bad state is denoted q, the probability A pays early is denoted x and the probability D pays early is denoted z.

If Bank D knows the state, and can observe A's move at period 1, there are effectively two games: the good-state game and the bad-state game. Each is equivalent to a standard sequential game. In what follows, I focus exclusively on the bad-state game.

PROPOSITION 12. *In the full information game, (delay; early) is the subgame perfect equilibrium in the bad-state game. A will delay at period 1 and D will pay early at period 2.*

The proof is in the appendix.

In this case, the equilibrium in the bad state game is for A to play the safe strategy (delay at period 1) and for D to pay early at period 2. The risk of spillover to the domestic system has been eliminated. Intuitively, if D has information, not just about the state, but about A's move, it can react to A's choice of strategy. If A has played the risky strategy, D will choose to delay to ensure it loses no liquidity should A default (given condition 4). A knows D will react in this way; therefore A knows there is no possibility of using an early payment from D to recycle liquidity later on in the game. Therefore, A will choose to play the safe strategy (given condition 3).

3.3. Summary: Information Asymmetry. We see that the probability of spillover can change as we change the degree of information asymmetry. However the relationship is not necessarily monotonic. The four cases are summarized below:

TABLE 7. Probability of spillover in each state

No information (baseline)	Knows action but not state
$q\gamma p_A$	$q\gamma p_A$
Knows state but not action	Full information
$qd^h d^f/\gamma p_A \pi\beta \gtrless q\gamma p_A$	0

In the most extreme case, when the domestic-only bank (D) neither observes the state, nor A's actions, the ex ante probability of spillover is $q\gamma p_A$: in the other extreme case, when D observes both the state, and A's move, the risk of spillover is eliminated. However, in between these two extremes, the probability of spillover may increase, decrease, or remain the same. It remains the same if D observes A's move, but not the state (version 2). Yet, if D only observes the state (version 3), the resulting mixed strategy equilibrium can lead to an increase or a decrease in the probability of spillover, depending on parameter values. Specifically, the risk of spillover would be reduced if domestic delay costs were sufficiently low or the cost of spillover to domestic agents is sufficiently high (π).

In the appendix, I discuss necessary conditions on collateral. However, it is worth noting that endogenising collateral decisions would not change the basic mechanism highlighted in this paper, whether one considers the posting decision ex ante, or allows for extra posting intraday. These decisions would simply add a further complexity to the model, in which bank A in particular had to trade off liquidity costs against the expected cost of default. The underlying incentive, to rely on incoming funds by exploiting an information asymmetry, would still be present. This incentive is the key component of the model.

At this point, it worth reflecting on the contribution of this model to existing analysis in this area. Indeed, the motive to recycle incoming liquidity, and the disruptive nature of liquidity sinks, have both been analyzed in the literature before. What I have identified is the potential for spillover of shocks if some participants have more information than others about a specific shock.

For this reason, the mechanism in the model could also be applied to a large, single payment system. For instance, it would apply to a situation where each bank observes shocks to its direct counterparties, but not to its counterparties' counterparties. This would be relevant in a large system where counterparty connections between banks are not highly concentrated.

4. Policy Implications

This brings us on to consider the policy implications of the model. Given the central bank wishes to avoid the spillover effect, a key implication for policy is for the central bank to foster information flows across locations. Given that the domestic central bank requires information about the offshore system, it needs to foster frequent dialogue and contact with the offshore settlement agent. This may be easier to achieve if the offshore settlement agent is another central bank, or at least a foreign government institution, as opposed to a private institution.

However, a high degree of information must be shared in order for spillover risk to be reduced. As we saw from the above discussion, the best outcome is for domestic participants to learn *not just that a bad state has occurred*, but also to observe the actions played by other participants in other systems. In other words, the domestic participants would need to learn whether multinational banks had exposed themselves to risks offshore (by playing risky strategies i.e. paying early offshore despite a bad state). Notice it is not sufficient simply for participants to observe the strategies of other participants; as we saw above, the equilibrium outcome is not improved if D observes A's strategy but not the state.

If it is hard for such information flows to occur, what does the model imply for existing policy? Crucially, it suggests that a standard policy stance taken by central banks may actually increase the risk that problems in other systems spill over into their own system. Central banks generally prefer participants to make early payments, and some try to incentivise participants to do so. In so doing they increase delay costs, d_h, which were a key parameter above. The motivation comes from the risk of operational problems within the system later in the day, which could cause greater problems if banks have delayed payments until the last minute. However, in the framework of this model, if Bank D is incentivised to play early, this increases the exposure of the system to spillover. This is because it encourages dual-system participants to take risks. Without sufficient information flows, enabling domestic participants to react to new information in advance and thus change strategy, the explicit encouragement by the central bank to pay early may increase such incentives for dual-system participants.

Finally, it is worth considering the implications for policy in the context of spillover within a single system. If the system is sufficiently large such that each bank only knows of problems with its direct counterparties, and not other banks, there is room for the central bank to make public such problems. In this way, all banks can react to the state, even if they are not aware of how other banks are responding to the news. As we saw in the model, however, this may or may not reduce the potential for spillover.[12]

In recognising the difficulty of eliminating all information asymmetries, I now turn to consider an alternative possible tool available to policymakers: the implementation of a Liquidity Saving Mechanism.

5. Extension: Liquidity Saving Mechanisms

In this section, I extend the model to consider the impact of a Liquidity Saving Mechanism (LSM). Recall that if an LSM is in place in a given system, this alters the rules of the system: an agent is allowed to make a payment contingent on receiving a specific incoming payment within that system. Using the model above, we can see that a liquidity saving mechanism in the offshore system could reduce the probability of default in the domestic location.

It does not matter to the game exactly when uncertainty about the offshore problem is resolved, as long as A must make its decision beforehand. We can therefore reinterpret γ as the probability that O makes a payment by the end of period 1 in the bad state.

[12]An alternative, of course, would be to flood the system with liquidity and relax collateral constraints as the Federal Reserve did on September 11th, 2001.

If Bank A could not make its payment at period 1 contingent on any incoming payment offshore, the game would be as above. However, suppose Bank A could queue its morning payment to the offshore non-strategic agent; this would ensure that the payment will only be made if A receives a simultaneous incoming payment offshore. In this case, Bank A's expected cost from queuing the payment would be γd^f, whereas its expected cost from simply paying early (not-queuing) would be as before ($\gamma p_A \beta$). Notice that if Bank A queues the payment at period 1, there is no risk of default in the domestic location. Even if D plays an early strategy Bank A will prefer to queue rather than pay early as long as

$$\gamma p_A \beta > \gamma d^f$$

Note that this is a weaker condition than Condition 1). If the following parameter condition holds,

$$\gamma p_A \beta < d^f < p_A \beta$$

then an LSM in the offshore location would eliminate the possibility of default, where otherwise A would have played a risky strategy.

Could an LSM in the domestic location affect the equilibrium? In the current set-up, it would make no difference to the game; Bank A's strategy at period 2 makes no difference to the game, so there would be no benefit to making D's payment contingent on any simultaneous incoming payment. So a standard LSM, in which payments can only be made contingent on other payments in the same system, would not help in the domestic location. Moreover, even if it were possible, there would be no benefit to making D's payment contingent on A's strategy in the offshore location; as we saw above, the probability of A's default is not altered if D can observe A's strategy (but not the state).

I am therefore left with the interesting conclusion that the introduction of an LSM in one system can benefit another system, as long as there are multi system participants. In this sense, I have identified positive externalities from the introduction of an LSM in any given system. This suggests policymakers should coordinate across systems, and on an international basis, in encouraging regulatory moves towards LSMs.[13]

[13]Of course, this externality result only applies if we relate the model to two separate systems. If we consider this a model of spillover within a given system, as mentioned earlier, there is no externality to consider.

Unfortunately, however, the nature of the externality means that policymakers trying to reduce spillover in their own system, are dependent on the introduction of LSMs in other systems. They cannot merely solve the problem by introducing an LSM into their own system.

6. Conclusion

In this paper, I have addressed the question of what risks might be posed from one settlement system to another. I have focused on financial stability implications, and shown that informational asymmetries could play a significant role in affecting the degree of spillover from the offshore system. Such informational asymmetries would affect the strategic behavior of participants, both domestic and dual-participant banks. The latter might have an incentive to exploit such asymmetries, and thereby play a risky strategy in the offshore system.

In order for spillover to be reduced, significant information flows must take place. It may not be sufficient for domestic participants just to learn the state, or just the strategy of the other player.

In the absence of full information flows, this model shows that encouraging banks to make payments early may not be as riskless as it is often perceived to be. In the context of this model, dual-participant banks would have greater incentive to exploit asymmetries if they knew domestic banks were going to pay early.

I show that the presence of a liquidity saving mechanism in the offshore system may be a solution. It has positive externalities for the domestic system since it changes the behavior of the multi-system participant and so reduces the risk of spillover to the domestic system. These externalities have implications for regulatory debates concerning the introduction of liquidity saving mechanisms in a variety of systems. They imply that policymakers are somewhat reliant on international coordination in effectively reducing spillover between systems.

Finally, it is worth reflecting that the mechanism highlighted in the model could also apply to information asymmetry between participants *within* the same, large, payment system. Indeed, the key ingredient is that banks are aware of operational problems experienced by their counterparties, but not those of their counterparties'

counterparties. The implications of the model therefore follow through for the management of information flows between participants in the same location. Although the likelihood of information asymmetries is reduced if participants are located near to each other, there is nevertheless a small risk that spillover could occur in this manner. Information flows are therefore important, not just between systems, but within a domestic system itself.

Appendix 1: Proofs

Information Asymmetry Baseline: Spillover Equilibrium. In this baseline case, Bank D is unaware of both the state of the world, and Bank A's strategy at period 1.

Proof of Proposition 1:

These conditions ensure that each bank (A and D) is playing a best strategy, given the strategy played by the other. Condition 1 ensures that, if D plays the early strategy at period 2, then A will find that paying early in the bad state at period 1 is a best strategy. Condition 2 ensures that, if A played early at period 1, then paying early at period 2 will be a best strategy for D. In the good state, paying early is always a dominant strategy for A since there is no cost involved in doing so.

Information Asymmetry: D knows A's strategy, but not the state. *Proof of Proposition 2:*

If D plays early in equilibrium, A will play a risky (early) strategy at period 1, for the reasons discussed in version 1 (since $\gamma p_A \beta < d^f$). Bank D's belief about the state, on the equilibrium path, is given by $q(early) = q$. Since $\gamma q p_A \pi < d^h$, D will play an early strategy at period 2. For any off-equilibrium belief, $q(delay)$, Bank A will have no incentive to deviate. If Bank A delays, D will play early: to which A's optimal response is to play early.

Proof of Proposition 3:

Suppose there were a separating equilibrium in which A played early in the good state and delayed in the bad state. Bank D would then play early in both states. However this means A would have an incentive to deviate in the bad state, and play early. Alternatively, suppose there were a separating equilibrium in which A played early in the bad state and delayed in the good state. Then D would delay if it observed an early strategy and would play early if it observed delay by A. In this case, A would have an incentive to deviate in the bad state because it would face lower expected costs from delaying than from playing early (since $\gamma\beta > d^f$).

Information Asymmetry: D knows the state but not A's strategy. *Proof of Proposition 4:*

If A plays early, D's best response is to play late. If D plays late, A's best response is to play late. If A plays late, D's best response is to play early. There is no case in which both banks are playing a best (pure) strategy.

Proof of Proposition 5:

At these values of x and z, A and D are indifferent between playing early and late. They therefore randomize over both strategies, according to these probabilities.

Full Information Game. *Proof of Proposition 6:*

I solve this using backwards induction. If Bank A plays an early strategy at period 1, Bank D will pay late given Condition 4. If Bank A plays a late strategy, Bank D will play early, given $(0 < d^d)$. Bank A therefore faces expected cost $\gamma\beta$ from playing an early strategy, and cost d^f from playing a late strategy. Given Condition 3, Bank A will play a late strategy. Therefore, in equilibrium, A will play late at period 1 and D will play early at period 2.

Appendix 2: Robustness, Parameter Conditions and Payment Values

The game, as portrayed, occurs over a single day, across two locations with overlapping time zones. Yet each settlement period is distinct, and non-overlapping with any in the other location.

These assumptions are not crucial for the basic results, however. As mentioned earlier, the only vital element is that a liquidity sink occurs offshore that sufficiently reduces the liquidity available to the multinational bank; in a subsequent period in the domestic location, the multinational bank then defaults. This basic result, and similar (albeit not identical) strategic behaviour, could occur in a game with simultaneous settlement periods. That game would be more complex and may involve more settlement periods; however, the basic intuition would still apply.

In the next subsection, I give a description of necessary restrictions on collateral and payment values.

Payment Values. What are the necessary ingredients in this equilibrium? In addition to the parameter conditions described in the equilibrium conditions, there are four conditions necessary for payment and collateral values. In what follows, I denote maximum possible payment values by an upper bar, and minimum payment values by a lower bar.

a): $m_{AO} + m_{AD} + \overline{s_{AD}} > C_A$

This condition ensures that if A has to make a large payment at period 4, then it will default if it has previously made a payment to its offshore counterparty and not received any incoming payment.

b): $m_{AD} + \overline{s_{AD}} \leq C_A$

This condition ensures that A will have sufficient collateral if it only makes payments in the domestic location.

c): $m_{AO} + m_{AD} - m_{DA} + \underline{s_{AD}} < C_A$

This condition ensures that A *will not* default, even if it makes a payment offshore in period 1 (m_{AO}), as long as D pays early at period 2 (m_{DA}), *and* A only receives a small payment instruction at period 4 ($\underline{s_{AD}}$).

d): $m_{AO} + m_{AD} - m_{DA} + \overline{s_{AD}} > C_A$

This condition ensures that A *will* default if it makes a payment offshore in period 1 and then receives a large payment instruction at period 4 ($\overline{s_{AD}}$), even if D pays early at period 2. As long as payment values are non-negative, condition d) implies condition a).

