ESSAYS ON FINANCIAL CONTRACTS AND BUSINESS CYCLES

Alfred James Michael Duncan

BSc BCA

Submitted in fulfillment of the requirements for the degree of PhD in Economics.

Adam Smith Business School

College of Social Sciences

University of Glasgow

September 2015

ABSTRACT

This dissertation studies the intersection between the sharing of individual specific risks and business cycle risks. Individual specific or idiosyncratic risk sharing is typically hampered by moral hazard, and in Chapter 2 we propose a new theory of debt finance as an effective mechanism for sharing idiosyncratic risks. But business cycle or systemic risk sharing is also affected by the means of idiosyncratic risk sharing. Departures from full systemic risk sharing can dampen the incentive compatibility constraint allowing a greater degree of idiosyncratic risk sharing (Chapter 1). Entrepreneurs' productive risk can quickly transform into low employment, as wages fall below marginal revenue products of labour (Chapter 3). Market prices for systemic risk insurance do not necessarily internalise balance sheet externalities, resulting in excessive swings in leverage and factor market wedges of inefficiency (Chapter 4). Sometimes, agents have private information about the risks faced by their projects, and how they correlate with the broader economy. When this is the case, optimal systemic risk sharing arrangements must allocate business systemic risk in a way that deters entrepreneurs from herding among their peers (Chapter 5).

CONTENTS

List of Figures

List of Tables

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Alfred Duncan

INTRODUCTION

Émile Zola's novel *Germinal* is a story of hardship and tensions between workers and capital owners in Post Revolutionary France. It is clear that even during 'good' times, the workers of Village 240 in the coal mining town of Montsou have little to eat and suffer terrible, dangerous working conditions. Gross inequality between workers and capitalists appears timeless in Montsou. Yet it is when the price of coal falls, bringing down wages and employment that brings to light important questions about how this fall in coal prices should be shared across workers, entrepreneurs and rentier shareholders. And it is these questions of fairness that are amplified when workers and entrepreneurs alike have near exhausted any reserves and lines of credit. The interaction between the sharing of idiosyncratic and systemic risks is the main theme of this dissertation. What we show is that the problem of how to share business cycle risk is deeply tied to the problem of sharing idiosyncratic risks, when idiosyncratic risk sharing is made difficult by information asymmetries relating to effort and individual specific luck.

The families in Montsou experience a variety of idiosyncratic or individual specific financial risks in addition to aggregate, or systemic risks, associated with commodity prices. While there is some insurance within multiple income households, the Maheu family, who are followed closely throughout the novel, earns income solely from coal mining. Individual miners enjoy different compensation packages, subject to the risk of finding high quality coal in their allocated seams. They also face the very real risk of unemployment. Sickness and injury pose further serious risks to household consumption, with only partial insurance provided by employers. Credit markets appear to be open to households and entrepreneurs alike to help smooth consumption in the wake of income risks. But the debts incurred by households and entrepreneurs do not offer any insurance against aggregate or systemic risks. Furthermore, under the weight of debts drawn during the previous recession in the case of the Maheu family, or during the boom in the case of the entrepreneur Deneulin, both the Maheu family and Deneulin find themselves unable to increase their borrowing to maintain consumption and investment during the current downturn—or at least unwilling to bear the potential risks associated with further borrowing.

The Maheu family's mining yield depends on individual specific luck associated with the quality of the seam allocated to them, it also depends on their efforts exerted at the coal face. It is straightforward to see why their compensation scheme remunerates them according to the quantity and quality of the coal in their tubs, even though this leaves them highly exposed to idiosyncratic risk. A constant wage income independent of the amount of coal returned would clearly result in a moral hazard problem with imperfect monitoring at the coal face.

Similarly, the entrepreneur Deneulin is the sole shareholder in his firm. His income and net worth are heavily dependent on the revenues earned at his mines, after returning fixed interest and principal payments to creditors. Again, there is a clear moral hazard problem explaining exactly why creditors would be unwilling to share a large amount of idiosyncratic risk with Deneulin, who has a large degree of control over the investment projects and labour bargains that are crucial determinants of his firm's success or failure.

But what is not so obvious is why Deneulin and the Maheu family appear to have limited access to insurance against systemic risks which are outside their control but to which they are heavily exposed, apparently moreso than other agents in Zola's novel. Neither the Maheu family nor the entrepreneur Deneulin are 'responsible' for the falls in the prices of coal and other commodities which so severely reduce incomes earned in Montsou. There is no obvious moral hazard problem preventing them from writing systemic risk insurance contracts with other agents in the economy who are in a better position to weather the current storm.

In Zola's novel, it appears that markets for insurance against systemic risks could be open. Labour contracts are reasonably sophisticated and the long term nature of mining employment relationships within Montsou certainly allow for the possibility of risk sharing between workers and shareholders. The wealthy rentier Léon Grégoire has ample opportunity to arrange a risk sharing agreement with his cousin Deneulin, but resists these opportunities apparently in accordance with his very low tolerance for risk. In sum, it is not clear that the individual actors in Montsou are constrained by arbitrary restrictions on contracts that prevent any risk sharing that would be possible under a richer set of contracts.

So why do systemic risk sharing arrangements in Montsou appear to be neither fair nor efficient? Would more effective risk sharing have reduced the massive costs of hunger, unemployment and civil unrest provoked by the fall in coal prices?

The first chapter of this dissertation considers the relationship between idiosyncratic and systemic risk sharing in a simple two-period endowment economy. A straightforward but intriguing result is shown: information asymmetries relating to individual specific risks do in fact hamper the sharing of systemic risks under constrained efficient allocations. The intuition behind this result is that any agent falsely claiming to have suffered bad luck, and hiding away their ill-gotten gains, has a greater demand for exposure to business cycle risk than truthful unlucky agents. Optimal risk sharing contracts can use this differentiated demand for protection or exposure to business cycle risk to identify truth-telling and misreporting agents. Partially closing markets for business cycle risk further exacerbates the difference between the demands of fraudulent and truth-telling agents, helping identify those who are lying about their luck. This allows increased sharing of individual specific risks across agents without encouraging fraud.

Straight away, in a simple endowment economy model with a basic private information constraint, we can see how information asymmetries or *frictions* that hamper the sharing of individual specific risks will also affect the sharing of business cycle or systemic risks that are commonly observed. That is, they are not obviously affected by information asymmetries.

The next step in our analysis is to develop a richer theory of debt contracts as a response to information asymmetries about individual specific risks. The first chapter assumes that individual specific risks were private information not available at any cost to monitors or outsiders, an assumption that quickly yields simple non-contingent debt-like contracts as an optimal contract used by individual agents to manage idiosyncratic risks through delaying or bringing forward consumption between the current and future periods.

But, it is perhaps unrealistic to think of private information as being unobtainable to outside monitors at any cost. In practice, financial contracts typically allow for some form of costly information acquisition by monitors, and these *audits* reveal a useful signal of the veracity of the claims made by the reporting agent. In order to consider how risk sharing evolves throughout the business cycle, we need a model that allows for endogenous variation in the degree of risk sharing between agents, and can explain important features of real life contracts, such as debt, default, and limited liability.

In Chapter 2 we develop such a model. We relax the assumption of strict private information considered in Chapter 1 and we permit an audit technology that reveals an ex post signal testifying to the veracity of the claims made by the monitored agent, at some cost and with imperfect precision. Essentially, as the cost of audit increases, or the signal weakens in precision, the model converges to strict private information as in Chapter 1. What we show is that it is the imprecision of the audit technology, rather than the cost, which is important for explaining the standard form of debt contract. These debt contracts allow for limited risk sharing except following extreme circumstances. Following low reports, an auditing process resembling default is initiated, and the eventual repayment leaves the entrepreneur with strictly positive real wealth. That is, the entrepreneur enjoys limited liability. When the precision of the audit technology is high, optimal contracts resemble equity. These equity contracts exhibit a high degree of risk sharing even if audit costs are relatively high, and even as firms' profits go from good to great. Any marginal increase in revenues results in a marginal increase in repayments, or dividends, to outside investors.

The model considered in Chapter 2 focuses on the funding of productive investments by entrepreneurs, like Deneulin in Zola's novel. Deneulin is the sole shareholder in his firm, having offered an equity stake to M. Grégoire, but having been unable to agree terms. In the end, Deneulin raises all his external finance through debt issuance. This leverage through debt finance leaves Deneulin highly exposed to volatility in coal prices. When coal prices fall, Deneulin struggles to cope with the risk, even following reductions in wages. Eventually, Deneulin decides that he cannot absorb the risks of production, prices and wages, and he sells his firm to The Company, the firm owning all of the other mines in Montsou. Following the sale, wages fall further, even though coal prices appear to have stabilised. It appears as though any wedge that had existed between workers wages and workers' marginal revenue product has increased.

In Chapter 3, we develop a financial macroeconomic model with wage-earning households and credit constrained, risk averse entrepreneurs. Entrepreneurs raise external finance for their contracts via the contracts studied in Chapter 2. We show that the information asymmetries that encourage the use of debt finance result in a time-varying wedge between wages and the marginal productivity of labour. A lack of entrepreneurial risk sharing has turned into an inefficiency wedge in the labour market, dramatically increasing the costs of systemic risk for wage earning households. This wedge tends to be counter-cyclical in response to productivity shocks, smoothing the path of real wages. But the labour wedge is procyclical in response to financial shocks, amplifying their effects on output and employment.

Deneulin's debts do not include contingencies for variation in the price of coal, the key macroeconomic risk his firm is exposed to. A market for systemic risk insurance, either through derivative-like contracts or through contingencies written into standard debt finance contracts would allow Deneulin to insure his firm against volatile swings in coal prices, if he desired to do so.[1] In Chapter 4, we study the role of markets for systemic risk insurance in our macroeconomic framework. Could these markets help internalise the costs to wage earning households of the time varying labour wedge, and discourage large swings in entrepreneurs' leverage? The short answer is that systemic risk markets in our environment can actually make things worse. The prices of systemic risk insurance do not internalise the balance sheet externalities that amplify business cycles in our model. The invisible hand prioritises individual-level consumption insurance over business cycle stabilisation, allowing the labour and capital market wedges of inefficiency to vary excessively. Indeed, when systemic risk markets are open in our model, risk tolerant entrepreneurs tend to sell insurance against downturns, increasing their exposure to these business cycles and thereby the volatilities of their net wealth and leverage.

It was not Deneulin's fault that coal prices fell. Yet it cost him his fortune and resulted in the closing of his mines which paid workers higher wages than the mines owned by his competitors. A bailout could have saved Deneulin, and helped to keep his mines open. Possibly even a bout of high inflation could have saved him through a real reduction in his nominal debts. Certainly at a first glance this looks like it could be helpful in Montsou, protecting Deneulin from events out of his control and encouraging investment and production throughout the downturn. On the other hand, one can't help but wonder whether Deneulin was acting in the interests of the village in the first place when he levered his considerable net worth toward an investment in the re-opening of old coal mines. Most of the working families in Village 240 relied solely on wages from coal mining. If Deneulin had instead invested in a different industry, providing jobs whose wages were not tied to the volatile price of coal, then the insurance benefits of diversification in household income sources could perhaps have been beneficial. The reader cannot be certain that in Montsou, Deneulin's investment in coal mining was inefficient, and that the invisible hand was not allocating investment resources toward their best use. But what we do show in Chapter 5 is that a monetary policy of nominal output targeting can have precisely the effect of preventing the invisible hand from allocating resources efficiently into diversi-

[1] Of course, these same markets would allow Deneulin to increase his exposure to the volatility of coal prices, if he so desired.

fied projects. While Deneulin cannot control the price of coal, it was certainly his choice to invest in coal production, knowing the risks this entailed. Shielding Deneulin from the volatility of coal prices would have provided further incentives ex ante to devote additional resources into herding with other coal producers in Montsou, rather than investing in projects that would have provided diversification for the wage earners of Montsou.

Before moving on to the main chapters of this dissertation, we review some of the main stylised facts and puzzles motivating this analysis, before considering the role of policy intervention in incomplete-market economies suffering from difficulties with risk sharing.

STYLISED FACTS AND MAJOR PUZZLES

We start with the prevalence of debt finance in modern (and also historical) economies.

Fact 1. *The widespread use of debt finance is an important determinant of economies' responses to shocks, and disturbances emerging from within the financial sector are an important source of business cycle fluctuations.*

Schularick and Taylor (2012) document the paths of *credit booms gone bust*. Rapid increases in credit growth they argue leave economies vulnerable to shocks that have devastating consequences. The financial accelerator model of Bernanke, Gertler, and Gilchrist (1999) and the credit cycles model of Kiyotaki and Moore (1997) have been influential in helping to form a consensus around the interaction between debt finance and business cycle volatility. These papers show that when entrepreneurs finance their projects with debt, fluctuations in asset prices cause volatility in entrepreneurs' leverage ratios, raising financing costs and reducing the demand for investment. In turn, the required deleveraging and reduction in investment further reduces asset prices, amplifying the cycle.

Extensions of these models have also shown that shocks from within the financial sector, that is, disturbances that affect the intermediation of credit directly are also important determinant of business cycle volatility (Christiano, Motto, and Rostagno, 2014, Nolan and Thoenissen, 2009).

Fact 2. *During periods of financial stress, the prevailing distribution of burden sharing of business cycle risk is often seen as unfair and/or inefficient.*

In Zola's novel *Germinal* it is not until the recession that the seeds of socialist revolution are sown. This is in spite of the fact that the inequalities in consumption during Zola's recession appear to have changed little since the preceding boom—both the capital owners and wage earners have lost a large amount of their income. This highlights the political tensions amplified during episodes of financial stress, and these tensions are further exacerbated when political actions appear to favour one group over another. That is, when Wall Street receives a bailout but Main Street does not.

Mian and Sufi (2015) and Shiller (2008) among others have argued that existing market arrangements regarding mortgage finance leave homeowners unfairly exposed to volatility in broader trends in real estate prices. This is not just a fairness argument but an efficiency argument: These authors argue that there are opportunities for mutual benefit between borrowers and lenders if mortgage contracts were linked to house prices. In other words, the state-contingent intertemporal marginal rates of substitution for homeowners differs from non-homeowners. Homeowners would be willing to give up some consumption during periods of rising house prices in turn for insurance against falls in house prices, and other households would be willing to take the other side of this trade, giving up some consumption in downturns in order to enjoy greater consumption during periods of rising house prices.

On the other hand, during the Global Financial Crisis of 2007-08, several Governments including the United States and the United Kingdom have extended financial support directly to ailing firms, particularly in the banking sector but also in the auto industry in the case of the United States. The presumption behind this financial support being that the weight of losses on these firms would result in inefficient outcomes through unemployment, balance sheet externalities and potential disruption of the monetary transmission mechanism.[2] Even Adam Smith argued that the *invisible hand* could not be relied on to ensure the efficient allocations of systemic risk, calling for some regulation of lending and deposit issuance (Smith, 1776, II (ii) 94).

Fact 3. *During periods of financial stress, there is high unemployment.*

Perhaps the most striking feature of financial crises is the labour market response.

[2]Calomiris and Khan (2015) provide an excellent discussion of the stated objectives of the Troubled Asset Relief Programme (TARP), which was the primary vehicle for US bank recapitalisation during the crisis. Goolsbee and Krueger (2015), provide a useful and broadly favourable review of the rescue of the US automotive industry.

During the Great Depression, the unemployment rate in the United States peaked at 26% in May 1933.[3] During the recent Global Financial Crisis, unemployment in the United States peaked at 10% in October 2009.[4]

It is helpful to think about the employment and wage responses to financial stress through the idea of the *labour wedge*, the difference between the real marginal product of labour (the economy's work-consumption marginal rate of transformation) and workers' consumption leisure marginal rate of substitution:

$$\frac{-U_N(C,N)}{U_C(C,N)} = (1-\tau_N)F_N(K,N).$$

Here, the left hand side is the consumption-leisure marginal rate of substitution, τ_N is the labour wedge and $F_N(K,N)$ is the marginal product of labour, which is the marginal rate of transformation from work to consumption. Defining the labour wedge in this way, we can see that the effects of the labour wedge resemble variations in payroll and labour income taxes. Chari, Kehoe, and McGrattan (2007, Figure 1) estimate that the increase in the labour wedge between 1929 and 1934 is over 30%.[5]

Fact 4. *The unfavourable tax treatment of equity finance relative to debt finance cannot explain the prevalence or the features of standard debt contracts.*

Throughout this study, we motivate the use of debt finance as an alternative to equity finance by appealing to information asymmetries. These information asymmetries make it difficult to enforce the the link between dividend repayments and good or bad luck associated with equity contracts. An alternative motivation for the prevalence of debt finance is the preferable tax treatment enjoyed by debt finance over equity finance for incorporated firms in many countries.[6] Typically, firms' earnings attract corporate income tax before their disbursement as dividends to shareholders, whereas debt interest payments can be deducted from revenues when calculating taxable income.