Combining these conditions, we obtain the following conditions on relative payment values:

$$m_{AO} > 0, \tag{6.1}$$

$$\overline{s_{AD}} > \underline{s_{AD}}, \tag{6.2}$$

$$m_{AO} > m_{DA}. \tag{6.3}$$

The final one is non-trivial: the payment A makes to its offshore counterparty in the morning (period 1) must be strictly greater than the payment D makes to A in period 2. This ensures that D's incoming payment to A does not fully offset the loss of liquidity A faces by paying offshore.

The conditions on collateral can be simplified to:

$$m_{AO} + m_{AD} - m_{DA} + \overline{s_{AD}} > C_A^* > \max(m_{AD} + \overline{s_{AD}},\ m_{AO} + m_{AD} - m_{DA} + \underline{s_{AD}})$$

If $\overline{s_{AD}} - \underline{s_{AD}} > m_{AO} - m_{DA}$ then the lower bound condition simply requires that Bank A has sufficient collateral to meet the maximum possible debit position in the domestic location.

What condition ensures that A does not, ex ante, choose to hold more units of collateral than C_A^*? On the equilibrium path in version 1 of the game above, the probability that A faces a payment requirement of greater than C_A^* is $\gamma q p_A$. If I denote the ex ante per unit cost of collateral as λ, then A will choose to hold no more than C_A^* as long as

$$\gamma q p_A \beta < \lambda(m_{AO} + m_{AD} - m_{DA} + \overline{s_{AD}} - C_A^*)$$

Finally, we have the necessary and sufficient condition on C_D:[14]

$$m_{DA} + \overline{s_{DA}} \leq C_D$$

Since Bank A is free to play a late strategy, and delay the morning payment, Bank D must have sufficient collateral to make its afternoon payment (irrespective of incoming payments from A).

These characteristics of C_A and C_D do not seem unreasonable. Indeed, CHAPS participants frequently hold collateral in excess of actual liquidity usage within the system.[15] Nevertheless, it also seems likely that a bank will not hold sufficient collateral to cover the maximum of the sum of all its gross positions, across all locations. The

[14]These are necessary conditions as long as we allow for the possibility that Bank A may play a late strategy, in either the domestic or offshore location.

[15]It is not, however, known to the author whether any bank holds collateral equivalent to their maximum possible gross debit position.

likelihood of requiring such a large amount of liquidity is small; in part, the possibility of recycling incoming payments in any location is high. By contrast, the opportunity costs of holding such an amount of liquidity may be high. Indeed, the existence of global liquidity pools within banks indicates banks optimally choose to hold less than the maximum sum of all their global gross positions.

Parameter values. In this section, I discuss the four parameter conditions determining the equilibria. I can combine the Conditions 1) to 4) in the following condition:

$$qp_A < \frac{d^h}{\gamma\pi} < p_A < \frac{d^f}{\gamma\beta} < 1$$

If we add to this the condition for the LSM result to hold, we obtain

$$qp_A < \frac{d^h}{\pi} < \gamma p_A < \frac{d^f}{\beta} < p_A < 1$$

The probability of an operational problem offshore q is likely to be low: Klee (2010) estimates that the probability of an operational problem on a given day is between 5% and 18%, depending on the severity of the operational problem. Given that q is low, the first inequality is unlikely to be violated. Equally, delay costs are likely to be small relative to the cost of default ($d^f < \beta$), so even allowing for a low γ, the condition to the far right is likely to hold. The key remaining parameter restrictions can therefore be written as:

$$\frac{d^h}{\pi} < \gamma p_A < \frac{d^f}{\beta}$$

The middle term is the product of the probability of O's default, conditional on the bad state, γ, and the probability that A faces a large, rather than small, payment instruction at period 4. The latter parameter could be anything: in the analysis below I set it to 0.5. It is difficult to estimate γ without access to sensitive data: Klee (2010) documents that over the period 1998 to 2007, each day of payments had an 18% probability of an outage and a 14% probability of a Fedwire extension. An extension for settling beyond the normal end-of-day cut-off time implies the operational problem had had significant disruptive effects on the system: however, this does not necessarily indicate a default. It seems reasonable to suppose that most operational problems are resolved within the day, even if they lead to disruption, so that default only occurs in relatively few cases: so γ is likely to be small.

In terms of economic meaning, the key parameter restriction really boils down to

$$\frac{d^h}{\pi} \ll \frac{d^f}{\beta}$$

This is not implausible. Since A's default costs β are likely to be at least as large as D's costs of lost liquidity π, this means the condition reduces to

$$d^h \ll d^f$$

Intuitively, this makes sense. Bank A must have a sufficient disincentive to delay payments in the offshore location, in order to take the risky strategy of paying early. Nevertheless, Bank D would be prepared to pay the delay cost in the domestic location, if there were no information asymmetry and it knew they were in the bad state. So d^h must be low relative to d^f.

Whilst I have been referring to these delay costs as location-specific, as noted in an earlier section they could also be bank-specific. In which case, I require the multinational bank to face larger costs of delay than the domestic only participant. This may be plausible if the multinational bank is larger and thus serves larger, more connected customers.

Consider Table 8, which provides four numerical examples of parameter values consistent with the conditions and discussions above, as well as the condition for the LSM result.

As shown in the discussion above, it is largely the ratios of parameters, rather than their absolute values, which matter. The two probabilities p_A and γ have a close relationship in the above condition. If one changes the other must compensate. Notice also, that these conditions can be satisfied even for fairly high values of q, the probability of the bad state (even though this seems unlikely to be high in practice). It should be acknowledged that, in all these examples, the ex ante probability that A defaults (given by $q\gamma p_A$ in Version 1) is relatively small.

Table 8. Parameters and Values

Parameters	Values				
q	0.05	0.05	0.05	0.05	0.8
γ	0.5	0.2	0.7	0.3	0.7
p_A	0.3	0.7	0.2	0.5	0.2
β	30	30	30	30	30
d^h	3	3	3	1	3
d^f	5	5	5	5	5
π	25	25	25	30	25
Ex ante prob. A defaults	0.0075	0.007	0.007	0.0075	0.112

Note: The probability of the bad state is given by q; the probability the offshore problem persists, conditional on a bad state, is given by γ; p_A is the probability A receives a large payment instruction at period 4; β is the cost to A of default and π the cost to D of lost liquidity. Delay costs in the offshore location are denoted by d^f and in the domestic location by d^h.

Appendix 3: Delay costs

Payment instructions, as highlighted in the main text, are independent events, each sent by different customers. As a result, we should think of a delay cost, the cost of failing to meet a legal deadline, as being associated with the payment itself. As such, the bank cannot avoid a delay cost if it turns out that its counterparty cannot make outgoing payments; whilst the liquidity may be offsetting for the banks, the payment instructions, as independent events, do not offset.

However, we may want to motivate the model in a different way and imagine the story is instead about banks making interdependent offsetting payments to each other. This would be relevant if the payments were not from external customers of banks, but rather related to exposures set up from lending and borrowing between banks themselves. In this case, it would be sensible to alter our assumptions about delay costs.

Notice that if we alter our assumptions about delay costs, the role of information asymmetry goes away; the risk of spillover is driven to a greater degree by the relative size of domestic delay costs and the cost of lost liquidity. If delay costs were contingent on a bank's counterparty surviving, then A would only pay d^f if he played late *and* O avoided default (similarly D would only pay d^h if he played late *and* A avoided default). In this case, conditions 1)-4) would no longer be a function of γ and, after cancelling terms, would reduce to the following:

$$\text{Condition 1) } p_A\beta < d^f,$$
$$\text{Condition 2) } \pi < d^h,$$
$$\text{Condition 3) } \beta > d^f,$$
$$\text{Condition 4) } \pi > d^h.$$

Notice that conditions 2) and 4) cannot both hold. D is less likely to face a delay cost, now it is contingent; however, the fact he has the same probability of paying the lost liquidity cost as the delay cost means the relative size of these two becomes more important. If condition 2) is satisfied rather than 4), the probability of spillover remains the same but D would not wish to change his strategy and play late if he knew A had played early in the bad state. This is because the cost of delay is high relative to the cost of lost liquidity.

The policy implication here is more extreme: since the spillover not longer depends on information asymmetry, it is crucial for policymakers to ensure low delay costs. This would ensure condition 4) was satisfied instead of 2), so multi-system participants had no incentive to play a risky strategy. Although the equilibrium is therefore different, the policy implication is similar to the model in the main text: contrary to the standard policy of encouraging early payments, policymakers should ensure low delay costs to avoid spillover in this scenario.

CHAPTER 3

Consumer Credit and Payment Cards

1. Introduction

Debit or credit? Every day, millions of consumers stand at store checkout counters and make a payment decision: whether to pay by debit or by credit card. Since the retail price at the checkout is generally the same either way, this decision looks pointless. It is not. Financial incentives, merchants' interests, and available credit facilities do play an important role for consumer payment choice. Moreover, behind the scenes, billions of dollars are at stake.

In this paper, we study the pricing of payment cards. Since payment card networks are two-sided markets, we consider the optimal fees charged by the network to the consumer and to the merchant. Unlike most payment models, where consumer credit is not considered, our model is among the first to analyze payment network fees and competition by explicitly incorporating the different ways consumer credit is offered in debit and credit card networks. Specifically, we consider overdraft facilities and credit lines.

As is standard in the literature, we assume that part of the value of payment cards to consumers comes from the reduced need to hold cash. Specifically, both payment cards provide additional security over cash. They also enable merchants to avoid the cost of cash handling. We also consider the way in which cards help liquidity constrained consumers. If a payment card offers payment possibilities in extra states of the world, this is valuable to both consumers and merchants. The card fees, set by the payment network, then depend in part on the degree to which the network can extract surplus from consumers and merchants. The optimal combination of merchant fee and consumer fee is however determined by the need for the payment network to 'internalize' the network externalities from either side of the market (see Rochet and Tirole (2006) for discussion of two sided markets).

Debit and credit cards offer distinctly different credit possibilities for the consumer. A debit card enables its holders to make purchases and have these transactions directly and immediately charged to their current accounts. The consumer can access credit via her debit card as long as she has an overdraft facility on her current account. Typically, such credit faces immediate interest charges. By contrast, a credit card enables cardholders to make purchases up to a prearranged credit limit. Such credit is interest free for a limited 'grace' period, beyond which the consumer faces interest charges on any remaining negative balances. In short, debit and credit card networks operate different business models for supplying credit.

We show how the different models of credit affect equilibrium merchant and consumer fees, as well as the nature of competition between the two payment networks. We explore two cases; the first in which credit and debit card networks set merchant and consumer fees monopolistically, and the second in which the two networks compete for custom. First, when the credit card network behaves monopolistically, we find that default risk and funding cost are partly passed onto the merchant through the credit card merchant fee. Yet, debit card merchant fees do not share this feature in a debit-only world, as long as the only alternative to debit is cash. Debit merchant fees only depend on default risk and funding costs when we introduce competition with credit cards. In that case, however, we find that debit merchant acceptance actually increases with the probability of default, despite an increase in the merchant fee. This is because the credit card merchant fee responds more to the higher default risk, causing some merchants to switch from credit cards to debit cards.

Second, we find that monopolistic credit card fees also depend on overdraft interest rates, even though these are associated with the current account and therefore completely separate from the credit card network. The overdraft is an outside option for the consumer in one state of the world, so at the margin the credit line allows the consumer to 'save' on the costs of servicing the overdraft.

Third, when we turn to consider competition between a debit and a credit card model, we find that a degree of complementarity exists between debit and credit cards. Greater credit card acceptance increases profit for the bank that issues the debit card as the consumer can maintain a positive balance whilst using the 'grace' period on the credit card to make purchases. Competition drives payment fees down, but the complementarity results remains. As a result, the debit card bank may increase merchant fees at the margin in order to decrease debit card acceptance in favour of credit cards; this means debit card merchant fees may approach monopolistic levels. When we consider welfare maximising fees, we find that the competitive debit merchant fee is indeed high relative to the welfare-maximising case. Whilst that has also been observed in other papers (Bolt and Schmiedel (2011)), we have identified an extra wedge between the competitive and the socially optimal debit fee: this comes from the complementarity between the two business models.

The pricing of credit and debit cards is of particular relevance to policymakers and regulators. Policymakers have focused on the level of interchange fees, paid by the acquiring bank to the issuing bank (see below for more details). Following discussions with the European Commission, Mastercard has recently agreed to reduce interchange fees on cross-border European card transactions. Similarly Visa Europe has agreed to reduce such fees for cross-border debit card payments. US policymakers have also proposed setting a cap on interchange fees for debit and prepaid cards.

In addition, our results may be relevant for the realization of the Single Euro Payments Area (SEPA). The broad aim of the SEPA project is to enable closer European financial integration, through enhancing harmonization in the means of payment, treating all payments in the euro area as domestic payments. With respect to payment cards, the SEPA framework has focused on the need to increase competition and efficiency between card networks. Our paper attempts to shed new light on what competition between debit and credit cards and access to funds imply for optimal payment pricing of payment cards.

Our paper can be seen in the context of existing literature both on payment cards and in the field of consumer finance. Various papers including Baxter (1983), Rochet and Tirole (2002), Rochet and Tirole (2003a), Wright (2003, 2004) have analysed payment cards and two sided markets, focusing on the optimal combination of the consumer and merchant fees. They highlight the fact that neither side of the market takes into account the positive externality to the other side from one's own participation in the network. As a result, the network must find the optimal combination of fees to effectively ensure these externalities are internalized. For instance, there is no point setting an extremely high consumer fee and a low merchant fee, if this means no consumer will participate: in that case, the card is relatively worthless to merchants, irrespective of the fact they are paying a low fee. In a market where the network consists of two banks (the acquiring bank on the merchant's side and the issuing bank on the consumer's side), an interchange fee may be necessary to effectively enable one side to subsidize the other (see Rochet and Tirole (2006) for more details).