The disadvantageous tax treatment of equity finance relative to debt finance must

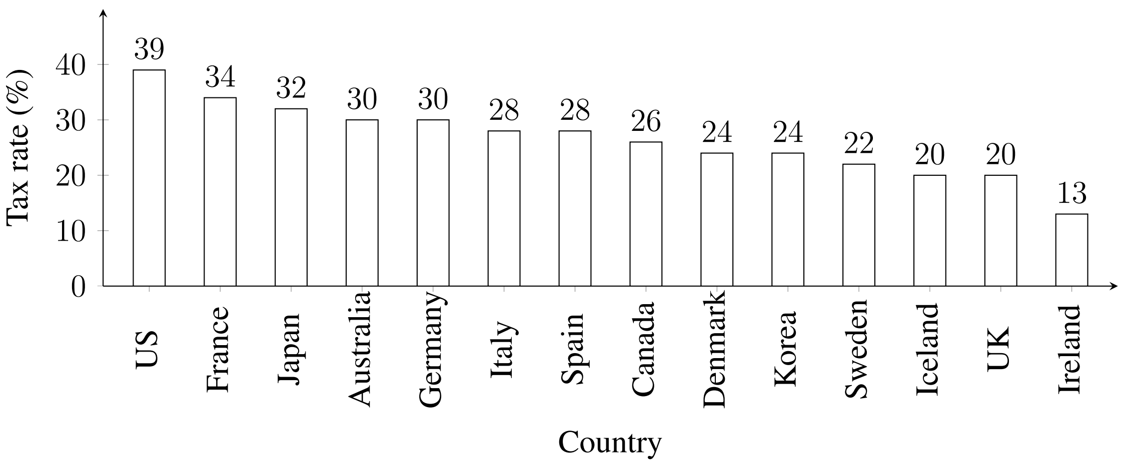
[3]Source: St. Louis Federal Reserve Fred Database. Unemployment Rate for United States (`M0892AUSM156SNBR`).

[4]Source: St. Louis Federal Reserve Fred Database. Civilian Unemployment Rate (`UNRATE`).

[5]Note that Chari, Kehoe, and McGrattan (2007, Figure 1) refer to $(1-\tau_N)$ as the labour wedge. Throughout this study, we will refer to τ_N as the labour wedge, which is also a common definition, and is more intuitive in our opinion. Under our definition, it will typically be the case that when the labour wedge is large in absolute terms, actions and allocations are further from the first-best efficient allocations.

[6]Notably, this is the motivation employed by Jermann and Quadrini (2012), who we will refer to frequently in later chapters.

Figure 1: Statutory corporate income tax rates. Selected OECD countries. For the tax year ending in 2015. Includes local government taxes where applicable (Source: OECD).



surely influence firms' financing decisions in favour of debt finance. Still, there are important reasons to believe that taxation plays only a minor role in explaining the prevalence of debt finance. Corporate income tax rates across countries vary considerably as shown in Figure 1. Debt remains the predominant form of external finance contract in countries with little or no debt-equity tax distortion. Further, the recent financial crisis starting in 2007 has not spared countries with small debt-equity tax distortions. In Ireland, severely hit by the global financial crisis, the corporate income tax rate is only 13%. In New Zealand, a generous company tax imputation credit scheme is employed, dramatically reduces any tax distortions between debt and equity finance. Nevertheless, New Zealand's equity markets are not very well capitalised by international standards (see for example Cameron, 2007) and the New Zealand economy suffered from the failure of highly-leveraged 'finance companies' in 2007-08 (NZHR, 2011).

Unlike differences in tax treatment, the information asymmetries including those we study in this volume do an excellent job not only of explaining why debt finance can be favourable to equity, but also in explaining some of the specific features of real world debt contracts, including default, the limited liability company, and interest rate spreads. The specific features of debt contracts, which determine to what extent risk is shared between agents across states, are crucial determinants of the extent to which debt finance provides an efficient mechanism for project finance and intertemporal consumption insurance.

If the prevalence of debt finance were the result of tax distortions alone, our policy prescription would be straightforward: remove the tax distortion. Immediately, agents would issue equity contracts and repay their debts, and the economy would revert to first best efficient allocations and actions. If the prevalence of debt finance is in part the rational response of agents to information asymmetries, then the policy problem is much more tricky.

Puzzle 1. *What are the important features of a microfoundation for debt finance?*

There remains much debate over the important features of a microfoundation for debt finance. In this study we focus on approaches to the microfoundation for debt that rely on some form of asymmetric information about the revenue of the firm. When this information is completely private, and only idiosyncratic risks are present, then it is widely known that the optimal risk sharing contract takes the form of a non-contingent debt contract (Cole and Kocherlakota, 2001). The seminal paper of Townsend (1979) showed that the introduction of an auditing technology can motivate what we might think of as a *standard debt* contract: Following moderate and high reports, the borrower makes a fixed repayment to the lender. Following low reports, the borrower is audited and a lower repayment is returned to the lender—a process that resembles bankruptcy. At least this is the case if lenders cannot use stochastic audit strategies.

If lenders can randomise their audits, then it turns out the story changes. Standard debt is no longer optimal. It becomes worthwhile to audit even relatively high reports. A low probability of audit combined with a large penalty can have a big deterrent effect preventing misreporting (Border and Sobel, 1987, Mookherjee and Png, 1989).

The contracts derived in the aforementioned studies require the lender to be able to commit to an auditing strategy ex ante, knowing that ex post, all borrowers are truth-telling and the audit costs incurred will be wasted. Krasa and Villamil (2000) show that relaxing this commitment assumption can restore standard debt contracts. The decision to audit is now taken ex post: if the expected returns to auditing are high, these audits will be undertaken with certainty.

It is surely the case that some lenders cannot credibly commit to audit schedules, but it seems unlikely that banks would be unable to make and sustain this commitment through delegation or some form of reputation equilibrium (Melumad and Mookherjee, 1989). We assume in this study that lenders can commit ex ante to ex post audit strategies, and show in Chapter 2 that standard debt contracts can indeed be optimal if costly audits provide only a noisy signal, and if borrowers are risk averse. The noise associated with the audit signal provides an incentive to use contracts with lower penalties where possible, limiting the deterrent effect from stochastic audit strategies and promoting the deterministic audit strategies that characterise debt. As the audit technology improves in its precision, the costs of audit errors decline, more severe penalties can be applied, audits are used more sparingly over a greater range of reported incomes and contracts converge to a form that resembles equity.

Another important approach towards a microfoundation for debt include the incomplete contracts literature, where debt contracts specifying a fixed repayment serve as a starting point for renegotiation following outcomes that could not have been predicted ex ante (Hart and Moore, 1988). Aghion and Bolton (1992) study optimal incomplete contracts where information about managers' effort is privately observable. Optimal contracts allocate control of assets and risk with managers, in a similar way to debt finance. One of the key advantages of the incomplete contracts approach is that it can explain the transfer of control rights from shareholders to debtholders in default.

Jensen (1986) argues that large amounts of free cash flow offers an opportunity to managers to extract pecuniary and non-pecuniary benefits from the firm. Debt finance restricts free cash flow, allocating some of these funds to creditors and thereby dampening the associated moral hazard problem. Ellingsen and Kristiansen (2011) describe a model with unobservable managerial effort and show that the optimal financial structure in this model can include a mixture of debt and equity.

Puzzle 2. *What are the mechanisms through which financial stress affects the* labour wedge*, the wedge of inefficiency between the economy's work-production marginal rate of transformation and workers' consumption-leisure marginal rates of substitution?*

In a frictionless environment, the labour wedge τ_N will equal zero. Households will equate their consumption-leisure marginal rate of substitution to the real wage rate, and firms will equate their marginal labour productivity to the same real wage rate. Denoting the real wage rate by W, our condition for first best efficiency is the following:

$$\frac{-U_N(C,N)}{U_C(C,N)} = W = F_N(K,N).$$

We can think of frictions acting on both sides of this market for labour. One the supply side preventing unemployed workers from entering jobs at acceptable wage rates, on the demand side preventing firms from offering wages that are as high as the marginal product of labour.

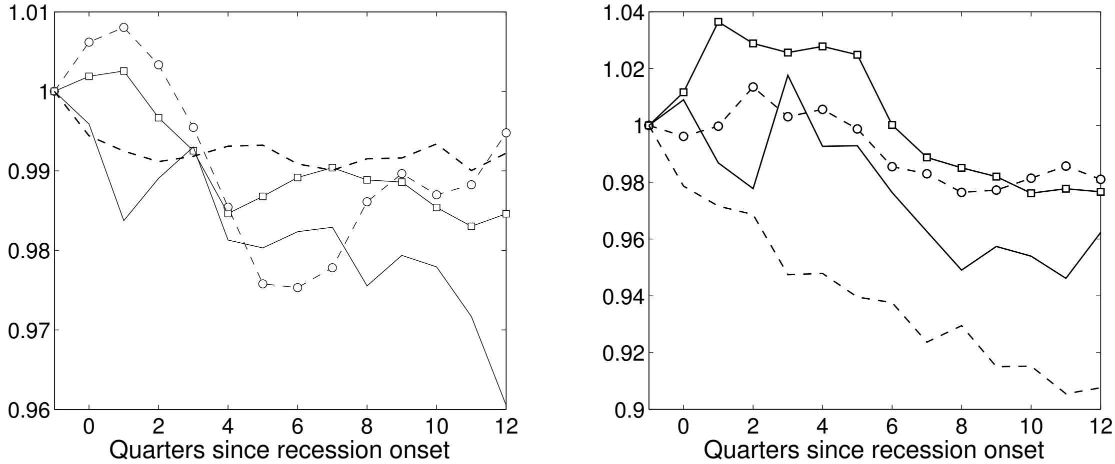
Figure 2 presents two measures describing the cyclical behaviour of the inverse of the demand side of the labour wedge, the ratio of real wages to marginal labour productivity. The left hand panel presents the path of the labour share of national income $\left(\frac{WN}{F(K,N)}\right)$, and the right hand panel presents the ratio of real wages to (average) nonfarm labour productivity $\left(\frac{W}{F(K,N)/N}\right)$. Differences in the two series result from variation in the sample of the economy considered and in methods of measurement. Under the assumption of a Cobb-Douglas production function ($F(K,N) = AK^{\alpha}N^{1-\alpha}$), average labour productivity is proportional to marginal labour productivity ($\alpha F(K,N)/N = F_N(K,N)$) and these series present the paths of $W/F_N(K,N)$ over these four recessions. We focus on the demand side of the labour wedge because this is where our model described in Chapter 3 predicts the financial friction to directly affect the labour wedge. Both measures show modest short run decreases in the labour wedge at the onset of the 1974 and 1981 recessions (or increases) in the ratio of wages to labour productivity), followed by increases in the longer run. Both measures show large persistent increases in the labour wedge in the recent 2001 and 2008 recessions. This behaviour of the demand side of the labour wedge is broadly consistent with the predictions of our model, with negative technology shocks producing a short run increase and long run fall in the labour wedge as in the 1974 and 1981 recessions, and negative financial shocks producing a large increase in the labour wedge, as in the 2001 and 2008 recessions.

Table 1 gives a mapping of the key frictions that together form the labour wedge, broken up into price setting rigidities, matching frictions and financial frictions.

The first class of frictions is price setting rigidities, also referred to as *sticky prices* or New Keynesian frictions.[7] When product or labour markets are imperfectly competitive, and firms and workers face non-trivial costs of price or wage adjustment, markups will be time varying. Retail markups are the difference between the marginal revenue produce of labour and marginal labour costs or wages per unit of effective labour, and wage markups are the difference between real wages and households' consumption-leisure marginal rate of substitution. That is, when retail markups are large, wages offered by firms are rel-

[7]See for example, the models described by Galí (2008) and Woodford (2003)

Figure 2: Labour income and productivity in US NBER recessions (Recession starting 1974Q1 dashed o, 1981Q4 solid □, 2001Q2 dashed, 2008Q1 solid). Left hand panel: Labour share of National Income (BEA NIPA Table 2.1). Right hand panel: Nonfarm real wages relative to labour productivity (BLS `PRS85006102`, `PRS85006092`. Nominal wages deflated by the consumer price index, St. Louis Federal Reserve FRED Database `CPIAUCSL`). Both series normalised to one at the onset of the respective recessions.



atively low. When wage markups are large, wages offered by firms are above the level that would be acceptable to an unemployed worker, but the quantity of employment demanded is low, resulting in unemployment—unemployed workers are willing to work at current wage rates but cannot find a job. Importantly, while steady state markups are the result of market power, and desirable to the individuals holding that power (retail markups are desirable to firms and wage markups are desirable to workers) variation in both retail and wage markups over the business cycle is mostly undesirable and the consequence of costs associated with changing prices. When New Keynesian markups are high, this is because firms and workers would like to offer lower prices and wages respectively, but are prevented from doing so.

The second class of frictions is matching frictions and follows the literature on vacancy posting and worker-employee matching pioneered by Diamond (1982) and Mortensen and Pissarides (1994). In typical labour search and matching models, firms' vacancy postings suffer a real resource cost, which must be reclaimed in expectation through the difference between real wages and marginal labour productivity. When the unemployment rate is low, the labour market is considered to be *tight* and it is difficult to find suitable workers for particular job openings, this increases the expected resource cost to firms of hiring additional workers, and results in a rationing equilibrium where not all unemployed workers can be matched in any given period, and wages can remain above workers' consumption-leisure marginal rates of substitution.

Table 1: Breaking down cyclical variation in the labour wedge

	Class of friction		
	Price setting frictions	Matching frictions	Financial frictions
Demand side			
Marginal labour productivity			
-	Retail markups		
-		Vacancy costs	
-			Risk premium
= *Real wages*			
Supply side			
Real wages			
=	Wage markups		
+		Vacancy rationing	
+ *Consumption-leisure marginal rate of substitution*			

Financial frictions can also directly affect the labour wedge. In Jermann and Quadrini (2012), this is a consequence of tax distortions that encourage entrepreneurs to use imperfect debt products to finance their wage bill. In Chapter 3, we show that wages offered by entrepreneurs will be lower than labour's marginal product if risk averse entrepreneurs cannot fully pass on idiosyncratic productivity risks to outside investors. That is, if entrepreneurs must finance their projects with debt.

These three classes of frictions interact with each other. Financial frictions acting solely on the intertemporal wedge can still create and amplify volatility in the labour wedge when New Keynesian price setting frictions are present through the effects of the financial friction effects on the natural rate of interest and real wages in the flexible price counterfactual.[8] This interaction between financial and New Keynesian frictions goes in both directions. As we show in Chapter 3, New Keynesian price setting frictions can slow down the adjustment of real interest rates to shocks, making it difficult for firms to restore their financial health during periods of financial stress. Financial frictions in the model we describe in Chapter 3 will also drive an efficiency wedge in vacancy posting when search and matching frictions are present.

[8]The intertemporal wedge between the consumption-savings marginal rate of substitution and the gross marginal product of capital.

Puzzle 3. *What is the extent to which agents can hedge against business cycle risks?*

Ideally, business cycle risk should be allocated such that (1) those agents who are more risk tolerant bare a greater share of business cycle risk, providing insurance to those who are less risk tolerant, and (2) those agents whose actions have a greater impact on the extent of business cycle risk themselves bear a significant portion of that business cycle risk, such that their incentives are aligned with an optimal solution to the trade-off between systemic risk and expected income if such a tradeoff exists.

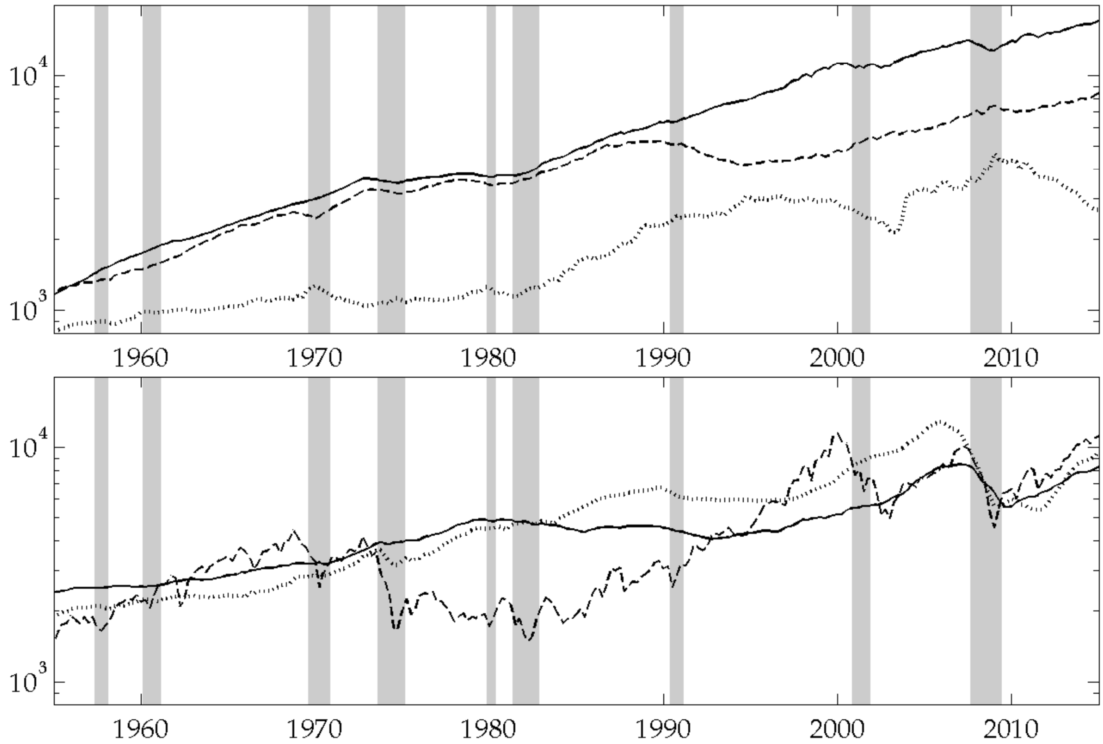
It can be difficult to imagine how real world contractual arrangements can achieve this allocation of systemic risk. There don't appear to be markets open for trade in Arrow securities contingent on a full set of all possible future paths for GDP, inflation or productivity. At least, it is not clear that individual households hold sophisticated portfolios of these assets. The debt contracts that dominate intertemporal trade between across households and between households and firms do not typically allow for automatic writedowns of principal or interest in response to systemic risks.

One plausible interpretation of the lack of trade in GDP or productivity linked securities is that in practise, these systemic risk insurance markets are closed. Perhaps the necessary contracts are too complicated to define and enforce, so agents respond by just using relatively simple debt contracts that do a good job of sharing idiosyncratic risk in the presence of information asymmetries.

An alternative interpretation, and one that we will seriously consider throughout this work, is that the same simple that do a good job of managing idiosyncratic risk are also effective at managing systemic risk. Under this interpretation, the reason we see little trade in complicated securities contingent on business cycle risks is not because these markets are closed but instead that there is actually little opportunity for mutually beneficial trade in these securities.

Figure 3 presents the recent historical evolution of US household and non-profit asset holdings. The top panel displays holdings of relatively safe assets: pension assets (which are typically diversified portfolios of debt and equity instruments), money-like instruments and credit securities. The lower panel presents holdings of more risky assets: equity in non-corporate firms, equity in corporate firms and equity in real estate. The picture that emerges from Figure 3 is that households do in fact have considerable choice

Figure 3: US Household and non-profit asset holdings, major components (2005 $USD billions). Top panel: Pension assets (—), currency, deposits and money market mutual fund shares (- -), credit securities (· · ·). Bottom panel: Equity holdings in non-corporate firms (—), Equity holdings in corporate firms (- -), Equity held in real estate (· · ·). Source: Federal Reserve Board Z.1 Tables. Shaded bars indicate NBER recession dates. Nominal series converted into 2005 dollars via deflation by the Consumer Price Index, St. Louis Federal Reserve FRED Database CPIAUCSL)



over their exposure to real estate, stock market and duration risks, and their portfolios of these assets evolve considerably over time. Even without markets for derivative contracts contingent on house price indexes, individual households certainly have some choice over how much real estate risk they are exposed to. The same goes for other systemic risks realised through equity prices.

Puzzle 4. *When markets for individual risk sharing are hampered by moral hazard, is there a conflict between competitive and socially optimal arrangements for systemic risk sharing?*

Mian and Sufi (2012) argue that in the context of housing risk, there is no conflict between competitive and socially optimal arrangements for systemic risk sharing. In their view, the problem is that markets for housing market risks are simply not open. If they were, leveraged mortgage borrowers with high marginal propensities to consume out of income would not be so exposed to house price volatility, and house price fluctuations would not have such dramatic effects on aggregate demand.

Farhi and Werning (2013) suggest that systemic risk sharing markets cannot restore constrained efficiency when markets suffer from New Keynesian price setting rigidities and consequently from aggregate demand externalities. Systemic risk insurance prices do not take into consideration the effects of insurance transfers on the demands for goods in markets that are heavily distorted by price setting rigidities. When monetary policy is constrained, the equilibrium allocations can exhibit large inefficiency wedges. This is not to say that allocations would improve if systemic risk markets were closed in Farhi and Werning (2013)'s model, but it does mean that there is some conflict between agents' trade in systemic risk insurance markets and constrained efficiency.

Carlstrom, Fuerst, and Paustian (2014) study systemic risk sharing markets in the financial accelerator model of Bernanke, Gertler, and Gilchrist (1999). Krishnamurthy (2003) and Nikolov (2014) introduce systemic risk sharing markets into the credit cycles model of Kiyotaki and Moore (1997). Each of these studies shows that the introduction of systemic risk markets eliminates or nearly eliminates the financial amplification mechanisms that amplify business cycles in these frameworks. There is no conflict between competitive trade in systemic risk markets and socially efficient outcomes, and indeed the resulting equilibria when systemic risk markets are open are approximately identical to the first best efficient outcomes. In these models, when systemic risk markets are open, entrepreneurs buy insurance against downturns. The resulting flows of wealth from households to entrepreneurs during periods of financial stress immediately restore entrepreneurs' balance sheets, dampening the volatility in leverage that amplifies recessions when systemic risk markets are closed.

But this raises a few questions. First, the argument of Mian and Sufi (2012) appears to suggest that efficient systemic risk insurance trade would result in flows toward indebted households during downturns, while the analysis of Carlstrom, Fuerst, and Paustian (2014) suggests that these transfers would flow toward entrepreneurs in downturns. Consider the case of AIG, the insurance group that was spectacularly rescued by an equity injection from the US government in 2008. Prior to the crisis, AIG had dramatically increased their exposure to the volatility of US house prices through sales of credit-default-swaps linked to US subprime mortgages. Effectively, AIG sold insurance against volatility in US house prices, the opposite of what would be predicted by Carlstrom, Fuerst, and Paustian (2014), Krishnamurthy (2003) and Nikolov (2014). In Chapter 4, we show that the aforementioned studies can be thought of as special cases within the broader financial macroeconomic framework. In general, entrepreneurs' demand for systemic risk insurance does not necessarily align with social efficiency. In periods of financial stress, entrepreneurs' return to wealth is high, but is also more risky, as it is more difficult to pass on this risk to lenders during these periods. Ex ante, entrepreneurs heavily discount the high returns earned during periods of financial stress, reducing their demand for systemic risk insurance. There is a conflict between privately optimal trade in systemic risk insurance markets and broader efficiency objectives.

In Chapter 1 we outline an alternative mechanism through which systemic risk insurance conflicts with idiosyncratic risk insurance. In an endowment economy, we show that misreporting agents receiving high idiosyncratic shocks have a different demand for systemic risk than truth-telling low reporting agents, and this information can be used to identify these misreporting agents. Perturbing the allocation of systemic risk away from the competitive allocation can relax the truth-telling constraint on agents receiving high idiosyncratic shocks, allowing for greater sharing of idiosyncratic risks.

A MECHANISM DESIGN APPROACH

The primary methodological approach employed in this work follows the tradition of mechanism design.[9] Specifically, the mechanism design approach to the study of financial markets is the search for environments where simple real world institutions emerge as a response to information asymmetries and other forms of transactions costs. Technically, this approach takes the starting point of analysis to be arbitrarily sophisticated mechanisms, institutions or contracts that could emerge to implement the optimal allocations given the information asymmetries and transactions costs present. Then, we consider whether the optimal outcomes obtainable through these sophisticated mechanisms could be implemented with simpler, real world contracts and institutions.

It might be useful to compare the mechanism design approach with a popular alternative approach, which is to start from an environment where many markets are exogenously closed, and then to introduce simple financial instruments one by one. This latter approach favours simple contractual forms (such as non-contingent debt) that might appeal to real-world contracts.

[9]Some examples which we have found influential in encouraging this approach include Green and Oh (1991), Kiyotaki (2010) and Wallace (2010).

The key difference between the two approachs is that under the mechanism design approach, simple institutions must emerge as optimal. Under the latter approach, simple institutions can be favoured over more sophisticated institutions by preference for simplicity. This latter approach tends to offer more support for policy interventions, if policy institutions can replicate the payments associated with the missing sophisticated contracts and institutions, and in particular if policies are not restricted to the simplicity assumed to restrict private contracts. Under the mechanism design approach, there is typically a lesser role for policy interventions. Any sophisticated policy introduced by the policy maker is subject to the same information asymmetries that prevented the relevant markets from opening in the first place.

It is important to note this distinction, which we will continue to refer to throughout this work. In our view, neither approach is unambiguously favourable for policy experiments in general, and at times we will employ the second approach where it provides useful insights. It is plausible that in practise, agents prefer to use simple, standardised contracts that may reduce transactions costs not present in our model, and that policymakers may have a comparative advantage in developing state-contingent contracts in some settings.[10] On the other hand, it is certainly the case that in many circumstances, markets are closed as a direct result of information asymmetries or other frictions which cannot be overturned by policymakers.

THE ROLE OF POLICY

Throughout our analysis we will consider two distinct motivations for policy. In some of the models we study, agents will be restriced arbitrarily from trade in securities that are contingent on particular sets of states. For example, agents may be restricted to contracts that are not contingent on commonly observable business cycle risks. In this case, agents may *want* to trade contracts contingent on technology, output or inflation, but cannot, and we may look for policy interventions *contingent on these same risks* that may improve outcomes. Typically, but not always, we can think of these interventions as replicating the payments that would have occured in the missing market.

[10] An example that macroeconomists will be familiar with is the study of monetary policy in sticky-price New Keynesian models. Retailers are exogenously restricted from indexing their posted prices to inflation, but monetary policymakers are free to adjust interest rates in response to inflation. This approach assumes that the monetary policy authority has a comparative advantage in adjusting their control variable in response to the observable variable, inflation.

This motivation for policy intervention can be somewhat unsatisfying. We cannot explain why agents are not trading in securities contingent on certain states, yet we expect that the policymaker can easily commit to policy plans that are contingent on these same states. We may rightly wonder why individual agents are passing up opportunities for mutually beneficial trade.

The second motivation for policy will be in environments where we allow agents in the economy to trade in a full set of state-contingent securities, subject to information asymmetries. In these cases, the policymaker, who suffers from the same information asymmetries as the agents, cannot replace the markets that have been endogenously closed as a consequence of the friction. Nevertheless, there may still be a role for policy interventions in these environments (Geanakoplos and Polemarchakis, 2008, Greenwald and Stiglitz, 1986). But the desirability of these policy interventions, an example of which is described in the following sections, depends on the extent to which agents are free to trade long term contracts and rights to participate in markets in future periods (Kilenthong and Townsend, 2014). It is easiest to see the practical implications of the arguments of Greenwald and Stiglitz (1986) and Kilenthong and Townsend (2014) through an example.

AN EXAMPLE: THE MARKET FOR HEALTH INSURANCE

The market for health insurance suffers from moral hazard.[11] Health insurance reduces the cost of ill-health to the policy holder, and to that extent it reduces the incentive to live a healthy lifestyle. Health risks taken by policyholders impose costs on their insurers, and this conflict between policyholders and insurance companies can increase the costs of policy provision, decrease the extent to which health risks can be pooled across agents and also result in the excessive taking of health risks.

Greenwald and Stiglitz (1986) show us that applying a subsidy on healthy behaviour, such as a gym membership, can improve welfare by aligning the price schedule of the policyholders with the social benefits of the healthy behaviour that are shared across all insurance policyholders. This dampens the moral hazard problem, and can result in a Pareto welfare improvement even if the subsidy is funded by distortionary taxes.[12] Note

[11]The health insurance market also suffers from adverse selection, as individuals have private information about their own health before purchasing insurance. This example will focus on the implications of moral hazard.

[12]For an interesting discussion of this result and some qualifications, see Dixit (2003).

that this intervention does not require the government to have any special information about the behaviour of the policyholder that is not also known to all other agents in the economy. Borrowing the language of Dixit (2003), it is assumed that the government is benevolent, but not necessarily omniscient.

Kilenthong and Townsend (2014) show that decentralised trade could also achieve the same constrained efficient outcome, as long as agents can commit to restrict their future trade to particular markets. It may be difficult to think about what that means at first, what it means for two agents in the real world to be trading "in different markets", but it is not as bizarre as it sounds and in fact is not uncommon.

Currently, a major UK health insurance firm offers its customers a 50% discount on gym membership fees.[13] Presumably, this health insurance firm finds it worthwhile to encourage healthy lifestyles of policyholders (note that this offer would also to encourage applicants who have a high demand for gym memberships, dampening the problem of adverse selection). If private firms like this insurance company can bundle products in a way that dampens the effects of moral hazard and adverse selection, this must surely weaken the arguments of Geanakoplos and Polemarchakis (2008) and Greenwald and Stiglitz (1986), reducing the desirability of government intervention in these markets.

In this example of a bundled offer of health insurance and gym services, policyholders of this health insurance company face a different price schedule, and are effectively trading in a different market for gym memberships than non-policyholders. This means that policyholders' marginal rate of substitution between gym services and other goods does not equate to the marginal rate of transformation faced by the producers of gym services. It also means that policyholders' marginal rate of substitution between gym services and other goods will not equate to other consumers' marginal rates of substitution, as they face a different price schedule of gym services relative to other goods.

With the voucher in hand, it may be in an individual policyholder's best interest to sell the voucher to another agent who has a greater demand for gym services, equating ex post marginal rates of substitution between gym services and other goods across policyholders and non-policyholders. This ex post trade would eliminate any ex ante improvement in the moral hazard and adverse selection problems that was obtained through the voucher

[13] `http://www.pruhealth.co.uk/vitality/partners/virgin-active/` Accessed 5 July 2015.

bundling scheme. It is clear that if private agents are to trade in bundled products of financial services, markets must have mechanisms for market exclusion for certain agents who would wish to trade at current market prices. In our example, gym membership vouchers offered by health insurance companies must not be transferable. The extent to which these non-transferability exclusions are difficult to enforce may be a useful determinant of the potential gains from policy intervention. When non-transferability exclusions are straightforward to enforce in markets, then market mechanisms may work well to internalise externalities associated with information asymmetries. When non-transferability exclusions are difficult to enforce, perhaps there is greater scope to for government intervention through Pigouvian taxes along the lines of the suggestions of Geanakoplos and Polemarchakis (2008) and Greenwald and Stiglitz (1986).

Many of the policy interventions we consider in this paper will be subject to the critiques developed by Kilenthong and Townsend (2014). We have not been able to find an equivalent to our health insurance - gym membership bundle that would apply in the specific scenarios we consider, and we think that within the specific examples we consider, the re-trading restrictions that are required to implement the constrained efficient competitive equilibria described by Kilenthong and Townsend (2014) are unlikely to be enforceable. Nevertheless, we encourage the reader to bear this critique in mind when considering the policy problems we derive in later chapters.

ROADMAP

Chapter 1 shows in an endowment economy model how departures from full systemic risk sharing can dampen the incentive compatibility constraint, allowing a greater degree of idiosyncratic risk sharing. In Chapter 2 we propose a new theory of debt finance as an effective mechanism for sharing idiosyncratic risks, based on the costly state verification framework. In Chapter 3, we show how entrepreneurs' productive risk can quickly transform into low employment, as wages fall below marginal revenue products of labour. Chapter 4 considers markets for systemic risk insurance, showing that the market prices for this insurance do not necessarily internalise balance sheet externalities, resulting in excessive swings in leverage and factor market wedges of inefficiency. Chapter 5 presents a model where sticky nominal factor compensation contracts motivate herding externalities under nominal income targeting monetary regimes.