So far, no paper has explicitly studied the impact of overdraft facilities and access to credit on the pricing decisions for card payment networks. Chakravorti and To (2007) introduce a credit line into their model of credit cards, but do not consider periods beyond the 'grace' period and therefore do not consider the relevant interest charge for credit. Moreover, their paper lacks an analysis of competition between credit and debit cards. Our paper builds on the modelling framework of Bolt and Chakravorti (2008) and, in particular, Bolt and Schmiedel (2011), but extends that work to consider consumer credit. In so doing, we attempt to bridge the gap between the payment card literature and that of consumer finance.

We structure the paper as follows. In section 2, we present the model, while in sections 3 and 4 we consider the optimal prices in a world just with a debit card, and subsequently a world just with a credit card. We refer to these as the monopolistic pricing models. In section 5, we consider optimal prices in a world where the credit and debit cards compete with each other. In Section 6 we discuss welfare implications, while in Section 7 we discuss possible extensions and robustness. We conclude in Section 8.

2. Model

The basic model closely follows that of Bolt and Schmiedel (2011) (and also Bolt and Chakravorti (2008)). In our model, there are three types of agents: consumers, merchants and payment network providers. All agents are risk neutral. Banks are considered to play the role of payment network providers. As in Bolt and Schmiedel (2011), we use a three party network. In other words, we combine the issuing bank and the acquiring bank into a single network provider. This enables us to focus simply on the merchant and consumer fees, without also having to solve for an interchange fee.[1] There are two periods in the model, period 1 and period 2 (which we will respectively refer to as 'day' and 'night').

2.1. Consumers. Consumers are homogeneous and maximise linear utility. They obtain utility $v > 0$ from consuming a single, indivisible good, which they would purchase from a merchant. Each consumer is matched with a single merchant near the beginning of period 1, after deciding whether to subscribe to a particular payment card. If she is able to make a purchase, she pays price p and therefore obtains (net) utility from consumption equal to $v_0 = v - p$ where $v_0 \geq 0$. For simplicity we normalize $p = 1$. However, there is no guarantee that she will make a successful purchase at a given merchant.

The first potential friction in making a payment comes from liquidity. The consumer may or may not receive positive initial income. As a result, she may have insufficient funds available (through a combination of initial income and credit) to make the purchase. We discuss below the specific nature of her income shocks and available credit.

Second, the merchant may or may not accept the card to which the consumer has subscribed. If not, the consumer must rely on cash to make the payment, which itself faces a cost.[2] We model the cost of cash in a reduced form way, by assuming she will be mugged with positive probability, $1 - \rho$, on her way to make the purchase; in that case, she will be unable to purchase and consume the good. This follows other papers which associate the cost of cash with theft, such as He et al. (2005).

During period 2, night time, the consumer receives a second income shock. Income may arrive early at the beginning of period 2, or late at the end of period 2, or not at all. The only value of receiving period 2 income comes from the ability to pay back

[1]This follows the same approach as in Bolt and Schmiedel (2011). As an alternative to the three party network with merchants, consumers and the network provider, one can also consider a four party network. This would explicitly model the acquiring bank and the issuing bank, instead of combining the two into a single network provider. However, as Bolt (2006) discusses, the two models are equivalent if either the issuing bank or the acquiring bank are perfectly competitive. If this is the case, there is a close linear relationship between the optimal interchange fee, paid between the acquiring and issuing bank, and the consumer and merchant fees.

[2]We ignore potential benefits of cash usage.

Table 1. Income streams: timing and shocks

Probabilities		Income			Total Income Received
Period 1	Period 2	1	2_{early}	2_{late}	
δ	γ_E	x_1	x_2		$x_1 + x_2$
	γ_L	x_1		x_2	$x_1 + x_2$
	$1 - \gamma_E - \gamma_L$	x_1			x_1
$1 - \delta$	γ_E	0	x_2		x_2
	γ_L	0		x_2	x_2
	$1 - \gamma_E - \gamma_L$	0			0

debt obligations from period 1. Before describing the credit options offered by each network, we consider the specific income shocks in more detail.

2.2. Income shocks and default. Period 1 income is given by $x_1 > 0$ and period-2 income by $x_2 > 0$. We assume that period-1 income is insufficient to cover the purchase, whilst period 2 income is greater than the price of the good. In other words,

$$x_1 < 1 < x_2.$$

At the beginning of period 1, the probability the consumer receives income x_1 is given by δ; otherwise she receives zero. In period 2, the probability she receives income early is given by γ_E and the probability she receives income late is given by γ_L. With the remaining probability she receives no income in period 2: $1 - \gamma_E - \gamma_L$. Note that the probability she receives income in period 2 is completely independent of the period-1 income shock.

Given the independence between period-1 and period-2 income shocks, there are six possible outcomes in the game as a whole. Table 1 summarizes these outcomes. In the table, we consider the total amount of income received by the end of period 2, gross of any outgoing payments. The upper panel depicts the case of positive period-1 income shock, the lower panel a zero period-1 income shock.

Regardless of period-1 income, the consumer must use credit for the purchase (since $x_1 < 1$). From Table 1, the consumer will default in two states, conditional on having purchased the good. Therefore, the ex ante probability of default, conditional on the consumer making a purchase, is given by $1 - \gamma_E - \gamma_L$.

Given the probabilities and income shocks described above, ex ante expected income is equal to:

$$E(I) = \delta x_1 + (\gamma_E + \gamma_L)x_2.$$

Since we assume consumers are ex ante solvent, this implies that $E(I) > 1$, or rearranging

$$1 - \delta x_1 < (\gamma_E + \gamma_L)x_2.$$

There are two distinct differences between credit and debit cards in our model. Both relate to the nature of credit offered in association with the two systems. Firstly, we assume that the consumer always has access to an overdraft associated with her current account, while the credit card offers a credit line to the consumer. If she holds a debit card, the consumer can use her overdraft facility to make payments via this card. Whilst this debt will immediately accrue interest charges, the credit line of the credit card offers the consumer a free 'grace' period. In effect, the credit line associated with the credit card will not accrue interest charges until after the first period.

Secondly, we assume the credit line is larger than the overdraft facility, thus enabling the consumer to make payments in more states of the world. Specifically, we assume the overdraft limit is sufficient to cover the purchase if the consumer received period-1 income, but insufficient if there was no income received; as a result she will be unable to purchase the good. By contrast, the credit limit on the credit card is sufficient to cover the purchase, even if the consumer received no period-1 income. This captures the fact that both credit and debit cards are used alongside credit facilities, but the credit card enables payment in extra states, relative to the debit card.

2.3. Merchants. We assume merchants are heterogeneous. Specifically, we assume merchant i receives profit $\pi(i)$ from a sale, where $\pi(i)$ is uniformly distributed between 0 and p (where $p = 1$). As discussed in Bolt and Schmiedel (2011), this captures the idea that merchants vary in their profit margins due to differing production costs. If the merchant accepts a payment card and the consumer uses this to make a purchase, the merchant must pay a per-transaction merchant fee (f). Alternatively, if the consumer uses cash, the merchant will face a per transaction cost of cash handling (h). As is standard in the literature, we assume the no-surcharge rule holds:

in other words, merchants are prohibited from setting a different price to cash-paying consumers as opposed to card-paying consumers.

2.4. Payment networks. We consider two different payment networks: a credit card network and a debit card network. Each network chooses per-transaction merchant fees $f_j \geq 0$, $j = C, D$ and consumer fixed fees F_j, $j = C, D$ to maximise profits. This mirrors what occurs in many countries; consumers generally pay fixed fees while merchants pay per-transaction fees.

We assume the network is a monopolist on the consumer side: in other words, the network will extract the maximum fixed fee that the consumer is willing to pay, for a given proportion of merchant acceptance. Following that, the network will choose the merchant fee to maximise network profits.

The profits are a function of the fee revenues as well as the payment network costs. First, they bear a per-transaction processing cost $c_j \geq 0$, $j = C, D$. Second, they bear the expected costs of offering credit and bearing consumer default. It is to this issue we now turn in considering the pricing of credit.

2.5. Interest rates. As described above, regardless of period-1 income, the consumer must use credit to purchase the good. In this section, we consider the costs of credit. We assume the overdraft limit on the debit card (and by extension if the consumer pays by cash) is sufficient to cover the purchase if the consumer received period-1 income, but insufficient if there was no income received. If the consumer uses her overdraft, then she will be in debt by an amount $m_d \equiv 1 - x_1$. This debt will accrue interest at rate r_d from period 1, until she repays using period-2 income.

By contrast, we assume the credit limit on the credit card is sufficient to cover the purchase, even if the consumer received no period-1 income. If the consumer uses her credit card, she will not face any interest accrual for period 1. This is known as the 'grace' period. However, if she is unable to repay the debt using period-2 early income, she will face interest charges over period 2 at rate r_c. As a result, she will not use any initial income to partially repay the debt until the beginning of period 2, since she doesn't face interest charges until that period.[3] Her expected debt in period

[3]For simplicity, we assume she earns no interest in her current account: in which case, she is indifferent between keeping her initial income in her account or using it to pay off some of her credit card debt

1 will be equal to $p = 1$, regardless of early income. Her expected debt in period 2, on which she accrues interest charges, amounts to $m_c \equiv 1 - \delta x_1$ (note that $m_c > m_d$).

For the bank providing the overdraft, the expected cost of credit (including default) is

$$ECo = m_d[r + (1 - \gamma_E)r + (1 - \gamma_E - \gamma_L)],$$

while the expected revenues, for given r_d, are

$$ER_O = r_d(\gamma_E + 2\gamma_L)m_d.$$

For the credit card provider, the expected cost (including default) is given by

$$EC_C = r + m_c[(1 - \gamma_E)r + (1 - \gamma_E - \gamma_L)],$$

while the expected revenues, for given r_c, are

$$ER_O = r_c \gamma_L m_c.$$

Notice that there is no term for γ_E in the credit card provider's expected revenues because the first period credit will be free for the consumer.[4]

It turns out our results for merchant fees and acceptance rates do not depend on the specific assumption we make regarding the way interest rates are set on credit lines (r_c). Moreover, most of our *key* results do not depend on the specific value of r_d, the overdraft interest rate. Intuitively, this is because the network is a monopolist on the consumer side, and effectively engaging in multi-part pricing, comprising of the fixed fee and the interest rate. If the network sets the consumer interest rate equal to the zero profit level (i.e. $ER_j(r_j) = EC_j$ for $j = C, D$), then the network will just extract the full consumer surplus through the fixed fee F_j, which we consider in a later section. If the network alternatively sets a lower interest rate and makes a loss on the credit portion, it will extract the difference through the higher fixed fee which the consumer is willing to pay. In both cases, the merchant fee will be the same, since this

in period 1. However, we could reasonably suppose she earns ε interest in period 1; this would make her strictly prefer to maintain the balance without changing the results.

[4]As discussed later, we assume that the overdraft cannot be used to pay off some of the credit card debt. Notice that if it could, then the bank supplying the overdraft would take on some of the consumer default risk otherwise borne by the credit card provider.

is the result of network profit maximisation, taking into account that the maximum consumer fee will depend on the merchant fee: $F_j(f_j)$ for $j = C, D$.

We therefore proceed in the following way. We assume that the overdraft credit is set as a zero Net Present Value (NPV) loan. One way to motivate this is to think of the credit part as a competitive 'aftermarket', implying consumers could substitute other loans such as store credit for overdrafts. It is an 'aftermarket' in the sense that it is determined separately from the market in which card fees are determined. This assumption makes no difference to the debit merchant fee, for the reasons given above. However, as we will see, the overdraft interest rate will affect the credit card merchant fee. For simulation purposes, it is therefore helpful to have the assumption of zero NPV in order to relate r_d to underlying parameters.[5]

The per period simple interest rate r_d must therefore solve:

$$r_d(\gamma_E + 2\gamma_L) = r + (1 - \gamma_E)r + (1 - \gamma_E - \gamma_L).$$

This gives r_d as follows:

$$r_d = r_d(r, \gamma_E, \gamma_L) = \frac{(2 - \gamma_E)r + (1 - \gamma_E - \gamma_L)}{\gamma_E + 2\gamma_L}. \tag{2.1}$$

We can easily show that equilibrium r_d decreases with γ_E and γ_L and increases with r.

In order to pin down r_c, we assume the credit line of the credit card competes with the same outside credit options as the overdraft, in the states with positive period 1 income when the consumer could use either cash or credit card. We also assume that the network cannot make r_c contingent on the state: so r_c is entirely pinned down by the outside credit options.[6] This is entirely without loss of generality, given the network extracts the full consumer surplus. Moreover, credit line interest rates will

[5]While we do not think the zero NPV assumption is necessarily an accurate description of reality, we feel it is a helpful baseline to consider. We model interest revenue as simple interest so that the lender receives revenue of $2r$ if the capital is left untouched over two periods. This keeps the notation simpler without changing the qualitative results.

[6]Since we assume the credit card adds value by enabling the consumer to make a purchase in one extra state of the world, it must be that none of the outside credit alternatives are sufficiently large to help the consumer when there is no period 1 income. Nevertheless, if the credit card network cannot charge different interest rates for different marginal units of credit, and if r_c cannot be contingent on the state, the outside credit options will still pin down r_c.

have no affect on the debit network's optimisation, so the specific way r_c is set makes no difference to our results.