CHAPTER 1

PRIVATE INFORMATION AND RISK SHARING IN AN ENDOWMENT ECONOMY

When individuals have private information about their own luck and income, the sharing of idiosyncratic risks is hampered by moral hazard. It turns out that these frictions acting on idiosyncratic risk sharing also affect the optimal sharing of systemic risks. Restricting trade in systemic risk insurance can assist agents in identifying those who have misreported their idiosyncratic or individual specific risks. This relaxes the incentive compatibility constraint relating to moral hazard and increases the extent to which idiosyncratic risks can be shared.

CHAPTER 1

INTRODUCTION

This chapter considers the role of debt finance in an economy with aggregate risk, and serves as a further, analytical introduction to the key themes of this dissertation. More specifically, how do individual-specific information asymmetries, which limit the risk sharing of idiosyncratic risks, interact with the sharing of common shocks that are freely observable to all agents?

We explore this question within the simple 2 period endowment economy model described by Green and Oh (1991) and Kiyotaki (2010). The model can also be thought of as a two-period version of the model of Cole and Kocherlakota (2001). We'll derive the well-known result that in an economy with no aggregate risk but with private information about idiosyncratic risk, constrained efficient allocations can be implemented by competitive trade in simple risk-free debt contracts (Proposition 1.1).

Then we will introduce aggregate (or *systemic*) risk through a common shock to endowments. When aggregate risk is present, constrained efficient allocations cannot be supported by risk-free debt contracts alone (Proposition 1.2). Compared with constrained efficient allocations, the competitive allocations under simple debt contracts allocates too much risk to the borrower agents who received a low endowment in the first period. A Pareto welfare improvement can be obtained by either (1) the opening of markets for securities contingent on the common shock, or (2) through a policy intervention that replicates the missing market for securities contingent on the common shock by redistributing wealth from creditors to debtors when the common shock is small, and from debtors to creditors when the common shock is large.

We'll then show that opening systemic risk markets (or replicating their effects with fiscal transfers) does not result in constrained efficient allocations (Proposition 1.3). Constrained efficient allocations require deviations from full across-aggregate state risk sharing that are not consistent with sequential trade in aggregate state-contingent debt contracts.

Compared with the competitive equilibrium with sequential trade in aggregate state-contingent debt contracts, the constrained efficient allocations dampen the sensitivity of low reporting agents' consumption to aggregate risks. This perturbation relaxes the truth-telling constraint, as this aggregate risk insurance plan is undesirable to misreporting high endowment households who would prefer to bear a greater share of aggregate risk in return for higher expected consumption.

This Chapter concludes with discussions about the efficiency of risk sharing in decentralised markets in economies that experience booms and recessions, the difficulties of determining whether or not risk markets are closed, and if closed, whether or not this closure is inefficient, and the importance of assumptions about systemic risk sharing arrangements for policy analysis in macroeconomic models. These discussions are set in the wider context of the literature on business cycle stabilisation policy.

1.1 THE MODEL

A unit measure of ex ante identical agents live for two periods. Agents enjoy consumption with c according to utility function $U(c)$, where U is in the Decreasing Absolute Risk Aversion (DARA) class of preferences ($U', -U'' > 0, A'(c) < 0$, where $A(c) = -U''(c)/U'(c)$ is the Arrow-Pratt measure of absolute risk aversion).[1] Agents' discount second period instantaneous utility by factor β. In the first period, individual agents receive endowment y_l with probability π_l and y_h with probability π_h, where $y_l < y_h$ and $\pi_l + \pi_h = 1$. In period 2, all agents receive common endowment z, where $y_l < z < \pi_l y_l + \pi_h y_h$. That is, the second period endowment received by all agents is less than the expected first period endowment, but greater than the first period endowment received by low income agents. There exists a durable good, which converts the period 1 consumption good into the period 2 consumption good. First period savings x return Rx units of the period 2 consumption good, where $R = 1/\beta$. Without loss of generality, the consumption enjoyed by an individual agent earning y_l in period 1 is denoted c_{1l}.

Within the class of problems we consider in this Chapter, the revelation principle holds, and we can consider the constrained efficient allocations of any given problem to be attainable by a direct mechanism implemented by a benevolent social planner. Throughout this chapter, we'll consider the planner's solutions to the problems we consider, before considering whether or not these solutions can be obtained through decentralised trade using only a small set of simple contracts.

[1] The DARA class of preferences includes as a subset the class of utility functions exhibiting Constant Relative Risk Aversion (CRRA), where relative risk aversion is defined as $cA(c)$.

The social planner aims to maximise the ex ante expected discounted utility of agents. Note that there is no conflict between agents at time zero, before the idiosyncratic risk y is drawn. The planner's objective function is

$$\max_{\mathbf{c},x} \pi_l U(c_{1l}) + \pi_h U(c_{1h}) + \beta[\pi_l U(c_{2l}) + \pi_h U(c_{2h})] \tag{1.1}$$

The planner's first and second period resource constraints are as follows and we attach Lagrange multipliers λ_1, λ_2 to them respectively:

$$\pi_l y_l + \pi_h y_h \geq \pi_l c_{1l} + \pi_h c_{1h} + x \tag{1.2}$$

$$z + Rx \geq \pi_l c_{2l} + \pi_h c_{2h} \tag{1.3}$$

The first period budget constraint (1.2) states that the sum of first period consumption across agents and savings (RHS) must be less than or equal to the total first period endowment income across agents (LHS). The second period budget constraint (1.3) states that total second period consumption (RHS) cannot exceed the sum of second period income and the gross return to first period savings (LHS). It is clear that in constrained efficiency requires that both resource constraints (1.2,1.3) are binding. If either constraint were not binding, it must be the case that there is an individual agent whose consumption could be increased without violation of any of the constraints faced by the social planner.

1.2 PERFECT INFORMATION

With perfect information, the planner solves (1.1) subject to the resource constraints (1.2,1.3). The first order necessary conditions can be written as follows

$$\begin{aligned} c_{1i}: &\quad \lambda_1 = U'(c_{1l}) = U'(c_{1h}) \\ c_{2i}: &\quad \lambda_2/\beta = U'(c_{2l}) = U'(c_{2h}) \\ x: &\quad \lambda_1 = R\lambda_2 \end{aligned}$$

From the first order necessary conditions, we can see that under the planner's solution, agents enjoy full consumption insurance, with consumption equated across high and low income agents. Agents also enjoy perfect consumption smoothing, with

$$c_{1l} = c_{1h} = c_{2l} = c_{2h} = \frac{1}{1+\beta}\left[\pi_l y_l + \pi_h y_h + \beta z\right].$$

This solution characterises the first-best efficient allocations in our model. The incomes of high endowment and low endowment agents are shared, as though all agents hold equity shares in each others' incomes. The storage technology is used to smooth the consumption of all households over the two periods.

1.3 Private information

Now, consider the same model, but where the planner cannot directly observe which agents have received the high endowments, and which have received the low endowments. It is also assumed that individuals savings held in the durable good cannot be observed by the planner.[2] Agents now have the option of lying about their endowment to the planner, and saving any excess income they do not wish to consume in the first period, earning return R on all savings. The revelation principle holds in our environment, and optimal allocations can be implemented by the planner, whose problem is now subject to the following truth-telling constraint to which we attach the Lagrange multiplier μ:

$$U(c_{1h}) + \beta U(c_{2h}) = V(c_{1l} + y_h - y_l, c_{2l}). \tag{1.4}$$

The value function V represents the expected discounted utility obtainable by an agent who receives a high endowment and fraudulently declares a low endowment. In the first period, they recieve a transfer from the planner equal to $c_{1l} - y_l$, which is added to their true endowment of y_h. In the second period they receive transfer $c_{2l} - z$, which they can add to their endowment z and the gross return from any private savings in the durable good.

It is clear to see that any similar constraint to ensure truth-telling from agents receiving a low endowment would not be binding under any optimal consumption plan. The primary objective of the planner is to provide insurance to agents receiving low endowments, and

[2]The assumption that storage is hidden is important. When storage is observable, misreporting high type agents are unable to smooth consumption. This inability to smooth consumption can be manipulated by the social planner to provide some consumption insurance across high and low type agents that is not possible when storage is hidden. See Green and Oh (1991) and Kiyotaki (2010) for details.

it is always in the interest of those agents to declare their endowments truthfully.

We now solve for the value attainable by a recipient of a high endowment who misreports their endowment before returning to the planner's problem.

1.3.1 THE VALUE OF MISREPORTING

Consider a recipient of a high endowment who reports a low endowment. We denote their consumption allocations in periods 1 and 2 by $\hat{c}_1$ and $\hat{c}_2$ respectively. As storage is hidden, this agent can use the storage technology to smooth consumption across the two periods. The misreporting agent solves the following problem

$$V(c_{1l} + y_h - y_l, c_{2l}) = \max_{\hat{\mathbf{c}},x} U(\hat{c}_1) + \beta U(\hat{c}_2)$$

subject to the resource constraints

$$c_{1l} + y_h - y_l \geq \hat{c}_1 + \hat{x},$$

$$R\hat{x} + c_{2l} \geq \hat{c}_2.$$

The left hand side of the first resource constraint adds the difference between high and low endowments (the hidden part of the endowment) to the consumption allocation of a truth-telling low endowment agent. The left hand side of the second resource constraint adds the gross return of any hidden savings to the consumption allocation of a truth-telling low endowment agent. The solution to this problem is

$$\hat{c_1} = \hat{c_2} = \frac{1}{1+\beta}\left[c_{1l} + y_h - y_l + \beta c_{2l}\right]$$

$$V(c_{1l} + y_h - y_l, c_{2l}) = (1+\beta)U\left(\frac{1}{1+\beta}\left[c_{1l} + y_h - y_l + \beta c_{2l}\right]\right) \tag{1.5}$$

1.3.2 The planner's solution

The planner maximises (1.1) subject to the resource constraints (1.2,1.3) and the truth-telling constraint (1.4) with the solution (1.5). The first order necessary conditions are

$$\begin{aligned}
c_{1l}: &\quad \pi_l \lambda_1 = \pi_l U'(c_{1l}) - \mu U' \left(\frac{1}{1+\beta} \left[c_{1l} + y_h - y_l + \beta c_{2l} \right] \right) \\
c_{2l}: &\quad \pi_l \lambda_2 = \pi_l \beta U'(c_{2l}) - \mu \beta U' \left(\frac{1}{1+\beta} \left[c_{1l} + y_h - y_l + \beta c_{2l} \right] \right) \\
c_{1h}: &\quad \pi_h \lambda_1 = \pi_h U'(c_{1h}) + \mu U'(c_{1h}) \\
c_{2h}: &\quad \pi_h \lambda_2 = \pi_h \beta U'(c_{2h}) + \mu \beta U'(c_{2h}) \\
x: &\quad \lambda_1 = R\lambda_2
\end{aligned}$$

It is straightforward to verify that the solution to this problem is

$$c_{1l} = c_{2l} = \frac{1}{1+\beta}[y_l + \beta z], \qquad c_{1h} = c_{2h} = \frac{1}{1+\beta}[y_h + \beta z]. \tag{1.6}$$

1.3.3 Competitive equilibrium with non-contingent debt

The solution described by Equation 1.6 is consistent with consumption smoothing over time by individual agents ($\beta U'(c_{2l})/U'(c_{1l}) = 1/R$), which under our specific restrictions on parameter values (notably $\beta = 1/R$) means that individual consumption paths are constant across time ($c_{1j} = c_{2j}$). But the solution also restricts the total present value of consumption of each agent to be equal to the present value of their endowment paths ($c_{1l} + \beta c_{2l} = y_l + \beta z$). This indicates that there is no sharing of the idiosyncratic endowment shocks across agents. There is no redistribution of present value wealth after endowments are realised in period 1.

Proposition 1.1 shows that these constrained optimal allocations described in (1.6) can be implemented through decentralised trade in one period non-contingent loans, where this loan market opens *after* endowments have been realised in period 1. These loan markets enable agents to bring forward or delay consumption from and to the future, which offers an improvement in welfare terms relative to autarky, but little insurance against endowment risks.

Proposition 1.1 *When aggregate income is constant, the constrained efficient allocations under private information with hidden storage can be implemented with decentralised trade in non-contingent one period debt contracts.*

The proof of Proposition 1.1 is contained in Appendix 1.A.

1.4 SYSTEMIC RISK

Now we introduce systemic risk through an aggregate endowment shock in period 2. The common endowment received in period 2 can take the values $z_L < z_H$, with probabilities ν_L, ν_H respectively, where $\nu_L + \nu_H = 1$.

What we're interested in is how the aggregate risk z is shared, and whether decentralised trade in the simple debt contracts we considered in the previous section can still implement constrained efficient allocations. We start by describing the planner's problem and the planner's first order necessary conditions before considering whether these conditions can be satisfied by decentralised sequential trade in non-contingent or aggregate state-contingent debt securities. Then we return to solve the planner's problem and derive the intuition behind our result that decentralised trade in these simple contracts cannot implement constrained efficient allocations.

1.4.1 THE PLANNER'S PROBLEM

When the economy suffers from systemic risk, the planner's objective function takes expectations of individual utilities across individual agents and aggregate states (z). The planner's objective function can be written as follows,

$$\max_{\mathbf{c},x} \ \pi_l U(c_{1l}) + \pi_h U(c_{1h}) + \beta \, \mathbb{E}_z \left[\pi_1 U(c_{2l}(z)) + \pi_h U(c_{2h}(z)) \right],$$

subject to the budget constraints,

$$\pi_l y_l + \pi_h y_h \geq \pi_l c_{1l} + \pi_h c_{1h} + x, \qquad (\lambda_1)$$

$$Rx + z \geq \pi_l c_{2l}(z) + \pi_h c_{2h}(z) \qquad z \in \{z_L, z_H\}. \qquad (\lambda_2(z))$$

The first period budget constraint is the same as in the earlier cases with no systemic risk. The second period budget constraints are contingent on the realisation of the common shock (z).

The incentive compatibility constraints also change. Agents report their endowments following realisation in period 1. Therefore, when deciding whether or not to report truthfully, they must take expectations over the allocation policy rule ($c_{2l}(z)$) and the distribution of common shocks (z). The first incentive compatibility constraint is

$$U(c_{1h}) + \beta \mathbb{E} U(c_{2h}(z)) \geq V(c_{1l} + y_h - y_l, c_{2l}(z)), \qquad (\mu)$$

where $V(c_{1l} + y_h - y_l, c_{2l})$ describes the value obtainable to a an agent who receives a high endowment in the first period but declares a low income to the social planner.

The first step in solving the planner's problem is to consider the value attainable by high endowment agents who misreport their endowment.

1.4.2 The value of misreporting

A recipient of a high endowment who reports a low endowment solves the following problem

$$V(c_{1l} + y_h - y_l, c_{2l}(z)) = \max_{\hat{\mathbf{c}}, x} U(\hat{c}_1) + \beta \mathbb{E} U(\hat{c}_2(z))$$

subject to the resource constraints

$$c_{1l} + y_h - y_l \geq \hat{c}_1 + \hat{x}, \qquad (\hat{\lambda}_1)$$

$$R\hat{x} + c_{2l}(z) \geq \hat{c}_2(z) \quad \forall z. \qquad (\hat{\lambda}_2(z))$$

The agent's first order conditions are

$$\begin{aligned} \hat{c}_1: &\quad \hat{\lambda}_1 = U'(\hat{c}_1) \\ \hat{c}_2(z): &\quad \hat{\lambda}_2(z) = \beta U'(\hat{c}_2(z)) \\ x: &\quad \hat{\lambda}_1 = R \mathbb{E} \hat{\lambda}_2(z) \end{aligned}$$

The most important parts of this solution for our purposes are that the intertemporal marginal rate of substitution for misreporting agents is equated to their marginal rate of transformation—the return to the savings technology R, and that the difference in state-contingent consumption allocations of the misreporting agents are equal to those same differences for low reporting truth-telling agents. That is,

$$\hat{c}_2(z_H) - \hat{c}_2(z_L) = c_{2l}(z_H) - c_{2h}(z_L).$$

What this shows is that any high earning agent who misreports their wealth must enter into the same systemic risk insurance plan as low endowment agents. We can summarise these two results as follows:

$$V_1 = R\mathbb{E}_z V_2(z)/P(z), \qquad \text{and} \tag{1.7}$$

$$\frac{V_2(z_L)/P(z_L)}{V_2(z_H)/P(z_H)} < \frac{U(c_{2l}(z_L))}{U(c_{2l}(z_H))}.^3 \tag{1.8}$$

As we'll see, this creates an opportunity for the planner to implement some insurance against *idiosyncratic* risks. Specifically, the planner can use the fact that systemic risk insurance plans that are desirable to low wealth individuals are likely to be undesirable to high endowment individuals. We can use this to elicit truth-telling from high endowment agents even when this results in a small transfer of wealth.

1.4.3 THE PLANNER'S FIRST ORDER NECESSARY CONDITIONS

The planner's first order necessary conditions are described by the following:

$$\begin{aligned}
c_{1l}: &\quad 0 = \pi_l U'(c_{1l}) - \pi_l \lambda_1 - \mu V_1(c_{1l} + y_h - y_l, c_{2l}(z)) \\
c_{1h}: &\quad 0 = \pi_h U'(c_{1h}) - \pi_h \lambda_1 + \mu U'(c_{1h}) \\
x: &\quad 0 = \lambda_1 - R\mathbb{E}\lambda_2(z) \\
c_{2l}(z): &\quad 0 = P(z)\pi_l \beta U'(c_{2l}(z)) - P(z)\pi_1 \lambda_2(z) - \mu V_2(c_{1l} + y_h - y_l, c_{2l}(z)) \\
c_{2h}(z): &\quad 0 = P(z)\pi_h \beta U'(c_{2h}(z)) - P(z)\pi_h \lambda_2(z) + \mu\beta P(z) U'(c_{2h}(z))
\end{aligned}$$

[3] This makes use of the assumption of DARA preferences.

1.4.4 THE PLANNER'S SOLUTION: KEY FEATURES

Eliminating μ from the planner's first order conditions with respect to the consumption allocations of high endowment individuals yields

$$\frac{U'(c_{1h})}{\lambda_1} = \frac{\beta U'(c_{2h}(z))}{\lambda_2(z)}$$

Which ensures first that high endowment agents receive systemic risk consumption insurance commensurate with the planner's marginal value of second period wealth,

$$\frac{U'(c_{2h}(z_L))}{U'(c_{2h}(z_H))} = \frac{\lambda_2(z_L)}{\lambda_2(z_H)}$$

and second, when combined with the first order condition for aggregate savings x, that the high endowment agents' intertemporal marginal rates of substitution equate to the intertemporal marginal rate of transformation specified by the storage technology, R,

$$\frac{\mathbb{E}_z \beta U'(c_{2h}(z))}{U'(c_{1h})} = \frac{1}{R}.$$

Turning to the low endowment households, eliminating μ from the first order conditions for their consumption allocations yields

$$\frac{U'(c_{1l}) - \lambda_1}{V_1(c_{1l} + y_h - y_l, c_{2l}(z))} = \frac{U'(c_{2l}(z)) - \lambda_2(z)}{V_2(c_{1l} + y_h - y_l, c_{2l}(z))/P(z)}.$$

Substituting the solutions 1.7 and 1.7 yields

$$\frac{U'(c_{2l}(z_L))}{U'(c_{2l}(z_H))} < \frac{\lambda_2(z_L)}{\lambda_2(z_H)}, \qquad \text{and}$$

$$\frac{\mathbb{E}_z \beta U'(c_{2l}(z))}{U'(c_{1l})} = \frac{1}{R}.$$

The first equation shows that low endowment agents are protected from systemic risk (z), to such an extent that their second period marginal utility is less sensitive to systemic risk than that of high endowment agents:

$$\frac{U'(c_{2l}(z_L))}{U'(c_{2l}(z_H))} < \frac{U'(c_{2h}(z_L))}{U'(c_{2h}(z_H))}. \tag{1.9}$$

The second equation shows that low endowment agents' intertemporal marginal rates of substitution are equated to the social marginal rate of transformation (the return on the storage technology) under constrained efficient allocations.

We'll show in the following sections that this is not consistent with competitive equilibria under trade in non-contingent *or* aggregate state-contingent loans, before returning to the intuition behind this result and the lessons that we can learn from it.

1.4.5 COMPETITIVE EQUILIBRIUM WITH NON-CONTINGENT DEBT

In period 1, upon realisaton of endowments, individual agent in receipt of endowment y_i can buy (+) or offer (-) unlimited quantities of non-contingent debt b_i, each unit of which returns one unit of the consumption good in the second period. The agent's problem can be written as follows:

$$\max_{\mathbf{c}_i, x_i, b_i} U(c_{1i}) + \beta U(c_{2i})$$

subject to the resource constraints

$$y_i \geq c_{1i} + x_i + Qb_i, \quad (\lambda_{1i})$$

$$Rx_i + b_i + z \geq c_{2i}(z). \quad (\lambda_{2i}(z))$$

In symmetric equilibrium, the total supply of one period bonds must be equal to zero:

$$\pi_l b_l + \pi_h b_h = 0.$$

The first order necessary conditions are

$$\begin{aligned} c_{1i}: &\quad \lambda_{1i} = U'(c_{1i}) \\ c_{2i}: &\quad \lambda_{2i} = \beta U'(c_{2i}(z)) \\ x: &\quad \lambda_{1i} = R\mathbb{E}\lambda_{2i}(z) \\ b: &\quad Q\lambda_{1i} = \mathbb{E}\lambda_{2i}(z) \end{aligned}$$

Proposition 1.2 *With systemic risk present, the competitive equilibrium with non-contingent debt only is not constrained efficient.*

The proof of Proposition 1.2 is contained in Appendix 1.B.

The problem with the non-contingent debt contracts is that while they do provide some intertemporal insurance in the form of consumption smoothing, they do not provide sufficient insurance against systemic risks. Creditors (receivers of high endowments) and debtors (low endowments) face identical absolute consumption risks in the second period with respect to the systemic risk. But debtors have higher expected marginal utility in the second period than debtors, and any absolute decrease in consumption results in a greater increase in marginal utility than that suffered by a creditor following the identical absolute change in consumption (under DARA preferences). That is, in the competitive equilibrium with non-contingent debt,

$$\frac{U'(c_{2l}(z_L))}{U'(c_{2l}(z_H))} > \frac{U'(c_{2h}(z_L))}{U'(c_{2h}(z_H))},$$

which contradicts 1.9.

It is this difference in how each group's marginal utilities respond to the systemic risk that indicates that a market for systemic risk insurance, or an allocation mechanism replicating the missing systemic risk insurance could yield a Pareto welfare gain.

Now, we introduce a market for systemic risk insurance into our competitive environment.

1.4.6 The Competitive Equilibrium with State-Contingent Debt Contracts

We've shown that simple non-contingent debt contracts cannot implement constrained efficient allocations when our endowment economy suffers from aggregate or systemic risk. Ex ante, all individual agents are identical, but after the realisation of idiosyncratic risk, some agents have greater wealth and consumption than others. These low wealth individuals are less able and willing to bear systemic risk than the higher wealth individuals. Since the outcome of the systemic risk is common knowledge, the planner is able to construct a superior mechanism that does provide low wealth agents with some insurance against the systemic risk shock, resulting in a Pareto welfare improvement.

In this section, we consider whether decentralised trade could achieve constrained efficient allocations, if individual agents were able to trade a richer set of securities that allowed for payoffs that respond to the outcome of the systemic risk.

In period 1, upon realisaton of endowments, individual agent in receipt of endowment y_i can buy (+) or offer (-) unlimited quantities of state-contingent debt $b_i(z')$, each unit of which returns one unit of the consumption good in the second period if and only if the realisation of z is $z = z'$. Each security $b(z')$ trades at price $Q(z')$ in period 1.

The agent's problem can be written as follows:

$$\max_{\mathbf{c}_i, x_i, b_i} U(c_{1i}) + \beta U(c_{2i})$$

subject to the resource constraints

$$y_i \geq c_{1i} + x_i + Q(z) b_i(z), \quad (\lambda_{1i})$$

$$R x_i + b_i(z) + z \geq c_{2i}(z). \quad (\lambda_{2i}(z))$$

In symmetric equilibrium, the total supply of bonds contingent on state z must be equal to zero:

$$\pi_l b_l(z) + \pi_h b_h(z) = 0 \qquad \forall z.$$

The first order necessary conditions are

$$\begin{aligned} c_{1i}: &\quad \lambda_{1i} = U'(c_{1i}) \\ c_{2i}: &\quad \lambda_{2i} = \beta U'(c_{2i}(z)) \\ x: &\quad \lambda_{1i} = R\mathbb{E}\lambda_{2i}(z) \\ b(z): &\quad Q(z)\lambda_{1i} = \lambda_{2i}(z) \qquad \forall z \end{aligned}$$

We can see straight away that the agents in our economy do in fact utilise the state-contingent contracts. There is full consumption risk sharing with respect to the systemic risk, z:

$$\frac{U'(c_{2h}(z))}{U'(c_{1h})} = \frac{U'(c_{2l}(z))}{U'(c_{1l})} = Q(z) \qquad \forall z. \tag{1.10}$$

Proposition 1.3 *With systemic risk present, the competitive equilibrium with state-contingent debt is not constrained efficient.*

Proof. The proof of Proposition 1.3 follows directly from consideration of equation 1.10, which directly contradicts equation 1.9. ■

When systemic risk markets are open, and loan contracts can be written to be contingent on the aggregate state, there is full consumption insurance. High income and low income agents experience the same variation in marginal utilities across aggregate states.

But what equation 1.9 shows us is that we can do better than this full systemic risk insurance allocation by even further protecting low endowment agents from fluctuations in systemic risk. By making their income more 'sticky' relative to total income.

The reason why this increased protection of low endowment individuals to systemic risks is useful is that in decentralised trade, high wealth individuals are natural sellers of insurance against systemic risks, increasing their exposure to systemic risks in exchange for a small insurance premium paid for by low wealth individuals. Misreporting agents, like truth-telling high income agents, are also tolerant of systemic risks and would wish to sell insurance against them. But misreporting agents are restricted by the planner in their exposure to systemic risk, such that their absolute consumption allocation varies according to systemic risk to the same extent as that of truth-telling low income agents. This restriction discourages high income agents from misreporting their incomes, relaxing their incentive compatibility constraint and allowing the planner to achieve some sharing of idiosyncratic risks. This increased insurance against idiosyncratic risks provide an ex ante welfare gain to all agents.

The planner in our model identifies misreporting agents by their demand for a certain good. In our case, misreporting agents have a higher demand for exposure to systemic risk than truth-telling low type agents. If we were to relax the assumption that storage is hidden then we can also identify misreporting agents by their demand for savings, which is greater than truth-telling low reporting agents (Green and Oh, 1991). This identification of misreporting agents through their differentiated demands for loans and systemic risk is not unrealistic and should not be restricted to the blackboard. It is a feature of real world contract and law enforcement. For a dramatic example, readers might recall the 1990 film *Goodfellas*. In one scene, short-tempered Tommy DeVito (played by Joe Pesci) scolds his mob colleagues for purchasing fur coats and expensive cars immediately after a successful heist. DeVito's concern is that these conspicuous purchases might reveal their recent windfall to polices, identifying them as suspects.

It is important to stress that the insights developed in this section, and summarised by Proposition 1.3, can only be derived when we approach the problem of systemic risk sharing starting with a clear microfoundation for the use of non-contingent debt as a mechanism for managing idiosyncratic risk. That is, starting from the assumption that agents have private information about their individual specific idiosyncratic risks. If we had instead taken non-contingent debt contracts as the starting point, restricting contracts directly rather than as a response to the incentive compatibility constraint, we would have missed the interaction between systemic risk sharing arrangements and the incentive compatibility constraint. We would have missed the insight that a small departure from full systemic risk sharing can help agents share idiosyncratic risks.

1.5 CAN MARKETS IMPLEMENT THE CONSTRAINED EFFICIENT ALLOCATIONS?

We've seen that with simple one-period debt instruments, and with slightly more sophisticated one-period state-contingent debt contracts, competitive markets cannot implement the constrained efficient allocations described in the previous section. This does not mean that decentralised trade cannot implement constrained efficient allocations.

The mechanism implemented by the social planner at time zero provides an ex ante welfare improvement to all agents, who would wish to enter into the mechanism or a sophisticated set of contracts at time zero replicating the state-contingent transfers implemented by the social planner in periods 1 and 2. Kilenthong and Townsend (2014) show that decentralised trade in an appropriate set of multi-period contracts can implement constrained efficient allocations in a generalised incomplete markets setting in which our model can be thought of as a specific example.

These contracts would need to satisfy the condition (1.9). Equation 1.9 shows that under the constrained efficient allocations, the across-state marginal rates of substitution are not equated across agents. This means that under any constrained efficient mechanism, individual agents would want to re-open systemic risk insurance markets prior to the realisation of the systemic risk in period 2. Any re-opening of systemic risk insurance markets in period 1 would result in the equation of across-state marginal rates of substitution, and therefore would not be consistent with constrained efficiency.

The question of whether decentralised trade can implement constrained efficient allocations depends largely on whether agents can commit *not* to trade in future periods, or whether trading institutions can be arranged such that any two agents who would wish to trade in period 1 would be restricted from doing so.

This may be possible in some market environments. In the introduction to this dissertation we described a health insurance company who offered a bundled package combining health insurance with discounted gym membership in an effort to combat adverse selection, by encouraging applicants who place a high value on gym memberships, and moral hazard, by encouraging the use of gym facilities by health insurance customers. In our environment, the time zero bundling of commitments to loan and savings instruments as well as mechanisms for sharing the period 2 income at interest rates and prices that deviate from ex post market clearing rates and prices could plausibly enforce the deviations from period 1 intertemporal and across-state risk sharing associated with the constrained efficient allocations.

Green and Oh (1991) study a related example with no systemic risk but with observable storage where the constrained efficient mechanism requires similar restrictions from mutually beneficial trade in one-period loan contracts in period 1. They argue that elements of the solution can resemble a credit rationing behaviour, and that this behaviour can explain some of the features of rural US financial markets with monopolistic banks. An important feature of these markets is that the monopoly status of the bank means that there is no opportunity for mutually-benificial side trades in loan contracts.

1.6 Which markets are open?

In Green and Oh (1991), which studies a similar environment but with observable storage, efficient allocations require that loan markets are partially closed when future income is expected to be low.[4] Here, in a very similar framework, we've shown that efficient allocations require systemic risk markets to be partially closed when future income is highly risky or uncertain. In sum, these simple endowment economy models are suggesting that during recessions and periods of uncertainty, private information relating to individual specific risks can impair the functioning of non-contingent loan markets and systemic

[4] By 'partially closed' I mean that there exist individual pairs of agents who could realise mutual gains from trade in these markets at period 1.

risk insurance markets that would not normally be considered to be hampered by these types of individual specific information asymmetries. Frictions which restrict the sharing of individual specific risks also affect the sharing of aggregate risks.

It would be a mistake for the observer to conclude that the absence of apparently mutually beneficial trade in loan and insurance markets is conclusive evidence of constrained inefficiency that warrants policy intervention. Green and Oh (1991) provocatively refer to their result as an “efficient credit crunch”. Our result described here can be thought of as an efficient closing of systemic risk insurance markets, which perhaps doesn’t have the same ring as “efficient credit crunch”, but still results in a departure from what might be considered a fair allocation of business cycle risk. Some households experience greater swings in marginal utility over the business cycle than others.

On the other hand, it is difficult to argue that the financial market disruption and consequent shut down of credit markets in the wake of the 2008 financial crisis was the constrained efficient response of an ex ante optimal mechanism of appropriately bundled state-contingent securities. It is even more difficult to argue that widespread unemployment pursuant to the financial crisis was a necessary consequence of constrained efficient market arrangements, or that the distributional sharing of business cycle risks observed in 2008 was efficiently unfair. Credit crunches and closed insurance markets are surely not always and everywhere constrained efficient.

The picture that is emerging is that fiscal and monetary policymakers face a challenging problem when determining which markets are open and closed, and whether perceived arrangements are consistent with efficient mechanisms and contracts or whether there is a role for policy interventions in implementing Pareto superior allocations.

1.6.1 The planner’s problem in related literature

A number of recent papers in macroeconomics have considered economies where agents are restricted to trading in non-contingent one period debt contracts. Christiano and Ikeda (2011) consider the role of a series of government interventions in a sample of four popular financial macroeconomics models. In their analysis, which assumes that loan contracts are not contingent to systemic risks, transfers from creditor households to entrepreneurs during downturns tend to provide useful stabilisation gains.

In Koenig (2011) and Sheedy (2014), agents are restricted to nominal debt contracts, which are not contingent to aggregate risks. Monetary policy frameworks which aim to stabilise nominal GDP tend to redistribute real wealth from creditors to debtors during periods when real output is low. These real transfers, which do not occur under inflation or price level targeting frameworks, replicate the missing transfers that would occur if markets for securities contingent on aggregate states were traded.

In the above papers, the way that policy works is by replicating payments that agents would wish to implement themselves, and could implement with sequential competitive trade in one period systemic risk insurance markets. Carlstrom, Fuerst, and Paustian (2014) show in the popular financial accelerator model of Bernanke, Gertler, and Gilchrist (1999) that sequential trade in systemic risk insurance markets could bring the competitive equilibrium of the model very close to the first-best efficient alternative. Krishnamurthy (2003) and Nikolov (2014) show a similar result in the collateral amplification model of Kiyotaki and Moore (1997).