From this assumption, the expected costs to the consumer must be the same for the credit line as the overdraft, conditional on having high initial income.

$$r_c \gamma_L = r_d(\gamma_E + 2\gamma_L).$$

This gives r_c as follows:

$$r_c = r_c(r, \gamma_E, \gamma_L) = \frac{(2-\gamma_E)r + (1-\gamma_E-\gamma_L)}{\gamma_L}. \tag{2.2}$$

If r_c is set as in this condition, the credit card network makes a loss on the credit in expectation. Effectively, it is subsidising the consumer in period 1 in the credit aftermarket by allowing the consumer to have a debt equal to 1 but to only pay interest in period 2 on the remaining portion of her debt $1 - x_1$. To see that the credit card company makes a loss in this aftermarket, note that:

$$\begin{aligned} ER_C &= r_c \gamma_L m_c \\ &= m_c[r + (1-\gamma_E)r + (1-\gamma_E-\gamma_L)] < EC_C. \end{aligned}$$

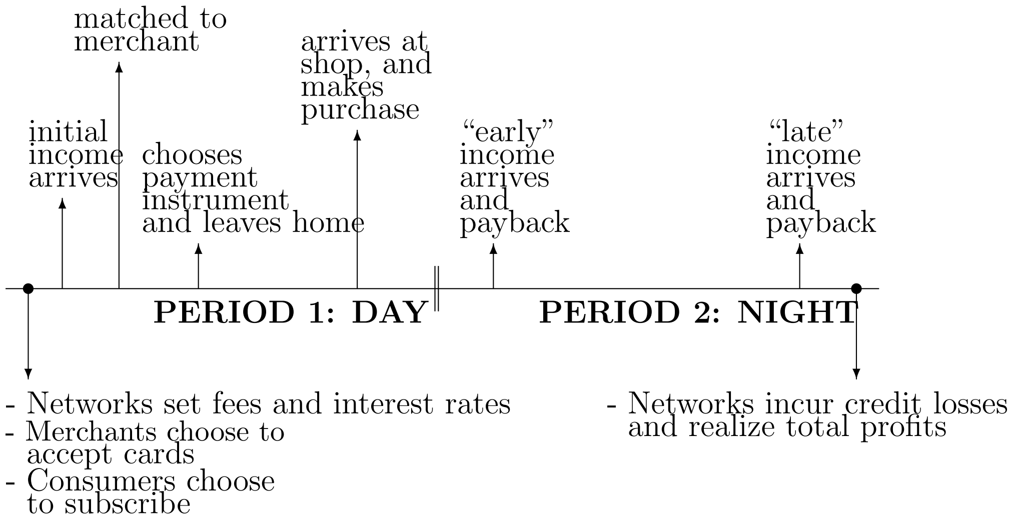
The loss will be captured by a cost term in the credit card provider's profit function. This loss is equal to

$$\begin{aligned} EC_C - ER_C &= r(1-m_c) \\ &= r\delta x_1. \end{aligned}$$

This directly shows the subsidy provided by the network given the consumer never has to pay interest on the x_1 part of her debt in period 1. The key point, however, is that our results do not change, whether the consumer is made to directly pay for the subsidy in higher interest rates, or whether the credit card provider bears the cost in the profit function. This is because the credit card provider extracts the full surplus from the consumer, so overall, when we account for the card fees, there is no subsidy.[7]

[7]Note that r_c explodes when γ_L approaches zero. With credit cards, consumers that receive late income carry all the funding and default cost. When γ_L is small, only a few consumers carry this burden and so pay very high interest rates. In the extreme, if $\gamma_L = 0$, no consumer pays interest on

Figure 1. Timing of events



2.6. Timeline. Figure 1 shows the timeline. In period 1, consumers choose whether to accept a card; they then get matched to a merchant and, hopefully, make a purchase using either cash (if they have not been mugged) or a payment card, if the merchant accepts it. In period 2, consumers either receive income (early or late) and repay any debts, or they default at the end of the game. Please note that the purchase and all previous actions take place near the beginning of period 1, so the consumer will be in debt for most of that period (possibly accruing interest). For artistic purposes, however, we have placed this near the end of the period.

It is also helpful at this stage to summarise the parameters in the model:

3. Debit Card Only Model

In this section, the consumer can either rely solely on cash to make a purchase, or decide to hold a debit card. The overdraft facility works similarly for cash as for the debit card. Therefore, the only benefit to holding a debit card comes from the risk of getting mugged, and losing cash.

its credit card loan (they receive grace or they default) and therefore the loan cannot be made NPV-zero. In this case, to recover cost, the burden must be shifted to merchants and consumers through higher payment fees. To avoid this exploding characteristic, we will mainly focus on distributions (γ_E, γ_L) that are not too 'skewed'. We ignore direct burden sharing mechanisms between merchants and credit card debtors. Naturally, the parameters that determine the interest rates will influence optimal merchant fees.

Table 2. Parameter List

Parameters	Meaning
v_0	net utility from consuming the good
$p = 1$	normalised price of the good
x_1	period 1 income
x_2	period 2 income
δ	probability of period 1 income
γ_E	probability of early period 2 income
γ_L	probability of late period 2 income
π_i	merchant i's profit from a sale
f_j	per transaction merchant fee where $j = C, D$
F_j	consumer fixed fee where $j = C, D$
c_j	per transaction network processing fee where $j = C, D$
m_d	overdraft debt
r_d	overdraft interest rate
m_c	credit line debt
r_c	credit line interest rate

3.1. Consumer's problem. The probability of getting mugged is $(1 - \rho)$. We denote by α^D the proportion of merchants who accept the debit card and F^D is the consumer's debit card fee. Recall that if debit cards and cash are the only payment instruments available to the consumer, she can only make a purchase if she receives a positive level of initial income; this occurs with probability δ. With probability $(1 - \delta)$, there is no purchase, no consumption, but also no default loss. Observe that $m_d \equiv 1 - x_1$ denotes the amount of debt when using the overdraft facility associated with the checking account.

The consumer will want to hold a debit card as long as:

$$\delta\rho v_0 - \delta(\gamma_E + 2\gamma_L) r_d m_d \ \leq \ \delta\left(\alpha^D + \rho(1 - \alpha^D)\right) v_o - \delta(\gamma_E + 2\gamma_L) r_d m_d - F^D.$$

The left-hand side is the payoff from just holding cash; in this case, the consumer can purchase the good only if she receives high initial income and is not mugged. If she makes a payment (which occurs with probability $\delta\rho$), then she will have to pay interest on her overdraft of size m_d. Note, however, if she gets mugged she will still have gone into her overdraft, having withdrawn 1, and thus will have to pay interest. In other words, we assume the consumer has no insurance against cash theft.

She will only have to pay interest in one period, if she receives an early second income shock (which occurs with probability γ_E). However, if she has to wait for positive income until period 2, she will have to pay twice the amount of interest; this occurs with probability γ_L.

On the right-hand side is the payoff from holding a debit card. The consumer can make a purchase with a debit card if she receives high initial income, and the merchant accepts the card. She can also rely on cash for the payment if the merchant does not accept the card (with probability $(1-\alpha^D)$), providing she is not mugged. Either way, the consumer must pay the debit card fee F^D.[8]

We continue to make the same assumptions about mugging. If a consumer is aware that she cannot pay by debit card, she will withdraw cash equal to 1. At this point, she faces a risk of being mugged, in which case she loses the money, and thus must pay interest on the overdraft until she can repay. The participation constraint can be simplified as follows:

$$F^D \leq \delta\alpha^D(1-\rho)v_0.$$

Note that the debit card allows the consumer to pay in one extra state, which occurs with probability $\delta\alpha^D(1-\rho)$. For this reason, the surplus from buying the good v_0, is multiplied by this term.

3.2. Merchant's problem. The merchant i receives profit $\pi(i)$ from a sale, where $\pi(i)$ is uniformly distributed between 0 and 1. His cost of handling cash is h whilst f^D is the merchant fee for accepting the debit card. His expected payoff from accepting cash is:[9]

$$Z_{cash}(i) = \delta\rho[\pi(i) - h],$$

and his expected payoff from accepting the debit card is:

$$Z_D(i) = \delta[\pi(i) - f^D].$$

[8]There is a chance that period-2 income does not arrive at all so that the consumer cannot pay for the fixed fee. Hence, the default probability 'artificially' increases the consumer willingness-to-pay for the card. However, the payment network would discount the high fixed fee with the same probability. Mathematically, this effect cancels out.

[9]Given π can be as low as zero, we assume the outside option for the merchant instead of enabling the purchase is given by $-h$. This normalization ensures that all merchants would prefer to accept cash than reject the purchase.

Merchants accept debit cards only when

$$Z_{cash}(i) \leq Z_D(i).$$

Since there is a level of profits $\bar{\pi}$ above which merchants will accept debit cards, we can write the proportion of accepting merchants as follows:

$$\alpha^D(f^D) \;=\; \Pr[\pi(i) \geq \bar{\pi}] \;=\; 1 - \bar{\pi} \;=\; 1 - \frac{(f^D - \rho h)}{1-\rho}.$$

3.3. Maximum consumer fee for debit cards. Using the function $\alpha^D(f^D)$, we can derive the maximum possible consumer fee as a function of f^D. This is obtained by finding the fee such that the consumer is indifferent between holding a debit card or solely relying on cash. It is given by:

$$F^D_{\max}(f^D) = \delta\left(1 - f^D - \rho(1-h)\right)v_0.$$

3.4. Debit card network. We make the standard assumption that the same bank operating the debit card network is the one to provide the consumer with a current account and associated overdraft facility. The Debit Card Bank (DCB) faces processing cost c^D per debit card transaction. The DCB is also able to earn interest on a positive balance in the customer account; we assume the bank takes this interest rate r as given. In addition, the bank charges interest rate r_d on any overdraft.

The DCB's payoff from issuing a debit card is:

$$\pi^{DCB} = F^D + \delta\alpha^D(f^D - c^D) + r[(1-\delta)\gamma_E x_2 + \delta\gamma_E(x_2 - m_d)].$$

The DCB receives the consumer fee regardless debit card usage. With probability $\delta\alpha^D$ the consumer will make a payment using the debit card, so the bank will receive the net per transaction payoff, which is a function of the merchant fee f^D.[10]

In addition to the per transaction fee, the bank earns interest on a positive balance in the customer account. A positive balance may exist for two reasons. If she did not make a purchase, but receives early income in period 2, the balance will be x_2

[10]As discussed earlier, the cost of funds do not enter the profit function as we have assumed the loan is zero NPV. Our results for the debit merchant fee are entirely independent of the way in which r_d is determined.

throughout that period. Alternatively, if she *did* make a purchase (or was mugged), and receives early income in period 2, the balance will be $x_2 - m_d = x_1 + x_2 - 1$ throughout that period. These two cases correspond to the third and fourth terms in the DCB's profit function.

Since the credit offered via the overdraft is priced perfectly competitively, the loan is zero NPV for the DCB. As a result, neither the revenues nor the costs from this loan show up in the profit function.

The DCB sets the optimal merchant fee by maximizing its payoff with respect to f^D, subject to

$$F^D = F^D_{\max}(f^D) \text{ and } \alpha^D = \alpha^D(f^D).$$

The optimal merchant fee is therefore:

$$f^*_D = \frac{1}{2}[c^D + 1 - \rho(1-h)] - \frac{1}{2}(1-\rho)v_0. \tag{3.1}$$

The merchant fee increases with the transaction cost faced by the bank and decreases with consumer surplus, v_0. When merchant extraction of consumer surplus is low, the debit card bank will set low merchant fees; this way, the acceptance rate will rise, thus increasing the value of the card to the consumer. As a result, the network can charge higher consumer fees. Note that the term v_0 is multiplied by $(1-\rho)$, the probability of the state in which debit cards enable payment when cash cannot.[11]

4. Credit Card Only Model

We now consider the case in which only a credit card is available to the consumer. We do, however, assume the consumer still has access to a current account, with an associated overdraft facility. The size of the overdraft facility is, once again, only sufficient to cover the desired overdraft if the consumer receives positive initial income. However, the credit line associated with the credit card enables the consumer to take out a larger loan. In this way, the credit card can enable payment in the no period 1 income case. Moreover, as with debit cards, credit cards insure against theft.

[11]Observe that the optimal debit merchant fee in (3.1) is the same as in the model of Bolt and Schmiedel (2011) without overdraft facility. This holds because the overdraft facility works similarly for cash as for debit cards and so presents no value added.

4.1. Consumer's problem. We denote by α^C the proportion of merchants who accepts the credit card and F^C is the consumer's credit card fee. Given merchant acceptance, recall that credit cards can be used in all states of the world regardless of period-1 income. Observe that $m_c \equiv 1 - \delta x_1$ denotes the average amount of debt when using the credit line associated with the credit card.

The consumer will want to hold a credit card as long as:

$$\begin{aligned}\delta\rho v_0 - \delta(\gamma_E + 2\gamma_L)r_d m_d \;\; &\leq \;\; \left(\alpha^C + \delta\rho(1-\alpha^C)\right) v_o - \alpha^C \gamma_L r_c m_c \\ &\quad - \delta(1-\alpha^C)(\gamma_E + 2\gamma_L)r_d m_d - F^C.\end{aligned}$$

If the consumer makes a payment with a credit card, she will have to pay interest on this credit line only if she needs to extend the credit for an extra period, having received no income at the end of period 1. If the merchant does not accept the card, and the consumer has to pay cash, she will then face the interest charges from the overdraft in each period, as previously discussed. This condition can be rearranged as follows:

$$F^C \leq \alpha^C(1-\delta\rho)v_o - \alpha^C \gamma_L r_c m_c + \delta\alpha^C(\gamma_E + 2\gamma_L)r_d m_d.$$

The consumer will never leave the house with cash if the merchant accepts a credit card. Indeed, given the way we have priced credit in the two models, the expected costs of servicing a credit line are the same as the equivalent costs associated with an overdraft, if the consumer has high initial income (it is only in this state where the consumer could use cash). This is because both are priced competitively. In other words,

$$(\gamma_E + 2\gamma_L)r_d = \gamma_L r_c$$

given,

$$r + (1-\gamma_E)r + (1-\gamma_E-\gamma_L) = (2-\gamma_E)r + (1-\gamma_E-\gamma_L).$$

from Conditions(2.1)-(2.2). Hence, since the consumer is indifferent regarding use of funds, she will certainly use her credit card so as to avoid mugging on her way to the store.