What these models share is a common feature that when systemic risk insurance markets are open, the invisible hand allocates resources in such a way that consumption insurance (equating marginal utilities of consumption) is compromised in favour of stabilisation of leverage and financial market stress. More specifically, the market prices of systemic risk insurance fully internalise the balance sheet externalities that drive departures from efficient outcomes in these models. In this way, the outcome of insurance trade is to dampen financial crises.

What we show in Chapter 4 is that these two quite different models constitute special cases of the general financial macroeconomic framework, in which there is no guarantee that the invisible hand will forego consumption insurance against stabilisation of the financial sector. Market prices of systemic risk insurance do not fully internalise balance sheet externalities. We also show that the payments and transfers associated with complete systemic risk markets can resemble simple realistic institutions such as deposit accounts. The implication is that financial stress is not necessarily conclusive evidence that systemic risk insurance markets are closed, and that attempts to open these markets may not bring the desired improvements in financial stability.

1.A PROOF OF PROPOSITION 1.1

Proof. In period 1, upon realisaton of endowments, individual agent in receipt of endowment y_i can buy (+) or offer (-) unlimited quantities of non-contingent debt b_i, each unit of which returns one unit of the consumption good in the second period. The agent's problem can be written as follows:

$$\max_{\mathbf{c}_i, x_i, b_i} U(c_{1i}) + \beta U(c_{2i})$$

subject to the resource constraints

$$y_i \geq c_{1i} + x_i + Qb_i, \quad (\lambda_{1i})$$

$$Rx_i + b_i + z \geq c_{2i}. \quad (\lambda_{2i})$$

In symmetric equilibrium, the total supply of one period bonds must be equal to zero:

$$\pi_l b_l + \pi_h b_h = 0.$$

The first order necessary conditions are

$$\begin{aligned} c_{1i}: &\quad \lambda_{1i} = U'(c_{1i}) \\ c_{2i}: &\quad \lambda_{2i} = \beta U'(c_{2i}) \\ x: &\quad \lambda_1 = R\lambda_2 \\ b: &\quad Q\lambda_1 = \lambda_2 \end{aligned}$$

The agents' first order necessary conditions can be rearranged to show that agents' optimal consumption profile exhibits constant consumption over periods 1 and 2:

$$c_{1h} = c_{2h} \quad \text{and} \quad c_{1l} = c_{2l}.$$

The value of c_{1i} that satisfies $c_{1i} = c_{2i}$ and the individual budget constraints for the agent receiving endowment i is

$$c_{1i} = \frac{1}{1+\beta}[y_i + \beta z],$$

which is identical to the solution 1.6. ■

1.B Proof of Proposition 1.2

Proof.

Consider a mechanism that replicates the consumption allocations that are identical to those enjoyed by agents under the competitive equilibrium with non-contingent debt. Non-contingent debt does not allow for transfers of wealth contingent on the common shock z. Given this, and the fact that the first order conditions of the individual agents under competitive trade with non-contingent debt result in the gross interest rate being equated to the gross return to hidden savings $(1/Q = R)$, solving the value function of misreporting agents yields the following:

$$V_1(c_{1l} + y_h - y_l, c_{2l}(z)) = U'(c_{1h})$$

$$V_2(c_{1l} + y_h - y_l, c_{2l}(z)) = \beta P(z) U'(c_{2h}(z)).$$

Substituting these solutions into the planner's first order necessary conditions, we obtain the following:

$$\mu = \pi_l \left[\frac{U'(c_{1l}) - \lambda_1}{U'(c_{1h})} \right] = \pi_h \left[\frac{\lambda_1 - U'(c_{1h})}{U'(c_{1h})} \right]$$

$$\mu = \pi_l \left[\frac{\beta U'(c_{2l}(z)) - \lambda_2(z)}{\beta U'(c_{2h}(z))} \right] = \pi_h \left[\frac{\lambda_2(z) - \beta U'(c_{2h}(z))}{\beta U'(c_{2h}(z))} \right] \qquad \forall z.$$

These conditions can be rearranged to obtain

$$\frac{\lambda_1}{U'(c_{1h})} - \frac{\lambda_2(z)}{\beta U'(c_{2h}(z))} = 0$$

$$\frac{U'(c_{1l})}{U'(c_{1h})} = \frac{U'(c_{2l}(z_H))}{U'(c_{2h}(z_H))} = \frac{U'(c_{2l}(z_L))}{U'(c_{2h}(z_L))}. \tag{1.11}$$

As non-contingent debt does not allow transfers between agents contingent on z, the individual agents' budget constraints specify that the absolute difference in consumption across the common shock (z) is equated across individual agents:

$$c_{2l}(z_H) - c_{2l}(z_L) = c_{2h}(z_H) - c_{2h}(z_L).$$

But, as non-contingent debt also does not allow transfers of wealth across agents, the consumption smoothing by individual agents specified by their individual first order necessary conditions means that

$$c_{2l}(z) < c_{2h}(z) \quad \forall z.$$

Under DARA preferences, it can be shown that $U'''(c) > 0$. It follows that under the competitive equilibrium under non-contingent debt contracts,

$$\frac{U'(c_{2l}(z_L))}{U'(c_{2l}(z_H))} > \frac{U'(c_{2h}(z_L))}{U'(c_{2h}(z_H))}.$$

This is a contradiction with 1.11, completing the proof. ■

CHAPTER 2

DISPUTES AND THE OPTIMALITY OF STANDARD DEBT

This chapter is co-authored with Charles Nolan.[1]

We show how the prospect of disputes over firms' revenue reports promotes debt financing over equity. These findings are presented within a costly state verification model with a risk averse entrepreneur. The prospect of disputes encourages incentive regimes which limit penalties and avoid stochastic monitoring, even when the lender can commit to stochastic enforcement strategies. Consequently, optimal contracts shift away from equity and toward standard debt. For a useful special case of the model, closed form solutions are presented for leverage and consumption allocations under efficient debt contracts.

[1]Professor of Economics, University of Glasgow. `Charles.Nolan@glasgow.ac.uk`. An earlier version of this chapter has been circulated as a discussion paper under the title *Disputes, Debt and Equity*, University of Glasgow Economics Discussion Paper 2014-21. The authors would like to thank Andrew Clausen, John Hardman Moore, Joel Sobel, Jonathan Thomas and Yiannis Vailakis in addition to conference and seminar participants at the University of Glasgow Macrotheory Seminar Series, the Reserve Bank of New Zealand, the 2014 Money Macro Finance Annual Conference, the 2015 Western Economics Association Annual Conference and the 2015 Royal Economic Society Annual Conference for helpful comments and discussions. We would also like to thank the Scottish Institute for Research in Economics for generous financial support. All errors are our own.

INTRODUCTION AND OVERVIEW

What form should an optimal contract take to handle the prospect of wrongful penalties? Our answer is that standard debt contracts are often optimal. We propose a theory of debt and limited liability based on inaccurate auditing in a costly state verification framework.

As we detail presently, costly state verification environments typically imply that equity contracts are optimal. However, that conclusion is shown to rest crucially on the efficacy of the audit technology. We introduce wrongful penalties through an imperfect audit technology.[2] With *perfect* auditing, optimal contracts can economize on audit costs ex post by committing to sufficiently severe penalties ex ante. Consequently, deterministic auditing is not optimal. Moreover, enjoying access to a perfect audit technology, agents may write contracts which pass on small fluctuations in revenue to security holders. However, a key insight under *imperfect* auditing is that the prospect of disputes and wrongful penalties restricts acceptable contracts to using smaller penalties than would otherwise be the case.[3] Smaller penalties come at the cost of encouraging fraudulent reports for any given repayment schedule. So, whilst increasing the frequency of audits can help with risk sharing, it also implies levying more wrongful penalties.

The key insight behind the result of this paper is the extent to which the marginal benefit of an increase in the probability of auditing depends on the precision of the audit technology.

When audits are perfect, all audited misreporting agents are exposed as misreporting their individual revenue. If penalties are sufficiently large, agents can be deterred from misreporting their even if the probability that they will be detected is low. When the probability of audit is sufficiently high, any marginal increase in the probability of audit yields no further incentive benefits, but does suffer resource costs. This means that deterministic audit strategies will not be used. On the other hand, starting from a contract with no risk sharing and no audits, the marginal benefits from a small increase in the probability of audit can be large, allowing a high degree of risk sharing with little resources spent on

[2]Our results carry over to other situations where the lender or bankruptcy court might erroneously dispute the borrower's revenue report.

[3]Our model is static, and our penalties are just units of the consumption good, but this reasoning carries over to alternative settings where alternative enforcement schemes such as non-pecuniary penalties or exclusion may be applied. When disputes are possible, even honest borrowers prefer the penalties for dishonesty to be smaller, all else equal.

audits, as even the small probability of audit can deter misreporting when combined with large penalties.

When audits are imperfect, the story changes. Even if the probability of audit is high, further increases in the probability of audit can always provide some marginal benefit by allowing contracts to achieve the same insurance profile with smaller penalties. As penalties are sometimes wrongly applied to truth-telling agents, any reduction in penalties provides an ex ante expected welfare gain.[4] On the other hand, starting from a contract with no auditing and no risk sharing, a small increase in the probability of audit yields a much smaller insurance benefit, for two reasons. First, penalties are limited by the welfare costs of wrongful penalties. Second, imperfect audits don't always catch misreporting agents, even if these agents are audited. Both of these effects mean that the welfare gains from a small increase in audits, starting from a low audit probability, are much smaller when the audit technology is imperfect than when the audit technology is perfect.

In sum, The marginal insurance benefits of an increase in the probability of audit are initially high but rapidly decreasing to zero in the probability of audit when the audit technology is perfect. When the audit technology is imperfect, the marginal insurance benefits of an increase in the probability of audit are initially relatively low, but unlike in the perfect audits case they do not rapidly decrease as the probability of audit increases, and they remain positive even as the probability of audit approaches one.

Faced by these very different marginal insurance benefit profiles offered by perfect and imperfect technologies, we find that the optimal contracts under perfect audits can look a lot like equity finance contracts, with low probability stochastic audits applied across a wide range of possible reports. A large amount of individual risk is passed on to outside investors and a small share of individual risk is retained by the entrepreneur.

Under imperfect audits, contracts can resemble standard debt. Under these standard debt contracts, moderate or high reports are not audited, and any marginal income risk between moderate and high states is absorbed by the entrepreneur, who simply pays the principal plus interest. The cost of wrongful penalties and the small benefits obtainable by audits that don't capture misreporting agents with certainty mean that the optimal audit probability across these states is zero. But in low states, where the marginal utility of

[4] Note that this feature of our model requires imperfect audits, and cannot be obtained with an exogenous restriction on penalties alone.

the entrepreneur is high, the insurance benefits from auditing start to outweigh the cost, even when these audits don't capture every misreporting agent and are accompanied by the occasional wrongful penalty. Further, the marginal insurance benefit of audits remains high even as the probability of audit increases.

The upshot is that we find that the optimal external finance contract often combines deterministic audits following low revenue reports with no audits for revenue reports above a certain cutoff. Following most reports, the borrower's repayment is independent of marginal differences in revenue: the borrower simply repays the principal plus interest. Moreover, when errors are rare, the optimal repayment following an overturned income report is just equal to the contracted coupon plus principal of the loan. Such features resemble standard debt contracts.[5,6] The key results are contained largely in Theorem 2.2 and Proposition 2.3 below.

In the previous chapter, we considered a model with private information, and showed that simple non-contingent debt contracts could be an efficient contract when there were no aggregate risks present. In this chapter we are relaxing the private information assumption. Rather than assuming that private information can never be attained by outside parties as we did in the previous chapter, here we are exploring what happens when private information is obtainable at some cost. What we show is that it is the quality of audit technologies, rather than the cost of audit technologies, which is the key to explaining the standard form of debt contracts that we see in practise, which typically exhibhit some bankruptcy mechanism providing some insurance to debtors following the worst possible outcomes. Within the simple private information environment we considered in the previous chapter, there was no such bankruptcy mechanism even following low reports.

[5]Earlier contributions to CSV problems with audit errors have focused on insurance problems in the context of a risky endowment. Haubrich (1995) shows that weakly informative audits are rarely used in efficient contracts. Alary and Gollier (2004) study an example with no commitment to audits, showing that the occurence of strategic default is dependent on the preferences of the agent. Imperfect signals are also commonly employed in the law enforcement literature. See Polinsky and Shavell (2007) for an excellent summary.

[6]A different literature assumes that project outcomes are observable, yet entrepreneurial actions are partially observable. Efficient contracts must encourage entrepreneurs to exert privately costly effort. In these models, the concepts of debt and equity finance are related solely to the optimal sensitivity of repayments to project outcomes. A recent example which rationalises a combination of debt and equity in this setting with partially observable actions and limited enforceability is Ellingsen and Kristiansen (2011).

DETERMINISTIC INCENTIVE REGIMES AND LEVERAGE

Introducing imperfect audits encourages both deterministic audit regimes, when risk sharing is considered to be of high value, and also the complete removal of audits, when risk sharing is considered to be of low value. The interaction between leverage and costly, imperfect auditing underpins the finding that deterministic incentive schemes are generally optimal. Note that leverage and audit probability are similar in that higher leverage increases expected consumption and the spread of consumption outturns; so too does a decrease in audit probabilities. So, for low levels of borrowing and hence low levels of risk, audits are less desirable. However, if borrowing is very high there will be a large impact on consumption if a low return is mistaken for a high return, what we call a Type-I error. That implies that there is an *endogenous borrowing limit* and that the audit probability goes to zero as borrowing approaches that limit. For intermediate levels of borrowing equilibrium auditing is typically deterministic; that is, of probability one.

That non-monotonic relationship between audit probabilities and leverage is perhaps surprising. However, more surprising is what we label a *bang-bang* result: Efficient contracts can jump from being non-contingent to standard debt contracts with deterministic auditing in low states, in response to marginal increases in project risk. Moreover, there is a discontinuous decrease in optimal leverage, an increase in default, an increase in expected monitoring costs and a drop in average consumption. We are able to characterize analytically the trade-offs that occur at the point when optimal contracts change in that way. At that point, there are two contracts which deliver the same level of utility; one a high-leverage/never-audit contract, the other a low-leverage/standard debt contract.

Audit costs also play an important role in determining optimal leverage. When audit costs are low, optimal leverage is such as to permit large gains from insurance or auditing. This is what Gale and Hellwig (1985) find in their seminal paper. In our case, 'extreme' incentive regimes tend to be optimal and auditing strategies are, again, deterministic.[7]

[7] Specifically, Gale and Hellwig (1985) also study the effects of audit costs and risk aversion on leverage in a costly state verification model with perfect and deterministic audits. Our analysis permits stochastic audit regimes, and finds alternative interactions between leverage and the contracting environment: leverage has a dramatic impact on the nature of the efficient contract in our model, and it is the joint determination of leverage and incentive regime which encourages debt contracts in our framework.

2.1 LITERATURE

Equity finance typically allows issuers to reduce repayments or dividends in bad times whilst reductions in the value of assets are shared between borrowers and lenders. Debt finance is more rigid. Debts are only reduced or discharged in bankruptcy, which follows large falls in income or asset values. So, surely it would be better if there was less debt and more equity?

Townsend (1979) was first to propose an explanation for the prevalence of debt contracts. He shows that when a risk averse borrower's income is costly to verify a standard debt contract is superior either to a strict debt contract, where repayments are constant across states, or a standard equity contract, where repayments are proportional to the borrower's income. The difficulty with the equity contract is that to ensure the borrower does not misreport income the investor needs to undertake a costly audit regardless of the report. A superior contract prescribes audits and risk sharing only following sufficiently low reports, when the borrower's marginal utility and sensitivity to risk are highest. If the borrower's income is sufficiently high, they make a fixed repayment and absorb any remaining income risk at the margin. Such a contract is the standard debt contract that is widespread in personal and corporate loan markets.

Townsend's analysis constrained agents to deterministic auditing regimes. However, he suggested a better contract might employ a stochastic auditing schedule (Townsend, 1979, Section 4). Perhaps following a very low report an audit would be highly likely, and following a high report less so. Using stochastic auditing schemes would allow more risk sharing across states with fewer resources spent on audits across a portfolio of loans. Border and Sobel (1987) and Mookherjee and Png (1989) confirm Townsend's conjecture. In fact, they show that deterministic audit strategies are *never* constrained efficient: Audit strategies should be stochastic, and the probability of audit should be positive even following relatively high revenue reports. Such a contract looks more like equity in the sense that income need not be low before the contract specifies risk sharing.