Of course, it was simply by assumption that the expected costs of servicing the credit line in the high income state were set the same as that of the overdraft. Yet, the credit card would not want to set the interest rate any higher, such that the consumer chose not to use the credit card. The higher interest rate would not directly affect the credit card network's profits because higher interest revenues in the profit function would be offset by a lower maximum fixed fee (given multipart pricing). However, the network's profits would indirectly reduce as the consumer uses her card for fewer transactions. Moreover, the maximum fixed fee would reduce by an amount equal to the cost savings on avoiding use of the overdraft $\delta\alpha^C(\gamma_E + 2\gamma_L)r_d m_d$. This would have a further effect on reducing network profits since there is no offsetting revenue term (given the network is not the same as the current account providing bank).

We further assume that if the credit line is taken down, the overdraft on the current account cannot be used to 'pay off' the credit line at the end of period 1. For instance, we assume the bank does not allow the overdraft to be used to pay off alternative debt; or at the very least, there exists a significant fixed cost to substituting overdraft debt for credit card debt.[12]

4.2. Merchant's problem. The merchant i receives profit $\pi(i)$ from a sale, where $\pi(i)$ is uniformly distributed between 0 and 1. His cost of handling cash is h whilst f^C is the merchant fee for accepting the debit card. His expected payoff from accepting cash is:

$$Z_{cash}(i) = \delta\rho[\pi(i) - h],$$

and his expected payoff from accepting the credit card is:

$$Z_c(i) = [\pi(i) - f^C].$$

[12]In some European countries, the overdraft is 'automatically' used to pay off outstanding credit card obligations at the end of the month. Hence, these consumers do not face a credit card interest rate but rather an interest rate on overdraft. However, consumers in the U.S. do not typically use overdrafts to pay off credit card debt, even if there are significantly lower interest rates on the former. This is sometimes called the 'credit card puzzle', and may be attributed to a specific behavioural trait or economic friction, but that discussion is beyond the scope of this paper (see e.g., Gross and Souleles (2002) and Telyukova and Wright (2008)).

Merchants accept debit cards only when

$$Z_{cash}(i) \leq Z_c(i).$$

Since there is a level of profits $\bar{\pi}$ above which merchants will accept credit cards, we can write the proportion of accepting merchants as follows:

$$\alpha^C(f^C) \;=\; \Pr[\pi(i) \geq \bar{\pi}] \;=\; 1 - \bar{\pi} \;=\; 1 - \frac{(f^C - \delta\rho h)}{1 - \delta\rho}.$$

This is different to the proportion associated with debit cards; the ρ here is multiplied by δ. This reflects the fact that the credit card allows for payment in both the high and low initial income states, unlike cash.

4.3. Maximum consumer fee for credit cards. Using the consumer's participation constraint, as well as α^C, we obtain the maximum consumer fee:

$$\begin{aligned} F^C_{\max}(f^C) \;&=\; [1 - f^C - \delta\rho(1-h)]v_0 - \\ &\quad \frac{[1 - f^C - \delta\rho(1-h)]}{(1-\rho\delta)}[\gamma_L r_c m_c - \delta(\gamma_E + 2\gamma_L) r_d m_d]. \end{aligned}$$

Unlike the debit fee, the probability of high initial income δ does not pre-multiply both terms; unlike the debit card, the credit card does not restrict the consumer to trade only in the high income state.[13]

The second term above captures the expected costs of credit; however, it is a function of both the credit line *and* the overdraft on the current account. In states where the credit card enables payment that would be impossible with cash, the relevant term for the expected cost of credit is simply $\gamma_L r_c m_c$. However, in the case of high period-1 income (which occurs with probability δ), the consumer *could* still use cash if she wished.[14] In this case, the relevant term is the *difference* between the cost of the credit line and the cost of the overdraft. It is this difference that captures the benefits (or otherwise) offered by the credit card.

[13]Notice that the maximum consumer fee becomes negative if $v_0 = 0$. Whilst the debit consumer fee is zero in this case, the credit consumer fee is negative since consumers would be paying higher expected interest costs under the credit card, than they would under the overdraft.

[14]Note of course that, if the consumer attempts to pay by cash, she will be mugged with probability $(1-\rho)$. In this case, she still enters her overdraft, even though she has not successfully made a purchase.

Given our earlier assumptions regarding r_d and r_c this difference is positive:

$$\gamma_L r_c m_c - \delta(\gamma_E + 2\gamma_L) r_d m_d = [(2-\gamma_E)r + (1-\gamma_E-\gamma_L)](1-\delta) > 0.$$

It might seem initially counterintuitive that this difference is non-zero, given they are priced competitively relative to the high income state. However, the loan is priced, conditional on the consumer requiring the loan in each case. Yet, when the consumer, ex ante, considers the value of a credit card she takes into account expected costs of the overdraft and the credit line; these are unconditional expected costs, before she knows the value of initial income. Since the credit card enables payment in one extra state of the world, the unconditional expected costs of credit via the credit card are higher than via the overdraft facility. Notice that the difference is decreasing in δ. As the probability of period-1 income increases, so does the probability of being able to pay using the cash and the overdraft facility. This increases the expected cost of the overdraft relative to that of the credit line on the credit card.

Indeed, even without assuming the loans are priced as zero NPV, it is difficult to imagine that the expected costs of the credit line to the consumer would be lower than the overdraft. After all, the positive difference above does not actually capture full costs of the credit line to the lender, relative to the overdraft. For this, we would need to add the term $\delta r x_1$ to the difference above, capturing the free credit equal to x_1 in the high initial state as part of the credit line.

4.4. Credit card network's problem. The Credit Card Network's (CCN) payoff from issuing a credit card is:

$$\pi^C = F^C + \alpha^C(f^C - c^C) - \alpha^C \delta r x_1.$$

Note that the consumer still has a current account, and overdraft facility, but neither of these show up in the credit card network's profit function. The final term reflects the expected loss on the credit line, as discussed in the section on interest rates. However, as discussed earlier, the profits would be equivalent if the network passed these costs onto the consumer, since the network extracts the full consumer surplus.

The network sets the optimal merchant fee by maximizing its payoff with respect to f^C, subject to

$$F^C_{\max} = F^C(f^C) \text{ and } \alpha^C = \alpha^C(f^C).$$

The optimal merchant fee for credit cards, as a function of r_c, is therefore:

$$f^*_C = \frac{1}{2}[c^C + 1 - \delta\rho(1-h)] + \frac{1}{2}\delta r x_1 + \frac{1}{2}[r + (1-\gamma_E)r + (1-\gamma_E-\gamma_L)]m_c - \frac{1}{2}\delta(\gamma_E + 2\gamma_L)r_d m_d - \frac{1}{2}(1-\delta\rho)v_0, \quad (4.1)$$

where we substitute out r_c.The fee is decreasing in the consumer's expected costs of servicing the overdraft. The intuition is straightforward. The overdraft, even in the absence of a debit card, offers an outside option to consumers in one state. By choosing to pay by credit card, not cash, the consumer avoids the expected costs of servicing an overdraft; if these are high, then the benefit of holding a credit card is high. In this case, the network can extract a large fee from the consumer, and is therefore able to reduce the merchant fee. In turn, high funding costs (r) for the credit card network will dampen the consumer maximum fixed fee resulting in a higher merchant fee to restore the balance. This has interesting implications. Effectively, the credit card competes with the overdraft facility in the state where cash could be used. It shows that the interest rate charged can impact the acceptance ratio of credit cards. An increase in the costs of an overdraft can lead to higher acceptance of credit cards.

In the comparative statics that follow we continue with our earlier assumption that the overdraft is a zero NPV loan, in order to pin down r_d. Recall, however, that our specific assumption about the credit line interest rate r_c will not affect the credit merchant fee.

The 'total' interest rate effect on merchant fees is then derived when we substitute $r_d = r_d(r, \gamma_E, \gamma_L)$, $r_c = r_c(r, \gamma_E, \gamma_L)$, $m_d = 1 - x_1$, and $m_c = 1 - \delta x_1$ in the optimal merchant fee f^*_C. This yields:

$$f_C^* = \frac{1}{2}[c^C+1-\delta\rho(1-h)]+\frac{1}{2}\delta r x_1+\frac{1}{2}[(2-\gamma_E)r+(1-\gamma_E-\gamma_L)](1-\delta)-\frac{1}{2}(1-\delta\rho)v_0. \quad (4b)$$

As with the consumer fee, the merchant fee is a function of the difference between the unconditional expected costs of servicing the credit line and the overdraft. It is also a function of initial income, since the consumer effectively gets free credit equal to her high initial income in the 'grace' period, rx_1. This equation also shows that higher funding rates r lead to higher merchant fees f_C^*. Higher defaults $(1-\gamma_E-\gamma_L)$ increase merchant fees as well. These effects make clear how merchants share the cost burden of credit card loans with consumers.

4.5. Comparison and comparative statics. In our model, the optimal debit card merchant fee f_D^* is not influenced by the funding cost or default risk. This derives from the fact that debit cards have no value added over cash regarding the use of the overdraft facility on the checking account. Debit cards only hedge against theft and that is why the probability of theft ρ plays an important role for the optimal merchant fee, as well as processing cost c_D.

In contrast, funding cost and default risk do affect the merchant fee on credit cards. In effect, merchants pay their 'fair' share with respect to credit card debt. If the network can extract lower surplus from consumers through a lower consumer fee, they will require merchants to pay a higher fee to compensate. An increase in r leads to an overall increase of f_C^*. In principle three effects are at play. One is because an increase in r leads to an increase in r_d, and as discussed above, this increases the saving the consumer can make from avoiding the costs of servicing the overdraft. This has a negative effect on the merchant fee as the credit card network tries to increase acceptance (a^C) to benefit from the higher extraction of surplus via the fixed consumer fee. The second is an opposing effect due to a lower consumer willingness-to-pay when credit card interest rates rise, making the credit card less acceptable to consumers and dampening the amount that the network can extract from consumers. The final effect is again a positive effect, coming from the subsidy provided to the consumer of free credit equal to x_1. This latter two effects dominate and therefore the CCN must increase the merchant fee when the funding cost rises.

$$\frac{\partial f_D^*}{\partial r} = 0 \quad \text{and} \quad \frac{\partial f_C^*}{\partial r} = \frac{1}{2}(2-\gamma_E)(1-\delta) + \frac{1}{2}\delta x_1 > 0.$$

A higher probability of early period-2 income γ_E increases the value of a credit card to consumers because it makes enjoying the grace period more likely. This allows a lower merchant fee, i.e.:

$$\frac{\partial f_D^*}{\partial \gamma_E} = 0 \quad \text{and} \quad \frac{\partial f_C^*}{\partial \gamma_E} = -\frac{1}{2}(1-\delta)(1+r) < 0.$$

Defining default $D = 1 - \gamma_E - \gamma_L$, and keeping γ_E constant, it is easy to show that

$$\frac{\partial f_D^*}{\partial D} = 0 \quad \text{and} \quad \frac{\partial f_C^*}{\partial D} = -\frac{\partial f_C^*}{\partial \gamma_L} = \frac{1}{2}(1-\delta) > 0.$$

That is, higher defaults lead to higher merchant fees. Once again, with higher default rates, the required interest rate on the credit line is higher; this reduces the maximum fee the network can charge consumers and so requires a higher fee from merchants. This effect is mitigated by a high probability of receiving period-1 income, i.e. δ large. When δ is large, the unconditional expected cost to the consumer of a credit line is not so much greater than an overdraft. In effect, the probability is low of being of being able to use a credit line in an extra state of the world.

For similar reasons, when the probability of receiving initial income rises then merchant fees go down

$$\frac{\partial f_D^*}{\partial \delta} = 0 \quad \text{and} \quad \frac{\partial f_C^*}{\partial \delta} = -\frac{1}{2}\left((2-\gamma_E-x_1)r + (1-\gamma_E-\gamma_L) + \rho(1-h-v_0)\right) < 0,$$

for sufficiently small v_0 and h. As δ increases, the unconditional expected cost of the credit line decreases relative to the overdraft (since there is an increase in the probability of being able to use the overdraft). Effectively, then, the credit card becomes more valuable to consumers. Since the network can extract a high fee from consumers, it will set a low merchant fee in order to maximize the network size. If the merchant fee is low, more merchants will accept the card and thus the card will become attractive to more consumers.

TABLE 3. Comparison between debit and credit cards

	funding cost r		default D		early income γ_E		initial income δ	
	1%	3%	5%	10%	50%	55%	95%	98%
f_D^*	0.00675	0.00675	0.00675	0.00675	0.00675	0.00675	0.00675	0.00675
α_D^*	0.42450	0.42450	0.42450	0.42450	0.42450	0.42450	0.42450	0.42450
f_C^*	0.02051	0.02357	0.02001	0.02051	0.02051	0.02000	0.03702	0.02051
α_C^*	0.34447	0.23574	0.36125	0.34447	0.34447	0.36141	0.39362	0.34447
r_d	0.08846	0.11154	0.04643	0.08846	0.08846	0.04778	0.08846	0.08846
r_c	0.28750	0.36250	0.14444	0.28750	0.28750	0.16125	0.28750	0.28750

Note: We set: $c^D = 0.0025$, $c^C = 0.005$, $h = 0.001$, $v_0 = 0$, and $\rho = 0.99$, $x_1 = 0.3$. Baseline parameters: $r = 0.01$, $\gamma_E = 0.50$, $\gamma_L = 0.40$ (but altered to $\gamma_L = 0.45$ for ($D = 5\%$)) and $\delta = 0.98$.

Table 3 illustrates the results[15]. As we can see from the table, the debit merchant fee and merchant acceptance rate do not depend on funding costs r, on the probability of default D or on the probability of different income shocks, γ_E or δ. However, the credit merchant fee increases with a higher cost of funding and a higher probability of default. We also see how the debit and the credit interest rates increase with the probability of default, and the funding cost, given they are priced in a competitive aftermarket.