That risk sharing comes at a cost. A cost that is not captured in the benchmark model. In order to ensure truth-telling when the probability of audit is low, audits that contradict the borrower's report can result in penalties far larger than the amount borrowed. If that audit technology were to contradict a *truthful* report, then the prospect of sizeable, wrongful penalties might render such contracts unacceptable to the borrower. Indeed, even if the entrepreneur were merely to *fear* that audits may not be perfect, or that their truthful report may be disputed by the lender or bankruptcy court, they would likely baulk at a contract that leaves open the prospect of large penalties following disputed reports. In short, equity-like contracts provide more insurance across states, but may exacerbate already bad situations for a borrower. Hence the motivation of this paper.

2.1.1 FURTHER FEATURES OF OPTIMAL CONTRACTS

As in Townsend's (1979) original analysis, our model motivates an endogenous form of limited liability.[8] Following default, entrepreneurs who successfully restructure their debts enjoy strictly positive consumption—they are not personally liable for the repayment of their firm's debts. That is at least the case if there is no dispute over the entrepreneurs' report. Following a disputed report, the entrepreneur is liable for the full debt repayment, plus potentially an additional fee if the audit technology suffers Type-II errors (that is, if the audit technology does not always identify misreporting high productivity entrepreneurs). Making entrepreneurs fully liable for debt repayment following disputes resembles the practise of *piercing the corporate veil*, which is relatively common in the United States in cases of corporate fraud.[9] In our model, the only form of fraud is misreporting income.

One feature of the equilibrium we consider in this paper is that the only entrepreneurs who are exposed as committing fraud are truth-telling agents. We would hope that in practise, most of the convictions for corporate fraud are not errors! Within our framework, purposeful fraud could certainly occur in response to poorly drafted contracts, or in departures from fully rational strategies, and in the real world such actions are likely to occur from time to time, indeed are likely to be more common than the Type-I audit errors that drive our results. Yet even with a very low probability of occuring, or merely the perception on the part of entrepreneurs that a Type-I error could occur, the main results of

[8] In applications of Townsend's framework with risk neutrality, including Gale and Hellwig (1985) and Bernanke, Gertler, and Gilchrist (1999), liability is only limited by the inability to pay, the lender simply takes everything upon default..

[9] An excellent discussion of the law and economics of piercing the corporate veil is given by Macey and Mitts (2014). In our model, the event of shifting liability to entrepreneurs following disputes is consistent with the authors' third justification of veil piercing, which is to prevent firm insiders from transferring corporate assets to themselves during bankruptcy reorganisation. In effect, this is exactly what the entrepreneur is accused of doing in the event of a Type-I error.

this paper will hold—debt-like contracts with limited liability will be desirable.

In our model, when the audit technology is relatively accurate, the costs associated with wrongful penalties and Type-I errors decline. Optimal contracts involve a high degree of risk sharing, with larger penalties with lower audit probabilities. Essentially, these contracts resemble equity even if the resource costs of audit are significant. In this sense, our framework nests both equity-like contracts and debt-like contracts as optimal under various parameterisations.

An important weakness of our model is that unlike Townsend (1979), we discretise the state space as in Mookherjee and Png (1989). This means that the probability of default in equilibrium is at best drawn from a discrete distribution and at worst fixed at the probability of a low draw in the two-state model. This discretisation is required for our analysis given the non-convexity of the incentive constraints and the absence of a single-crossing condition on these constraints. Unfortunately, this makes the model unsuitable for econometric analysis of the determinants of time-varying default probabilities. This also restricts the variability of credit spreads somewhat. In our model, credit spreads are more responsive to preferences for the sharing of risks across fixed probability default and no-default states than to the fluctuations in the expected resource costs of auditing (which drive fluctuations in credit spreads in the model described by Bernanke, Gertler, and Gilchrist, 1999).

On the other hand, we are able to consider optimal contracts, without imposing any ad-hoc restrictions on strategies, and the restriction to the simple two-state space results in tractable solutions. Further, as we show in Chapters 3 and 4, this model provides a straightforward way to generalise the preferences of entrepreneurs within financial macroeconomics models. Specifically, we show in Chapter 3 that this framework provides intuition and microfoundation to the well established empirical link between financial stress and the labour market wedge. In Chapter 4 we show that this framework can help us understand the perceived lack of systemic risk sharing, and the consequences for business cycle volatility and stabilisation policy.

2.1.2 Commitment

This paper, along with the aforementioned studies, considers an environment where the lender is able to commit ex ante to an incentive regime which is wasteful ex post. That commitment may indicate a concern for reputation, or delegation to a specialised auditor or bankruptcy court as in Melumad and Mookherjee (1989). Krasa and Villamil (2000) investigate what happens when lenders cannot commit to costly audits. That lack of commitment means the revelation principle does not hold and in equilibrium borrowers misreport their income with positive probability. It turns out that lack of commitment means that determinstic audits may be a feature of the optimal contract. Audits can only occur if the expected value of penalties levied following audits exceeds the audit costs. If true for a particular reported income, then this report will be audited with certainty. In short, for Krasa and Villamil (2000) the ability to commit implies equity-like contracts are preferable, whereas for us it does not.

The rest of the paper is set out as follows. Section 1 lays out the model environment and the nature of the auditing technology. Section 2 characterizes some key features of efficient contracts. In section 3 we present the perfect audits benchmark. Section 4 explores the imperfect audits case, and contains the key contributions of the paper. Section 5 presents comparative statics for a special case of the model where closed form solutions can be obtained. Section 6 provides a numerical example of a four state version of the model. Efficient contracts under perfect and imperfect audits are compared. Section 7 offers concluding remarks. Appendices contain formal arguments and proofs. Figures are contained in Appendix 2.E.

2.2 The environment

We study the one period problem of a risk averse and credit constrained entrepreneur. The entrepreneur has access to a special technology offering high returns which are uncorrelated with other projects undertaken in the economy.

The outcome of the project is initially private information to the entrepreneur, limiting the sharing of risk between the entrepreneur and their financier (the financial intermediary). Contract repayments are enforceable, but can only be conditioned on public information. The public information available to condition contracts includes any message sent by the entrepreneur, and any audit signal produced by the audit technology.

The entrepreneur makes a take-it-or-leave it contract offer to the financial intermediary, who is well-diversified and perfectly competitive. An efficient contract maximises the entrepreneur's expected utility.

2.2.1 THE ENTREPRENEUR

The entrepreneur enjoys consumption at the end of the period according to $U(x)$, where $U', -U'' > 0$, and $U'(0) = \infty$. The entrepreneur brings wealth α of the consumption good into the period. Combining the entrepreneur's wealth α with the net funds borrowed from the financial intermediary b, the project produces the consumption good according to stochastic gross return $(\alpha + b)\theta$. In this section, we restrict revenue to be drawn from one of two states, $\theta \in \Theta$, $\Theta = \{\bar{\theta}, \underline{\theta}\}$ and $\bar{\theta} > \underline{\theta}$. This restriction assists the intuition behind our key results, but is not essential for them. Section 6 extends the model to a four-state version, exploring optimal risk sharing across states in the perfect and imperfect audit models.

Following the realisation of their project, the entrepreneur can send a public signal indicating the state and subsequent revenues of the project. Messages m are drawn from $M = \{\bar{m}, \underline{m}\}$, where a message of $\bar{m}$ corresponds to reporting that the entrepreneur has received a high type shock $\bar{\theta}$, and a report of $\underline{m}$ implies a low type shock, $\underline{\theta}$.

2.2.2 THE FINANCIAL INTERMEDIARY

There exists a well-diversified financial intermediary who can make credible commitments to future actions.[10] Any contract involving the entrepreneur and the financial intermediary is small from the perspective of the financial intermediary's balance sheet. Further, the entrepreneur's return shock θ is uncorrelated with other shocks in the economy, and the returns of other assets/liabilities of the financial intermediary's balance sheet. It follows that the financial intermediary is risk neutral with respect to claims contingent on

[10] Efficient contracts will require commitment on behalf of the financial intermediary. One might think of this as sustained either through the intermediary's concern for its reputation, or through delegation to a specialist bailiff or auditor as in Melumad and Mookherjee (1989).

the entrepreneur's return shock θ.

The financial intermediary operates in a perfectly competitive market. Their opportunity cost of funds is given by ρ, and any contract offering an expected return on possibly state contingent loans exceeding ρ is acceptable to the financial intermediary. This condition is formalised in Definition 2.5. The opportunity cost of funds could be thought of as some combination of the interest rate paid by a risk free bond, the interest rate paid by the intermediary to their deposit holders, and the intermediary's administrative costs.

The following two assumptions ensure that there are available positive (but finite) gains from trade between the entrepreneur and financial intermediary.

Assumption 2.1 *Expected project returns exceed the financial intermediary's opportunity cost of funds,* $\sum_{\theta\in\Theta} \pi(\theta)\theta > \rho$.

Assumption 2.2 *In the low state, project returns are lower than the financial intermediary's opportunity cost of funds,* $\underline{\theta} < \rho$.

Assumption 2.1 ensures that there are economic gains from diverting resources to the entrepreneur's project, even when the entrepreneur has access to a deposit facility at the bank yielding a risk free return equal to the bank's opportunity cost of funds, ρ. Assumption 2.1 is strong enough to ensure that $b > -\alpha$.

Assumption 2.2 specifies that the entrepreneurs' projects are risky. In bad states, a project will yield lower returns than the risk free asset. Assumption 2.2 will be sufficient to ensure that leverage is finite under efficient contracts when type-I audit errors are present, a result shown in Proposition 2.3.

2.2.3 Audits

There exists an audit technology which produces a signal $\sigma \in \Sigma$ providing information about the outcome of the entrepreneur's project ex post. The action to undertake an audit is common knowledge, and so is the signal provided, σ. In other words, the entrepreneur knows if (s)he has been audited, and the result of the audit. It is assumed that an audit strategy, contingent on the entrepreneur's reports, can be agreed and committed to ex ante.

The audit technology cost is linear in assets, $\kappa(\alpha + b)$, where κ is a fixed parameter. The signal produced by the audit technology maps from the space of realised shocks θ as follows: If there is no audit, the audit signal is the empty set, $\sigma = \emptyset$. If there is an audit and the true state is $\bar{\theta}$, the audit technology reports $\sigma(\bar{\theta}) = \bar{\sigma}$ with probability $(1 - \eta(\bar{\theta}))$, and $\sigma(\bar{\theta}) = \underline{\sigma}$ with probability $\eta(\bar{\theta})$. If there is an audit and the true state is $\underline{\theta}$, the audit technology reports $\sigma(\underline{\theta}) = \underline{\sigma}$ with probability $(1 - \eta(\underline{\theta}))$, and $\sigma(\underline{\theta}) = \bar{\sigma}$ with probability $\eta(\underline{\theta})$.

Assumption 2.3 *Audits are informative:* $\eta(\bar{\theta}) + \eta(\underline{\theta}) < 1$.

Definition 2.1 *Conditional upon an audit, a* Type-I error *occurs when the audit technology signals a high type return when the true return is low,* $\sigma(\underline{\theta}) = \bar{\sigma}$. *A* Type-II error *occurs when the audit technology signals a low type return when the true return is high,* $\sigma(\bar{\theta}) = \underline{\sigma}$.

Audit strategies are defined in contracts, and implemented ex post by the financial intermediary. An audit strategy specifies the probability of audit, conditional upon the message sent by the entrepreneur, $q(m)$.

2.3 CONTRACTS

Definition 2.2 *A* contract *is an ordered set* $\Gamma = (b, q(m), z(m, \sigma), x(m, \sigma, \theta))$ *where* $b, q(m)$ *are publicly observed actions;* $z(m, \sigma) : \Theta \times \Sigma \to \mathbb{R}$ *is a function mapping publicly observed information to the financial intermediary's ex post receipt from the entrepreneur; and the entrepreneur's consumption allocations are specified by* $x(m, \sigma, \theta) : M \times \Sigma \times \Theta \to \mathbb{R}^+$.

A key motivation for this paper is the search for environments where debt contracts are efficient.

Definition 2.3 *We specify the following two benchmark contracts.*

a. A non-contingent debt *contract is a contract with constant repayments across all states and messages* $z(m_i, \sigma_j) = z(m_k, \sigma_l)\ \forall m_i, m_k \in \mathcal{M},\ \sigma_j, \sigma_l \in \Sigma$. *Any available audit signals are ignored, and therefore no audits are conducted (*$q(\underline{m}) = 0$*).*

b. *A* standard debt *contract specifies a constant repayment when either the entrepreneur's message or the audit signal is high, and a lower repayment following a verified low report* ($z(\bar{m}, \emptyset) = z(\underline{m}, \bar{\sigma}) > z(\underline{m}, \underline{\sigma})$). *All low reports are audited* ($q(\underline{m}) = 1$).

Note that debt contracts in our model do not restrict the entrepreneur borrower to zero consumption following default. In fact, in the examples that we consider, entrepreneurs will enjoy strictly positive consumption in all circumstances, even following a default. This positive consumption could represent income already paid to the entrepreneur during the life of the project, or rights to future earned income after the discharging of debts in bankruptcy.

Budget Constraints *State contingent budget constraints are specified as follows:*

$$(\alpha + b)\theta = z(m, \sigma) + x(m, \sigma, \theta) \qquad \forall (m, \sigma, \theta) \in M \times \Sigma \times \Theta. \tag{2.1}$$

The left hand side is the revenue received by the entrepreneur from their project, denominated in the consumption good. Following the repayment z, the remainder available for the entrepreneur to consume is x.

Definition 2.4 *A contract is* incentive compatible *if and only if* $m^*(\underline{\theta}) = \underline{m}, m^*(\bar{\theta}) = \bar{m}$ *solves the following problem:*

$$\begin{aligned} m^*(\theta) \in \arg\max_{m(\theta)} \, & (1 - q(m(\theta)))U(x(m(\theta), \emptyset, \theta)) + q(m(\theta))[1 - \eta(\theta)]U(x(m(\theta), \sigma = \theta, \theta)) \\ & + q(m(\theta))\eta(\theta)U(x(m(\theta), \sigma \neq \theta, \theta)), \qquad \theta \in \{\underline{\theta}, \bar{\theta}\} \end{aligned} \tag{2.2}$$

The consumption allocations on the right hand side of equation 4.10 are bundles enjoyed by misreporting agents. We can re-write the incentive compatibility constraint with respect to bundles consumed in truth-telling contracts by substituting in the budget constraints (2.1):

$$\begin{aligned} m^*(\theta) \in \arg\max_{m(\theta)} \, & (1 - q(m(\theta)))U(x(m(\theta), \emptyset, \theta)) \\ & + q(m(\theta))[1 - \eta(\theta)]U[x(m(\theta), \sigma = \theta, m(\theta)) + (\alpha + b)(\theta - m(\theta))] \\ & + q(m(\theta))\eta(\theta)U[x(m(\theta), \sigma \neq \theta, m(\theta)) + (\alpha + b)(\theta - m(\theta))], \qquad \theta \in \{\underline{\theta}, \bar{\theta}\} \end{aligned} \tag{2.3}$$

Definition 2.5 *A contract is* acceptable *to the financial intermediary if and only if*

$$\sum_{m\in\Theta}\Delta(m)\left[\sum_{\sigma\in\Theta\cup\emptyset}\Delta(\sigma|m,q(m))z(m,\sigma)-q(m)(\alpha+b)\kappa\right]\geq b\rho, \tag{2.4}$$

where $\Delta(\cdot)$ is an operator generating unconditional probability measures over its arguments. The state of nature θ is unobservable to the financial intermediary, therefore expectations in (2.4) are formed over the probability measure constructed over the entrepreneur's possible reports, which combines the likelihoods of shocks θ with the entrepreneur's ex post best response reporting strategy.

Definition 2.6 *A contract is* feasible *if and only if it is acceptable, and satisfies the budget constraints.*

Definition 2.7 *An incentive compatible contract is* efficient *if and only if it maximises the entrepreneur's utility subject to feasibility*

$$\max_{\Gamma}\sum_{\theta\in\Theta}\pi(\theta)\sum_{\sigma\in\Sigma\cup\emptyset}\Delta(\sigma|m,q(m))U(x(m,\sigma,\theta)) \tag{2.5}$$

subject to

$$(2.1),\ (4.10),\ (2.4),\quad q(m)\in[0,1].$$

Proposition 2.1 *In any efficient contract:*

1. *The financial intermediary's participation constraint (2.4) is binding,*
2. *high type reports are never audited,* $q(\bar{m})=0$*, and*
3. *the downward incentive compatibility constraint (equation 4.10, where* $\theta=\bar{\theta}$*) is binding when either (a) Type-I audit errors occur* $(\eta(\underline{\theta})>0)$*, or (b) utility is bounded below* $(U(0)=0)$*.*

A short description of Proposition 2.1 follows, while Appendix 2.A provides formal perturbation arguments for Parts 2 and 3.