Notice how the credit merchant fee decreases with a higher probability of period-1 income δ: the unconditional expected cost of servicing an overdraft increases, as the probability of using it increases. This decreases the relative cost to the consumer of using the credit card (as the overdraft, the outside option, increases in expected cost). As a result, the network can charge higher fees to the consumer and lower fees to the merchant. However, in these numerical examples, there is actually lower merchant acceptance of the credit card following an increase in δ: this is despite a lower merchant fee. There are actually opposing effects at work in determining the effect of δ on α^C; although a lower merchant fee increases the attractiveness of the card, the fact that δ increases the merchant's expected payoff from accepting only cash means that, at the margin, a given merchant is less willing to accept the card. For certain parameter values, such as those above, the latter effect dominates.[16]

[15]Many of these parameter estimates follow Bolt and Schmiedel (2011): see discussion in that paper.
[16]The comparative static for α^C with respect to δ is complex, so we omit it here. It is available from the authors on request.

5. Competition between Debit and Credit Cards

In this section, we examine competition between debit and credit cards.[17] We analyze the case in which the consumer multihomes and the merchant singlehomes (see discussion in section 7 on this issue). We also follow the preceding model and assume that the overdraft cannot be used to pay off the credit line in period 2. In other words, the consumer is committed to using the credit facility associated with the card he used for payment.

5.1. Consumers' participation. In what follows, α^i is the proportion of merchants who accept card i, where $i = C, D$. In addition, α denotes the proportion of merchants who hold *either* a debit or a credit card. Under the assumption of singlehoming merchants, this implies $\alpha = \alpha^D + \alpha^C$.

The consumer will hold both cards if:

$$\begin{aligned}\delta\rho v_0 - \delta(\gamma_E + 2\gamma_L)r_d m_d \;\leq\; & \left(\delta[(1-\alpha)\rho + \alpha] + (1-\delta)\alpha^C\right) v_o - \alpha^C \gamma_L r_c m_c \\ & -\delta(1-\alpha)(\gamma_E + 2\gamma_L)r_d m_d - \delta\alpha^D(\gamma_E + 2\gamma_L)r_d m_d - F_T,\end{aligned}$$

where F_T denotes the maximum total fee that the consumer is prepared to pay to hold both a debit and a credit card. For the reasons discussed earlier we continue to assume that the indifferent consumer will use a credit card, rather than use the overdraft.

We can rearrange to find the maximum total consumer fee, as a function of merchant acceptance:

$$\begin{aligned}F_T &= [(1-\rho)\alpha + (1-\delta)\alpha^C]v_o + \delta\alpha^C(\gamma_E + 2\gamma_L)r_d m_d - \alpha^C \gamma_L r_c m_c \\ &= [(1-\rho)\alpha + (1-\delta)\alpha^C]v_o - \alpha^C(1-\delta)[(2-\gamma_E)r + (1-\gamma_E-\gamma_L)].\end{aligned}$$

Note that individual contributions to the total fee will satisfy[18]

[17] See e.g. Armstrong (2006) and Guthrie and Wright (2007) for a comprehensive analysis of competition in two-sided markets.

[18] This breakout is derived by observing that merchants only accept one type of payment card (see discussion in Bolt and Schmiedel (2011)). As a result, the individual contribution from the debit card is given by the participation constraint on debit cards (where cash is the outside option): $\delta\rho v_o - \delta(\gamma_E + 2\gamma_L)r_d m_d \leq \delta\alpha^D v_o + \delta(1-\alpha^D)\rho v_o - \delta(\gamma_E + 2\gamma_L)r_d m_d - F_T^D$. The equivalent constraint for credit cards is: $\delta\rho v_o - \delta(\gamma_E + 2\gamma_L)r_d m_d \leq \alpha^C[v_o - \gamma_L r_c m_c] + (1-\alpha^C)\delta\rho v_o - (1-\alpha^C)\delta(\gamma_E + 2\gamma_L)r_d m_d - F_T^C$

$$F_T = F_D^T + F_C^T,$$

where

$$F_T^D = \delta\alpha^D(1-\rho)v_o,$$

and

$$F_T^C = \alpha^C(1-\delta\rho)v_o - \alpha^C(1-\delta)[(2-\gamma_E)r + (1-\gamma_E-\gamma_L)].$$

5.2. Merchants' acceptance. We assume that merchants singlehome; if they accept a card at all, it is *either* a debit *or* a credit card. In equilibrium, only merchants with high profit margins accept credit cards, since these are more costly: intermediate merchants accept debit cards, and low-end merchants accept cash.[19] Using the expected payoffs above, we can find the profit level above which merchants are prepared to accept debit cards $\bar{\pi}_d$ and likewise the profit level above which they will accept credit cards $\bar{\pi}_{dc}$;

$$\bar{\pi}_d(f_D) = \frac{f^D - \rho h}{1-\rho} \quad \text{and} \quad \bar{\pi}_{dc}(f_D, f_C) = \frac{f^C - \delta f^D}{1-\delta}.$$

This gives us the following acceptances:

$$\alpha(f_D) = 1 - \bar{\pi}_d(f^D) \quad \text{and} \quad \alpha^C(f_D, f_C) = 1 - \bar{\pi}_{dc}(f^D, f^C),$$

where debit card acceptance is:

$$\alpha^D = \alpha(f^D) - \alpha^C(f^D, f^C).$$

5.3. Networks' optimization. The CCN and the DCB engage in Bertrand competition.

5.3.1. *Debit card network.* The DCB, issuing the debit card, maximizes its profit function, with respect to f^D, subject to:

[19]Credit cards will be have higher merchant fees in equilibrium since they must cover higher expected costs of credit as well as higher per-transaction processing costs (assuming, as in practice, that $c^C > c^D$).

$$F_T^D = F_{T\max}^D(f^D) \text{ and } \alpha^C = \alpha^C(f^D, f^C).$$

However, its profit function is slightly altered from that of the debit-only world:

$$\begin{aligned}
\pi^{DCB} &= F^D + \delta\alpha^D(f^D - c^D) \\
&\quad +(1-\alpha^C)r[(1-\delta)\gamma_E x_2 + \delta\gamma_E(x_2 - m_d)] \\
&\quad +\alpha^C r[\delta x_1 + \gamma_E(x_2 - m_c)].
\end{aligned}$$

As in the no competition case, the bank can earn interest on positive balances, even in the absence of a credit card. However, the presence of the credit card affects both the frequency and size of the consumer's positive balance. This has positive and negative effects on the DCB's profit function. When the consumer pays by credit card, the DCB benefits from the delayed deduction of funds from the current account. Any funds remain in the current account for the duration of period 1, until the end of the 'grace' credit period. During this time, the DCB can earn interest on any positive balance, at market interest rate r. However, the credit card also enables the consumer to make a purchase in more states of the world. As a result, the size of the positive balance following early income will be smaller in period 2, relative to the no credit card case. This can be seen by rearranging the above profit function:

$$\begin{aligned}
\pi^{DCB} &= F^D + \delta\alpha^D(f^D - c^D) + r[(1-\delta)\gamma_E x_2 + \delta\gamma_E(x_2 - m_d)] \\
&\quad +\alpha^C r[\delta x_1 - (1-\delta)\gamma_E].
\end{aligned}$$

The last line captures this trade-off. It reflects an interesting case: if expected period-1 income δx_1 is sufficiently large, the DCB's profit function will increase with any increase in the proportion of merchants accepting the *credit* card. This is important. Although the credit and debit networks are in competition, there is also this element of complementarity between the debit card and the credit card. However, if the reverse holds, this complementarity will not exist.

The tradeoff continues to play a role when we solve for the optimal debit card merchant fee.

$$f^D(f^C) = \frac{1}{2}\frac{c^D(1-\delta\rho)+(1-\delta)\rho h+(1-\rho)f^C}{1-\rho\delta} - \frac{1}{2}(1-\rho)v_0 + \frac{1}{2}\frac{(1-\rho)}{1-\delta\rho}r[\delta x_1 - (1-\delta)\gamma_E]. \tag{5}$$

For a given f^C, the optimal merchant fee in the debit network is increasing in market interest rate, as long as δ is sufficiently high such that $\delta x_1 > (1-\delta)\gamma_E$. At the margin, if the DCB expects to earn a large amount on positive balances in period 1, it will set a high merchant fee to *discourage* debit acceptance in favour of credit cards. This allows for the possibility that, in equilibrium, we may observe higher debit merchant fees compared with credit merchant fees.

5.3.2. *Credit card network.* The profit function of the CCN remains unchanged, relative to the no competition case. That is:

$$\pi^{CCN} = F^C + \alpha^C(f^C - c^C) - \delta\alpha^C r x_1.$$

It now maximizes this profit function, with respect to f_d, subject to

$$F_T^C = F_{T\max}^C(f^C, f^D) \text{ and } \alpha^C = \alpha^C(f^D, f^C).$$

The optimal merchant fee for credit cards, having substituted in r_d, is therefore:

$$f^C(f^D) = \frac{1}{2}[c^C+1-\delta(1-f^D)] - \frac{1}{2}(1-\delta\rho)v_0 + \frac{1}{2}\delta r x_1 + \frac{1}{2}(1-\delta)[(2-\gamma_E)r + (1-\gamma_E-\gamma_L)]. \tag{6}$$

This is similar to the merchant fee in the credit-only model. The major difference is that the fee is a function of the debit merchant fee f^D, rather than the merchant's cost of cash, h.

The unique equilibrium merchant fees (f_D^{**}, f_C^{**}) are found from the intersection of the two best response functions, $f^D(f^C)$ and $f^C(f^D)$ (see appendix).

Table 4. Comparison between debit and credit cards: default and funding cost

Default ($r = 1\%$)	Monopoly				Competition			
	debit		credit		debit		credit	
	$D = 5\%$	$D = 10\%$	$D = 5\%$	$D = 10\%$	$D = 5\%$	$D = 10\%$	$D = 5\%$	$D = 10\%$
f^*	0.00675	0.00675	0.02001	0.02051	0.00492	0.00501	0.01703	0.01757
α^*	0.42450	0.42450	0.36125	0.34447	0.21797	0.23159	0.38944	0.36668
r_d	0.04643	0.08846	0.04643	0.08846	0.04643	0.08846	0.04643	0.08846
r_c	0.14444	0.28750	0.14444	0.28750	0.14444	0.28750	0.14444	0.28750
Funding cost ($D = 10\%$)	Monopoly				Competition			
	debit		credit		debit		credit	
	$r = 1\%$	$r = 3\%$	$r = 1\%$	$r = 3\%$	$r = 1\%$	$r = 3\%$	$r = 1\%$	$r = 3\%$
f^*	0.00675	0.00675	0.02051	0.02375	0.00501	0.00664	0.01757	0.02161
α^*	0.42450	0.42450	0.34447	0.23574	0.23159	0.19057	0.36668	0.24463
r_d	0.08846	0.11154	0.08846	0.11154	0.08846	0.11154	0.08846	0.11154
r_c	0.28750	0.36250	0.28750	0.36250	0.28750	0.36250	0.28750	0.36250

Note: We set: $c^D = 0.0025$, $c^C = 0.005$, $h = 0.001$, $v_0 = 0$, $\rho = 0.99$, $\gamma_E = 0.50$, $\delta = 0.98$, and $x_1 = 0.3$. Baseline parameters: $r = 0.01$ and 0.03, $\gamma_L = 0.45$ ($D = 5\%$) and 0.40 ($D = 10\%$).

5.4. Comparison and comparative statics. Table 4 compares competitive and monopolistic card fees for two different default levels ($D = 5\%$ vs. $D = 10\%$) and funding cost levels ($r = 1\%$ vs. $r = 3\%$).

First notice how an increase in default risk affects interest rates on debit and credit cards. As observed before, monopolistic debit card fees are not affected by default risk changes. The value of debit cards is driven solely by security concerns as they generate no advantage over cash with respect to the use of the overdraft facility. However, *competitive* debit card merchant fees *are* affected by default risk movements.

Notice that higher default leads to higher competitive debit card merchant fees but to higher debit card acceptance as well. (We show the comparative static in the appendix.) Total card acceptance decreases however. Intuitively, higher default increases the credit card merchant fee, allowing the competing debit merchant fee to rise as well. Although this has a negative effect on merchant acceptance of debit cards, this effect is smaller than the reduction in acceptance of credit cards. Since the merchants who no longer accept credit cards will switch to debit cards, this results in an overall increase in debit card acceptance.

Second, we observe higher competitive merchant fees when the funding cost increases. Note that debit card merchant fees, which were not affected by funding costs in the monopolistic case, may rise considerably in the competition case. They may even approach monopolistic levels. However, this is not primarily due to relaxation of competitive pressure given the rise in credit merchant fees—in fact, the latter rises by a lower proportion compared with the debit merchant fee. The effect is coming from the complementarity between debit and credit cards. The bank can benefit from a positive balance in the current account while the consumer enjoys the 'grace period'; the returns on the positive balance increase with r and so the bank substantially increases the debit merchant fee, to discourage debit card usage. As a result, debit card acceptance α^D actually decreases.

This demonstrates our key result. Although competitive pressures reduce fees for both cards (as found in Bolt and Schmiedel (2011), they also lead to an element of complementarity between debit and credit cards when the two cards offer different credit possibilities. As a result, debit card fees may actually be relatively high, despite competition from credit cards.

6. Welfare

We now turn to consider welfare maximising fees. In effect, we will find the fee in each case such that the optimal proportion of merchants are induced to accept the card. These results closely mirror those in Bolt and Schmiedel (2011), but we review them here as a means of comparing them with our new results in earlier sections. Given the repetition, however, we only review below the cash-only world, and derive the welfare in an environment with both cards. We leave the debit-only and credit-only environments to the appendix.

When we consider social welfare, interest rates and fees are merely transfers between agents. As a result, the complementarity effect which operates in the context of competition will be irrelevant for considering social welfare. However, welfare will be a function of the probability of default since this represents deadweight loss. Notice this explicitly enters the welfare function, unlike in the case of private optimisation, when default only featured through the interest rates. In effect, the benefit from cards accrues from the extra surplus v_0 but part of the cost comes from the possibility the consumer cannot repay the loan.

6.1. Cash-only economy. As a baseline case, the welfare in a cash-only economy is given as follows:

$$W^{cash} = \delta\rho v_0 + \delta\rho(\frac{1}{2} - h) - \delta(1 - \gamma_E - \gamma_L)m_d$$

Notice that there is still a positive probability of default in a world without payment cards; this is because the consumer can use her overdraft facility to withdraw cash for payment. Regardless of whether she is mugged, she will then default on repaying this debt if she receives no period 2 income.

6.2. Credit and Debit Cards. When both cards are present, welfare W^{DC} is given by:

$$\begin{aligned} W^{DC} &= \delta[(1-\alpha)\rho + \alpha]v_o + (1-\delta)\alpha^C v_o \\ &\quad + \alpha^C(\frac{2-\alpha^C}{2}) + \delta\alpha^D(\frac{2-\alpha-a^C}{2}) + \delta\rho(1-\alpha)(\frac{1-\alpha}{2} - h) \\ &\quad - \alpha^C c^C - \delta\alpha^D c^D - \alpha^C(1-\gamma_E-\gamma_L)m_c - \delta(1-\alpha^C)(1-\gamma_E-\gamma_L)m_d. \end{aligned}$$

The first term captures the expected benefit of the purchase to the consumer. The second term. in large brackets, represents the expected benefit to merchants, using the fact that merchants are distributed uniformly on the interval 0 to 1. The final terms capture the deadweight costs of debit cards; the expected transaction costs and the expected costs of default.

The welfare maximising fees are given by the following

$$\begin{aligned} &\arg\max_{f^D, f^C} W^{DC} \\ s.t.\alpha &= 1 - \frac{f^D - \rho h}{1-\rho} \;,\; \alpha^C(f_d, f_c) = 1 - \frac{f^C - \delta f^D}{1-\delta} \\ \text{and } \alpha^D &= \alpha(f^D) - \alpha^C(f^D, f^C). \end{aligned}$$

This yields the following two best response functions:

$$f^D_{opt}(f^C_{opt}) = c^D + \frac{f^C(1-\rho) - c^C(1-\rho) - (1-\delta)(1-\rho)(1-\gamma_E-\gamma_L)}{1-\delta\rho},$$

$$f^C_{opt}(f^D_{opt}) = c^C - c^D\delta - v_o(1-\delta) + (1-\delta)(1-\gamma_E-\gamma_L) + \delta f^D.$$

These equations can be solved to find the unique welfare maximising fees:

$$f^D_{opt} = c^D - (1-\rho)v_o,$$

$$f^C_{opt} = c^C + (1-\delta)(1-\gamma_E-\gamma_L) - (1-\delta\rho)v_o.$$

Using these fees, we can obtain the welfare maximising proportion of merchants accepting each type of card. The total card acceptance is given by

$$\alpha_{opt} = 1 + v_o - \frac{c^D}{1-\rho} + \frac{\rho}{1-\rho}h,$$

while credit card acceptance is given by

$$\alpha^C_{opt} = 1 + v_o + \frac{\delta}{1-\delta}c^D - \frac{1}{1-\delta}c^C - (1-\gamma_E-\gamma_L).$$

This means debit card acceptance is given by

$$\begin{aligned}\alpha^D_{opt} &= \alpha_{opt} - \alpha^C_{opt}\\ &= (1-\gamma_E-\gamma_L) + \frac{\rho}{1-\rho}h - \frac{1-\delta\rho}{(1-\delta)(1-\rho)}c^D + \frac{1}{1-\delta}c^C.\end{aligned}$$

Whereas the complementarity between cards was relevant in the privately competitive framework, this does not affect the socially optimal fees as discussed above. As a result, the debit merchant fee is no longer a function of expected costs of default and is identical to the socially optimal fee in the debit-only world. (Nevertheless, as in Bolt and Schmiedel (2011), the proportion of merchants accepting debit cards relative to credit cards still increases with the probability of default).

The table below compares monopolistic and competitive merchant fees, with the socially optimal fees, at default levels (D) of 10%. This table confirms the results of Bolt and Schmiedel (2011), that competitive fees are still large relative to the socially optimal level. However, there is an extra reason in this paper for the large wedge between competitive and socially optimal debit fees; as discussed above, competitive debit fees are inefficiently high in part due to the complementarity effect between debit and credit cards.

Table 5. Comparison between Monopoly, Competition and Social Optimality.

	Cash	Debit Only		Credit Only		Debit and Credit	
		Monopoly	Social	Monopoly	Social	Compet.	Social
f^D		0.00675	0.00250			0.00501	0.00250
α^D		0.42450	0.84900			0.23159	0.07650
f^C				0.02051	0.00700	0.01757	0.00700
α^C				0.34447	0.79766	0.36668	0.77250
Welfare	0.41553	0.41818	0.41906	0.42195	0.42501	0.42307	0.42503

Note: We set: $c^D = 0.0025$ $c^C = 0.005$, $h = 0.001$, $v_0 = 0$ $\rho = 0.99$, $\gamma_E = 0.5$, $\gamma_L = 0.4$, $\delta = 0.98$ $r = 0.01$, and $x_1 = 0.3$.

Note, also, that for high probabilities of period 1 income (δ high) there is very little extra benefit to a credit card, relative to a debit card. As a result, the proportion of debit card acceptance increases relative to credit card acceptance as δ increases.

7. Robustness

As in Bolt and Schmiedel (2011), we make certain simplifying assumptions regarding consumer and merchant type. First, consumers are homogeneous and merchants are heterogeneous in our model. Second, we assume that consumers will multihome while merchants singlehome. It is worth bearing in mind, however, that our key results do not qualitatively depend on these assumptions.

First of all, note that if both sides were homogeneous, there would be no element of competition required, leading to a trivial equilibrium. If instead we considered heterogeneous consumers, this would lead some of them to accept the card and others to reject; or, in the case of two cards, some consumers would accept both cards, while others just accepted one, or none at all. If merchants were homogeneous we would be back considering an 'all-or-nothing' corner solution, whilst if they were heterogeneous as well as consumers, this would lead to separation on both sides of the market.

In any case, however, these alternative assumptions would not change the element of complementarity between debit and credit cards. Nor does it change the fact that credit cards will incorporate default costs in a more direct way than debit cards. In the model above, we found that debit cards were relatively high in competition, due to the complementarity effect. This would not disappear if card networks were forced to compete either on different sides of the market, or on both. Whilst greater competition might lower fees all round, the complementarity effect would still leave debit card fees *relatively* high.

At this point, we should also note the assumption we have made about use of overdrafts and credit lines. We have assumed that consumers are willing and able to use their overdrafts via their debit cards. Moreover we have assumed that they cannot (or will not) use their overdraft to partially pay off their credit line.

In reality, we actually observe many different practices, some the result of cultural or behavioural characteristics. European consumers differ from US consumers regarding their credit card use. Europe consumers are less likely to use their card for credit purposes, but rather simply rely on the payment element. In some parts of Europe, consumers are able to use overdrafts with debit cards, while in others they are not (or do not choose to do so).

Sometimes, checking account balances are used to pay back outstanding credit card payments. Instead of revolving the credit card debt and paying interest rate r_c, consumers may now draw upon their overdraft facility for repayment and pay interest rate r_d. Although the interest effects in our model will somewhat be mitigated, the credit line channel will still affect payment fees, since $m_c - m_d > 0$.

Related to this observation is the fact that few European consumers pay interest on their credit card debt. Those loans are repaid at the end of the 'grace' period, or not at all, that is, consumers default. This implies that credit card loans cannot be zero NPV. In this case the cost of funds burden must be shifted explicitly towards merchants and consumers in the form of higher payment fees.

No model can hope to capture all different types of observed behaviour. Nevertheless, we have taken a key step in highlighting the important role of credit in payment card competition, and in so doing explored a hitherto ignored element of complementarity.

8. Conclusions

In this model we examine the role of consumer credit in both debit and credit card networks. We allow for the fact that the consumer will always have access to a current account, with an associated overdraft facility. This account is provided by the bank which would issue an associated debit card.

In the 'credit card only' world, the credit card effectively competes with the overdraft facility in the state where cash could be used. As a result, higher expected costs of servicing an overdraft will allow the credit card network to increase the consumer fee and lower the merchant fixed fee; this will increase the acceptance ratio of credit cards among merchants.

Our model also shows that cost of funds and default risk affect debit cards and credit cards in a different way. Specifically, in a 'debit card only' world, these factors have no effect on the merchant fee, while they do affect credit card merchant fees. In a competitive situation, these cost factors drive both cards, but credit card merchant fees are more affected than debit card merchant fees. Debit merchant acceptance actually increases with the probability of default, despite an increase in debit merchant fee, since some merchants switch from credit cards to debit cards.

The debit merchant fee also depends on funding costs in the context of competition. However, as a result, the debit card fee may be increased to discourage debit card acceptance at the margin. Effectively, we find there is a degree of complementarity, as well as competition, between the two networks. The bank providing the debit card and current account actually benefits from consumers using credit cards, if they have positive initial income. In effect the bank benefits from the 'free credit' period offered to the consumer by the credit card network, as the bank can earn interest on the balance that remains in the current account during this period. If the probability of initial income is high, therefore, this complementarity incentivises the bank to increase the merchant debit card fee and reduce merchant acceptance of the debit card.

These results help to inform current debates about the pricing of debit and credit card fees. Recent discussion has focused on whether there should be differential interchange fees for debit and credit cards. Although we do not explicitly model the interchange fee, it will be closely related to the merchant fee (see Bolt (2006)). We therefore shed new light on how to understand the different drivers at work in affecting debit and credit card fees.

Appendix 1: Derivation of Competitive Merchant Fees and Comparative Statics

Note: All algebraic expressions and numerical results in our paper are verified using Mathematica, version 8, and program files are available upon request.

The intersection of (upward-sloping) reaction functions $f^C(f^D)$ and $f^D(f^C)$ yields (f^{**}_D, f^{**}_C), where

$$\begin{aligned} f^{**}_D &= \frac{1}{4-3\rho\delta-\delta}\left(ab[1-\gamma_E-\gamma_L]+[ab(2-3\gamma_E)+3b\delta x_1]r+\right. \\ &\quad \left.[bc^C+2cc^D+2a\rho h+ab]-3bcv_0\right), \end{aligned}$$

$$\begin{aligned} f^{**}_C &= \frac{1}{4-3\rho\delta-\delta}\left(2ac[1-\gamma_E-\gamma_L]+[2ac(2-\gamma_E)-ab\delta\gamma_E+(b\delta+2c)\delta x_1]r+\right. \\ &\quad \left.[c(2c^C+\delta c^D)+2a\rho\delta h+2ac]-c(b\delta+2c)v_0\right), \end{aligned}$$

where:

$a=1-\delta$, $b=1-\rho$, and $c=1-\rho\delta$. Note that $4-3\rho\delta-\delta=4c-b\delta>0$.

For the partial derivative wrt default D, we find:

$$\begin{aligned} \frac{\partial f^{**}_D}{\partial D} &= -\frac{\partial f^{**}_D}{\partial \gamma_L}=\frac{ab}{4-3\rho\delta-\delta}>0, \\ \frac{\partial f^{**}_C}{\partial D} &= -\frac{\partial f^{**}_C}{\partial \gamma_L}=\frac{2ac}{4-3\rho\delta-\delta}>0. \end{aligned}$$

It easy to show that $\partial f^{**}_C/\partial D > \partial f^{**}_D/\partial D$. Furthermore, by combining our results for merchant fees with the conditions for merchant acceptance, we can show that debit card merchant acceptance is increasing in the default rate:

$$\frac{\partial \alpha^{**}_D}{\partial D}=\frac{c}{4-3\rho\delta-\delta}>0.$$

For funding cost r we find

$$\frac{\partial f^{**}_D}{\partial r}=\frac{ab(2-3\gamma_E)+3b\delta x_1}{4-3\rho\delta-\delta}>0,$$

for sufficiently large average period-1 income δx_1 if $\gamma_E \geq 2/3$, and:

$$\frac{\partial f_C^{**}}{\partial r} = \frac{2ac(2-\gamma_E) - ab\delta\gamma_E + (b\delta + 2c)\delta x_1}{4 - 3\rho\delta - \delta} > 0,$$

since $c > b$ and $2 - \gamma_E > 1$. For the probability of early period-2 income γ_E, we find:

$$\begin{aligned} \frac{\partial f_D^{**}}{\partial \gamma_E} &= -\frac{ab(1+3r)}{4-3\rho\delta-\delta} < 0, \\ \frac{\partial f_C^{**}}{\partial \gamma_E} &= -\frac{2ac(1+r) + ab\delta r}{4-3\rho\delta-\delta} < 0. \end{aligned}$$

Finally, defining $x^e = \delta x_1$, we find (keeping δ constant):

$$\frac{\partial f_D^{**}}{\partial x^e} = \frac{3br}{4-3\rho\delta-\delta} > 0 \quad \text{and} \quad \frac{\partial f_C^{**}}{\partial x^e} = \frac{(b\delta+2c)r}{4-3\rho\delta-\delta} > 0.$$

Appendix 2: Welfare Analysis

Debit-only economy. In a debit-only world, welfare is given by W^{debit} where

$$
\begin{aligned}
W^{debit} &= \delta[\alpha^D + \rho(1-\alpha^D)]v_0 \\
&+\delta\left\{\alpha^D(\frac{2-\alpha^D}{2}) + \rho(1-\alpha^D)(\frac{1-\alpha^D}{2} - h)\right\} \\
&-\delta\alpha^D c^D - \delta(1-\gamma_E-\gamma_L)m_d.
\end{aligned}
$$

The welfare maximising fee is given as follows:

$$
\begin{aligned}
&\arg\max_{f^D}\ W^{debit} \\
s.t.\ \alpha^D &= 1 - \frac{f^D - \rho h}{1-\rho}.
\end{aligned}
$$

Although the expected cost of default enters the welfare measure, it has no effect on the optimal fee because the probability of default does not depend on α^D, the probability of debit card usage. This follows from our assumption that the overdraft, rather than the debit card itself, provides the consumer with sufficient means to make the purchase in the high income state. The benefit of the debit card only comes from greater security over cash.

The optimal proportion of merchant acceptance is

$$\alpha^D_{opt} = 1 + v_0 - \frac{c^D}{1-\rho} + \frac{\rho h}{1-\rho},$$

which means the optimal fee is

$$f^D_{opt} = c^D - v_0(1-\rho).$$

Intuitively, the merchant fee increases in the network cost c^D, but decreases with the social benefit from debit cards. This benefit comes from the additional consumer surplus v_0 that can be obtained in states where cash is insecure. Note the merchant fee is the same as the welfare maximising debit fee in Bolt and Schmiedel (2011).

Credit-only economy. In a credit-only world, the welfare is W^{credit} where:

$$
\begin{aligned}
W^{credit} &= [\alpha^C + \delta\rho(1-\alpha^C)]v_o \\
&+\alpha^C(\frac{2-\alpha^C}{2}) + \delta\rho(1-\alpha^C)(\frac{1-\alpha^C}{2} - h) \\
&-\alpha^C c^C - \alpha^C(1-\gamma_E-\gamma_L)m_c - \delta(1-\alpha_c)(1-\gamma_E-\gamma_L)m_d.
\end{aligned}
$$

The optimal fee is given by

$$
\begin{aligned}
&\arg\max_{f^C} W^{credit} \\
s.t.\ \alpha^C &= 1 - \frac{f^C - \delta\rho h}{1-\delta\rho}.
\end{aligned}
$$

In contrast to the debit only economy, the probability of default in the credit-only economy is a function of α^C, the proportion of merchants accepting the card. The socially optimal proportion is given by

$$
\alpha^C_{opt} = v_o + 1 - \frac{[c^C - \rho\delta h + (1-\gamma_E-\gamma_L)(1-\delta)]}{1-\delta\rho},
$$

which is clearly decreasing in the expected cost of default. By extension, the optimal fee is

$$
f^C_{opt} = c^C - (1-\delta\rho)v_o + (1-\delta)(1-\gamma_E-\gamma_L).
$$

Conclusion

In these three chapters, I have explored various aspects of modern banking. Each issue addressed has relevance when considering what the next ten or so years hold for banks. The first chapter explores an area of key vulnerability for banks in the lead up to the recent financial crisis. It highlights how capital regulation under Basel I and II contributed to this vulnerability. This issue may well be addressed in Basel III as there have been suggestions that off-balance sheet items such as credit lines may face a 100% credit conversion factor. In the context of the paper, this would be a beneficial policy change. Yet, this is not the end of the story: further work is needed to consider how important are the beneficial effects of credit lines, such as the signalling role they play, and how optimal capital regulations should take this into account. Moreover, a change in capital regulations to discourage credit lines will encourage banks to capture surplus through some other means: the question remains how exactly they will go about doing so, and what will be the consequences for bank lending in good times, and in bad.

The relevance of large-value settlement systems will also continue, not simply due to the increasing interconnected nature of financial transactions. The financial crisis has encouraged the use of central exchanges for the settlement of certain trades, which were previously organised on a bilateral basis. Whilst such exchanges do not share all of the same characteristics as traditional large-value payment, their interconnectedness throws up similar questions regarding how problems in one system may spill over to another.

Finally, the next few years will undoubtedly see a growth in the area of retail payments, and retail banking in general. This trend is not simply spurred on by banks' need for alternative sources of profitability following the crisis: the development of technology means that payment cards will increasingly operate as part of a broader package. For instance, there are already linkages between mobile phones and payment cards in some parts of the world. Our paper takes into account that payment cards are twinned with other banking services, such as credit and current accounts. Future work in this area will need to consider even broader business models; models that combine cards with other forms of consumer service, such as mobile phones.

Bibliography

Afonso, G. and H.S. Shin (2009), 'Systemic risk and liquidity in payment systems', *FRB of New York Staff Report* .

Angelini, P. (1988), 'An analysis of competitive externalities in gross settlement systems', *Journal of Banking and Finance* **22**, 1–18.

Armstrong, M. (2006), 'Competition in two-sided markets', *RAND Journal of Economics* **37**, 668–691.

Bank For International Settlements (2009), 'Annual report 2009'.
URL: *http://www.bis.org/publ/arpdf/ar2010e.htm*

Bank of England (2008a), 'Inflation report: February 2008'.
URL: *http://www.bankofengland.co.uk/publications/inflationreport/ir08feb.pdf*

Bank of England (2008b), 'Credit conditions survey: 2008 q4'.
URL: *http://www.bankofengland.co.uk/publications/other/monetary/creditconditions survey090102.pdf*

Basel Committee on Banking Supervision (2009), 'Strengthening the resilience of the banking sector'.
URL: *http://www.bis.org/publ/bcbs164.pdf*

Baxter, W. (1983), 'Bank interchange of transactional paper: Legal and economic perspectives', *Journal of Law and Economics* **26**, 541–588.

Bech, M. and R. Garrett (2003), 'The intraday liquidity management game', *Journal of Economic Theory* **109**, 198–210.

Becher, C., M. Galbiati and M. Tudela (2007), 'The timing and funding of chaps sterling payments', *Bank of England mimeo* .

Bhalla, V.K. (2008), *Working Capital Management*, 7th edn, Anmol Publications Pvt. Ltd.

Blavy, R. (2005), 'Monitoring and commitment in bank lending behaviour', IMF Working Paper.

Bolt, W (2006), 'Retail payments in the netherlands: Facts and theory', *De Economist* **154**, 345–372.

Bolt, W. and H Schmiedel (2011), 'Sepa, efficiency, and payment card competition', *Annals of Finance, forthcoming* .

Bolt, W. and S. Chakravorti (2008), 'Consumer choice and merchant acceptance in payment media', Federal Reserve Bank of Chicago Working Paper.

Boot, A., A.V. Thakor and G.F. Udell (1987), 'Competition, risk neutrality and loan commitments', *Journal of Banking & Finance* **11**, 449–471.

Chakravorti, S. and R. Rosen (2006), 'Platform competition in two-sided markets: The case of payment networks', *Review of Network Economics* **5**, 118–142.

Chakravorti, S. and T. To (2007), 'A theory of credit cards', *International Journal of Industrial Organisation* **25**, 583–595.

Chari, V.V., L.J. Christiano and P.J. Kehoe (2008), 'Facts and myths about the financial crisis of 2008', Federal Reserve Bank of Minneapolis Working Paper.

Chateau, J.P. (2007), 'Basel-2 revised standard approach and beyond: Credit risk valuation of short-term loan commitments', CCF Research Conference, WHU.

Cho, I. and D.M. Kreps (1987), 'Signaling games and stable equilibria', *The Quarterly Journal of Economics* **102**, 179–222.

Committee on Payment and Settlement Systems (2006), 'Cross border collateral arrangements'.
URL: *http://www.bis.org/publ/cpss71.htm*

Committee on Payment and Settlement Systems (2008), 'The interdependencies of payment and settlement systems'.
URL: *http://www.bis.org/publ/cpss84.htm*

Committee on Payment and Settlement Systems (2011), 'Principles for financial market infrastructures'.
URL: *http://www.bis.org/publ/cpss94.htm*

Cornett, M.M., J. J. McNutt, P.E. Strahan and H. Tehranian (2010), 'Liquidity Risk Management and Credit Supply in the Financial Crisis', Working Paper.

Diamond, D.W. and P.H. Dybvig (1983), 'Bank runs, deposit insurance and liquidity', *Journal of Political Economy* **91**, 40119.

Diamond, W.D. and G.R. Rajan (2001), 'Liquidity risk, liquidity creation, and financial fragility: A theory of banking', *Journal of Political Economy* **109**, 287–327.

Financial Leadership Exchange (2008), 'Loan covenants in the credit crisis'.
URL: *http://www.executiveboard.com/businessweek/*

Gatev, E. and P.E. Strahan (2006), 'Banks' advantage in hedging liquidity risk: Theory and evidence from the commercial paper market', *The Journal of Finance* **61**, 867–892.

Gatev, E., P.E. Strahan and T. Schuermann (2009), 'Managing bank liquidity risk: How deposit-lending synergies vary with market conditions', *Review of Financial Studies* **22**, 995–1020.

Greenspan, A. (2008), *The Age of Turbulence*, Penguin Books.

Gross, D. and N Souleles (2002), 'Do liquidity constraints and interest rates matter for consumer behavior? evidence from credit card data', *Quarterly Journal of Economics* **117**, 149–185.

Guthrie, G. and J. Wright (2007), 'Competing payment schemes', *Journal of Industrial Economics* **55**, 37–67.

Hart, O. and J. Moore (1998), 'Default and renegotiation: A dynamic model of debt', *The Quarterly Journal of Economics* **113**, 1–41.

He, P., L. Huang and R.D. Wright (2005), 'Money, banking, and monetary policy', *International Economic Review* **46**, 637–670.

Holmstrom, B. and J. Tirole (1997), 'Financial intermediation, loanable funds, and the real sector', *The Quarterly Journal of Economics* **112**, 663–91.

Holmstrom, B. and J. Tirole (1998), 'Private and public supply of liquidity', *Journal of Political Economy* **106**, 1–40.

Houston, J.F. and S. Venkataraman (1996), 'Liquidation under moral hazard: Optimal debt maturity and loan commitments', *Journal of Banking & Finance* **20**, 115 – 133.

Institute for International Monetary Affairs (2004), 'Settlement systems of east asian economies', Japanese Ministry of Finance.

International Monetary Fund (2008), 'Global financial stability report: April 2008'.
URL: *http://www.imf.org/external/pubs/ft/gfsr/2008/01/index.htm*

Ivashina, V. and D.S. Scharfstein (2010), 'Bank lending during the financial crisis of 2008', *Journal of Financial Economics* **97**, 319–338.

Jokipii, Terhi and Alistair Milne (2011), 'Bank capital buffer and risk adjustment decisions', *Journal of Financial Stability* **7**, 165–178.

Jovanovic, B. and M. Ueda (1997), 'Contracts and money', *Journal of Political Economy* **105**.

Jurgilas, M. and A. Martin (2010), 'Liquidity-saving mechanisms in collateral-based rtgs payment systems', Bank of England Working Paper.

Kahn, C., J. McAndrews and W. Roberds (2003), 'Settlement risk under gross and net settlement', *Journal of Money, Credit and Banking* **35**, 591–608.

Kahn, C.M. and W. Roberds (2009), 'Why pay? an introduction to payments economics', *Journal of Financial Intermediation* **18**, 1–23.

Kanatas, G. (1987), 'Commercial paper, bank reserve requirements, and the informational role of loan commitments', *Journal of Banking & Finance* **11**, 425 – 448.

Kashyap, A.K. and J.C. Stein (2000), 'What do a million banks have to say about the transmission of monetary policy?', *The American Economic Review* **90**, 407–428.

Kashyap, A.K., R. Rajan and J.C. Stein (2002), 'Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking', *Journal of Finance* **57**, 33–73.

Klee, E. (2010), 'Operational outages and aggregate uncertainty in the federal funds market', *Journal of Banking and Finance* **34**, 2386–2402.

Kohn, M.G. (1999), 'Early Deposit Banking', Working Paper.

Loukoianova, E., S.N. Neftci and S. Sharma (2006), 'Pricing and hedging of contingent credit lines', IMF Working Paper.
URL: *http://www.imf.org/external/pubs/ft/wp/2006/wp0613.pdf*

Manning, M.J. and M. Willison (2006), 'Modelling the cross-border use of collateral in payment systems', Bank of England Working Paper.

Martin, A. and J. McAndrews (2008), 'Liquidity-saving mechanisms', *Journal of Monetary Economics* **55**, 554–567.

Meh, C.A., Quadrini. V. and Y. Terajima (2010), 'Limited nominal indexation of optimal financial contracts', Working Paper.

Merrouche, O. and J. Schanz (2010), 'Banks' intraday liquidity management during operational outages: Theory and evidence from the uk payment system', *Journal of Banking and Finance* **34**, 314–323.

Mills, D. and T. Nesmith (2008), 'Risk and concentration in payments and securities settlement systems', *Journal of Monetary Economics* **55**, 542–553.

Mosebach, M. (1999), 'Market response to banks granting lines of credit', *Journal of Banking & Finance* **23**, 1707–1723.

Renault, F., W. E. Beyeler, R. J. Glass, K. Soramaki and M. L. Bech (2009), 'Congestion and cascades in interdependent payment systems', Working Paper.

Rochet, J.C. (2007), 'Competing payment systems: Key insights from the academic literature', Toulouse School of Economics and IDEI mimeo.

Rochet, J.C. and J. Tirole (2002), 'Cooperation among competitors: some economics of payment card associations', *RAND Journal of Economics* **33**, 549–570.

Rochet, J.C. and J. Tirole (2003a), 'An economic analysis of the determination of interchange fees in payment card systems', *Review of Network Economics* **2**, 69–79.

Rochet, J.C. and J. Tirole (2006), 'Two-sided markets: A progress report', *RAND Journal of Economics* **37**, 645–667.

Rochet, J.C. and J. Tirole (2009), 'Credit card interchange fees', European Central Bank Working Paper.

Rochet, J.C. and J. Tirole (2011), 'Must-take cards: Merchant discounts and avoided costs', *Journal of the European Economic Association* **9**.

Schanz, J. (2009), 'How do different models of foreign exchange settlement influence the risks and benefits of global liquidity management?', Bank of England Working Paper.

Sufi, A. (2009), 'Bank Lines of Credit in Corporate Finance: An Empirical Analysis', *Review of Financial Studies* **22**, 1057–1088.

Telyukova, I. and R. Wright (2008), 'A model of money and credit, with application to the credit card puzzle', *Review of Economic Studies* **75**, 629–247.

Wright, J. (2003), 'Optimal card payment systems', *European Economic Review* **47**, 587–612.

Wright, J. (2004), 'Determinants of optimal interchange fees in payment systems', *Journal of Industrial Economics* **52**, 1–26.

Yun, H. (2009), 'The choice of corporate liquidity and corporate governance', *Review of Financial Studies* **22**, 1447–1475.