For Part 1, note that if the intermediary's participation constraint were slack, repayments following high type reports $z(\bar{m},\emptyset)$ could be reduced. That would increase the expected utility of the entrepreneur without breaching the incentive compatibility constraint.

The intuition for Part 2 is as follows: Let it be the case that audits are required to prevent low type agents from declaring high type reports. Such a contract must be increasing rather than reducing consumption risk relative to some strictly superior non-contingent contract.

Proposition 2.1 part 3 shows that if high type entrepreneurs strictly prefer to report truthfully their income, then it must be the case that either high type consumption could be transferred to low states, or the auditing probability and expense could be reduced, allowing a direct increase in expected utility, or relaxing the participation constraint respectively. Note that the proof provided for Proposition 2.1 part 3(a) does not require the entrepreneur to freely choose the audit probability q. When audit signals are imperfect, the incentive compatibility constraint is binding even when all low reports are audited.

Corollary 2.1 *Let the audit probability q be constrained arbitrarily. Subject to this constraint, the downward incentive compatibility constraint (equation 4.10, where $\theta = \bar{\theta}$) is binding for any efficient contract when Type-I audit errors occur* $(\eta(\underline{\theta}) > 0)$.

We can now re-write the problem as a Kuhn-Tucker problem:

$$\begin{aligned}
\mathcal{L} = {} & \bar{\pi}\, U(x(\bar{m}, \emptyset, \bar{\theta})) + \underline{\pi}(1 - q(\underline{m}))\, U(x(\underline{m}, \emptyset, \underline{\theta})) \\
& + \underline{\pi} q(\underline{m})(1 - \eta(\underline{\theta}))\, U(x(\underline{m}, \underline{\sigma}, \underline{\theta})) + \underline{\pi} q(\underline{m})\eta(\underline{\theta})\, U(x(\underline{m}, \bar{\sigma}, \underline{\theta})) \\
& + \lambda \left[\begin{array}{l} (\alpha + b)(\mathbb{E}(\theta) - \rho - \underline{\pi} q(\underline{m})\kappa) + \alpha\rho - \bar{\pi} x(\bar{m}, \emptyset, \bar{\theta}) - \underline{\pi}(1 - q(\underline{m}))x(\underline{m}, \emptyset, \underline{\theta}) \\ -\underline{\pi} q(\underline{m})(1 - \eta(\underline{\theta}))x(\underline{m}, \underline{\sigma}, \underline{\theta}) - \underline{\pi} q(\underline{m})\eta(\underline{\theta})x(\underline{m}, \bar{\sigma}, \underline{\theta}) \end{array} \right] \\
& + \mu \left[\begin{array}{l} U(x(\bar{m}, \emptyset, \bar{\theta})) - (1 - q(\underline{m}))U[x(\underline{m}, \emptyset, \underline{\theta}) + (\alpha + b)(\bar{\theta} - \underline{\theta})] \\ -q(\underline{m})(1 - \eta(\bar{\theta}))U[x(\underline{m}, \bar{\sigma}, \underline{\theta}) + (\alpha + b)(\bar{\theta} - \underline{\theta})] \\ -q(\underline{m})\eta(\bar{\theta})U[x(\underline{m}, \underline{\sigma}, \underline{\theta}) + (\alpha + b)(\bar{\theta} - \underline{\theta})] \end{array} \right] \\
& + \nu_0 q(\underline{m}) + \nu_1(1 - q(\underline{m})).
\end{aligned} \tag{2.6}$$

The Lagrange multipliers λ and μ are attached respectively to the participation constraint and the incentive compatibility constraint, and the Kuhn-Tucker multipliers, ν_0 and ν_1, to the upper and lower bounds on the probability of audit respectively. Proposition 2.1 ensures that the participation and incentive compatibility constraints are binding under any efficient contract. The upper and lower bounds on the audit probability $q(\underline{m})$ are occasionally binding constraints.

The first order necessary conditions are described in detail as they will be used at various points to establish certain facts about efficient contracts. Hence:

$$x(\bar{m},\emptyset,\bar{\theta}):\quad 0=\bar{\pi}\,U'(x(\bar{m},\emptyset,\bar{\theta}))-\lambda\bar{\pi}+\mu U'(x(\bar{m},\emptyset,\bar{\theta})) \tag{2.6a}$$

$$\begin{aligned}x(\underline{m},\emptyset,\underline{\theta}):\quad 0&=\underline{\pi}(1-q(\underline{m}))\,U'(x(\underline{m},\emptyset,\underline{\theta}))-\lambda\underline{\pi}(1-q(\underline{m}))\\&\quad-\mu(1-q(\underline{m}))U'[x(\underline{m},\emptyset,\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\end{aligned} \tag{2.6b}$$

$$\begin{aligned}x(\underline{m},\underline{\sigma},\underline{\theta}):\quad 0&=\underline{\pi}q(\underline{m})(1-\eta(\underline{\theta}))\,U'(x(\underline{m},\underline{\sigma},\underline{\theta}))-\lambda\underline{\pi}q(\underline{m})(1-\eta(\underline{\theta}))\\&\quad-\mu q(\underline{m})\eta(\bar{\theta})U'[x(\underline{m},\underline{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\end{aligned} \tag{2.6c}$$

$$\begin{aligned}x(\underline{m},\bar{\sigma},\underline{\theta}):\quad 0&=\underline{\pi}q(\underline{m})\eta(\underline{\theta})\,U'(x(\underline{m},\bar{\sigma},\underline{\theta}))-\lambda\underline{\pi}q(\underline{m})\eta(\underline{\theta})\\&\quad-\mu q(\underline{m})(1-\eta(\bar{\theta}))U'[x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\end{aligned} \tag{2.6d}$$

$$\begin{aligned}b:\quad 0&=\lambda(\mathbb{E}(\theta)-\rho-\underline{\pi}q\kappa)\\&\quad-\mu(\bar{\theta}-\underline{\theta})\left[\begin{array}{l}(1-q(\underline{m}))U'[x(\underline{m},\emptyset,\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\\+q(\underline{m})(1-\eta(\bar{\theta}))U'[x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\\+q(\underline{m})\eta(\bar{\theta})U'[x(\underline{m},\underline{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\end{array}\right]\end{aligned} \tag{2.6e}$$

$$\begin{aligned}q(\underline{m}):\quad 0&=-\underline{\pi}\,U(x(\underline{m},\emptyset,\underline{\theta}))+\underline{\pi}(1-\eta(\underline{\theta}))\,U(x(\underline{m},\underline{\sigma},\underline{\theta}))+\underline{\pi}\eta(\underline{\theta})\,U(x(\underline{m},\bar{\sigma},\underline{\theta}))\\&\quad+\lambda\left[\underline{\pi}x(\underline{m},\emptyset,\underline{\theta})-\underline{\pi}(1-\eta(\underline{\theta}))x(\underline{m},\underline{\sigma},\underline{\theta})-\underline{\pi}\eta(\underline{\theta})x(\underline{m},\bar{\sigma},\underline{\theta})\right]\\&\quad+\mu\left[\begin{array}{l}+U[x(\underline{m},\emptyset,\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\\-(1-\eta(\bar{\theta}))U[x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\\-\eta(\bar{\theta})U[x(\underline{m},\underline{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\end{array}\right]\\&\quad-\lambda(\alpha+b)\underline{\pi}\kappa+\nu_0-\nu_1.\end{aligned} \tag{2.6f}$$

The general problem is non-convex, owing to the uncertainty faced by misreporting high type agents. Numerical results in the following sections indeed confirm that multiple locally efficient contracts may result.

2.4 PERFECT AUDITS

In the introduction we stated that the interaction between leverage and costly, *imperfect* audits underpins the optimality of deterministic contracts. Before establishing that, and other, results it is insightful to analyse the case of *perfect* audits. We find, as did Mookherjee and Png (1989), that debt contracts are not optimal. Moreover, we go on to show that optimal leverage is unbounded, absent other restrictions.

Theorem 2.1 *(Mookherjee and Png (1989)) When audits yield correct signals with certainty ($\eta(\underline{\theta}) = \eta(\bar{\theta}) = 0$), standard debt contracts are inefficient, $q^*(\underline{m}) \neq 1$.*

The proof proceeds as follows. First, Lemma 2.1 shows that when audits are perfect, any allocation which is feasible under a standard debt contract can be achieved while the incentive compatibility constraint is slack. Then, Proposition 2.1 part 3 shows that the downward incentive compatibility constraint cannot be slack under any efficient contract. Therefore, all allocations which are feasible under a standard debt contract are inefficient.

Lemma 2.1 *Let audits yield correct signals with certainty ($\eta(\underline{\theta}) = \eta(\bar{\theta}) = 0$), consumption be positive in all states $x(\cdot) > 0$, and the probability of audit of low type reports be equal to 1, $q(\underline{m}) = 1$. Any feasible allocation can be implemented with the incentive compatibility constraint slack.*

Proof. To prove Lemma 2.1, first re-write the downward incentive compatibility constraint (equation 4.10, where $\theta = \bar{\theta}$), with $q = 1$ as follows: $U(x(\bar{m}, \emptyset, \bar{\theta}) \geq U[(\alpha + b)\bar{\theta} - z(\underline{m}, \bar{\sigma})]$. Under perfect audits, the observable pair $(\underline{m}, \bar{\sigma})$ correctly identifies misreporting high type entrepreneurs with certainty. Under a truth-telling equilibrium, the repayment $z(\underline{m}, \bar{\sigma})$ is not made by any agent. Any feasible allocation can be perturbed by increasing $z(\underline{m}, \bar{\sigma})$, which does not affect the participation constraint of the financial intermediary, does not affect the ex ante welfare of the entrepreneur, but does ensure that the incentive compatibility constraint is slack. ■

Proposition 2.2 *When audits are perfect ($\eta(\underline{\theta}) = \eta(\bar{\theta}) = 0$), sufficiently inexpensive ($\kappa < (\mathbb{E}(\theta) - \rho)/\underline{\pi}$) and projects enjoy constant returns to scale, efficient leverage and entrepreneurial consumption are infinite.*

Proof. Set the probability of audit equal to one, $q = 1$. Substituting the budget constraints (2.1) into the participation constraint (2.4) shows that when $\kappa < (\mathbb{E}(\theta) - \rho)/\underline{\pi}$, expected consumption will be rising in b. Lemma 2.1 states that the incentive compatibility constraint need not bind for any allocation satisfying the participation constraint, given any level of borrowing b. ■

Under perfect audits, auditing with a high probability allows us to equate the entrepreneur's ex post marginal utility across all states, regardless of leverage. When audits are sufficiently inexpensive, higher leverage permits higher entrepreneurial consumption in all states. Leverage in equilibrium is only bound by decreasing technological returns to scale, as in Gale and Hellwig (1985), or through general equilibrium effects.

If the probability of audit is sufficiently high, large penalties charged against misreporting high type entrepreneurs ensure that the incentive compatibility constraint is slack for any schedule of positive consumption allocations earned with positive probability. Any further audits would be wasteful, as the resource costs of additional audits would tighten the participation constraint of the fiinancial intermediary, and as the incentive compatibility constraint was already slack, no further risk sharing gains would be available from the additional audits.

Efficient allocations require that any agent who earns a low type return, declares their return truthfully, yet receives a high type audit signal $(\underline{m}, \bar{\sigma}, \underline{\theta})$ should face a repayment greater than their revenue $z(\underline{m}, \bar{\sigma}) > (\alpha + b)\underline{\theta}$. However, this outcome occurs with zero probability when audits are perfect.

2.5 IMPERFECT AUDITS

When type-I errors occur with positive probability ($\eta(\underline{\theta}) > 0$), Lemma 2.1 and subsequently Theorem 2.1 cease to hold; the outcome $(\underline{m}, \bar{\sigma}, \underline{\theta})$ occurs with positive probability in any contract with auditing ($q^*(\underline{m}) > 0$). Even if the audit probability is high, a further increase in the audit probability does increase the set of feasible consumption allocations available to the entrepreneur. Increasing the probability of audit under imperfect audits allows the incentive costs of contract enforcement to be defrayed more widely, increasing risk sharing across states. If audit costs are low, then increasing the probability of audit is worthwhile even when the probability of audit is already high.

Theorem 2.2 *Let borrowing be taken as given $b = \hat{b}$. When type-I audit errors occur with positive probability ($\eta(\underline{\theta}) > 0$), there exists some strictly positive audit cost $\hat{\kappa}$ such that for all $\kappa < \hat{\kappa}$, standard debt contracts ($q(\underline{m}) = 1$) are efficient.*

Proof. We consider an arbitrary efficient contract with interior audit probability $q(\underline{m}) \in (0, 1)$, and show from the first order condition for the probability of audit $\mathcal{L}_{q(\underline{m})}$ that if audit costs were sufficiently low, the initial contract could be strictly improved by an increase in audit probability $q(\underline{m})$.

To simplify notation, in this section we will define $B(b) := (\alpha + b)(\bar{\theta} - \underline{\theta})$. Also, as we are only considering allocations consistent with truth-telling, we will drop the report variable from the consumption allocation $x(\sigma, \underline{\theta}) := x(\underline{m}, \sigma, \underline{\theta})$.

Consider the first order necessary condition for $q(\underline{m})$ (2.6f), which can be re-written as follows:

$$\begin{aligned} \mathcal{L}_{q(\underline{m})}: \quad 0 = {} & \underline{\pi}(1 - \eta(\underline{\theta}))\left[U(x(\underline{\sigma}, \underline{\theta})) - \lambda x(\underline{\sigma}, \underline{\theta})\right] - \mu\eta(\bar{\theta})U[x(\underline{\sigma}, \underline{\theta}) + B(b)] \\ & + \underline{\pi}\eta(\underline{\theta})\left[U(x(\bar{\sigma}, \underline{\theta})) - \lambda x(\bar{\sigma}, \underline{\theta})\right] - \mu(1 - \eta(\bar{\theta}))U[x(\bar{\sigma}, \underline{\theta}) + B(b)] \\ & - \underline{\pi}\left[U(x(\emptyset, \underline{\theta})) + \lambda x(\emptyset, \underline{\theta})\right] + \mu\, U[x(\emptyset, \underline{\theta}) + B(b)] \\ & - \lambda(\alpha + b)\underline{\pi}\kappa + \nu_0 - \nu_1. \end{aligned} \tag{2.7}$$

Up to division by $q(\underline{m})$, the consumption variables $x(\underline{\sigma}, \underline{\theta}), x(\bar{\sigma}, \underline{\theta})$ enter $\mathcal{L}_{q(\underline{m})}$ in the same way that they enter the entrepreneur's problem $\mathcal{L}$ (equation 2.6). This means that the first order necessary conditions for $x(\underline{\sigma}, \underline{\theta}), x(\bar{\sigma}, \underline{\theta})$ (equations 2.6c and 2.6d respectively), also identify a stationary point of $\mathcal{L}_{q(\underline{m})}$, with respect to the consumption allocations of audited agents ($x(\underline{\sigma}, \underline{\theta}), x(\bar{\sigma}, \underline{\theta})$) and holding other variables constant.

This property has a straightforward economic interpretation: efficiently allocating consumption to audited agents, is the same problem as maximising the gain from additional audits, which is expressed by the first order condition $\mathcal{L}_{q(\underline{m})}$.

We can think of the first three lines of (2.7) as the gains attained from the information provided by a marginal increase in the probability of audit. The fourth line contains the marginal resource cost, plus Kuhn-Tucker multipliers associated with the upper and lower bounds on the audit probability.

Here it is important that audits are imperfect, which by Corollary 2.1 ensures that the incentive compatibility constraint is binding, and $\mu > 0$, regardless of $q(\underline{m})$. Were audits perfect, sufficiently high audit probabilities would result in slackness in the incentive compatibility constraint ($\mu = 0$), leaving the first order conditions for $x(\underline{\sigma}, \underline{\theta}), x(\emptyset, \underline{\theta})$ equated.

Consider the allocation $\hat{x}(\underline{\sigma}, \underline{\theta}) = \hat{x}(\bar{\sigma}, \underline{\theta}) = x(\emptyset, \underline{\theta})$. This allocation would leave the sum of the first three lines of (2.7) equal to zero. But this allocation is not a stationary point of $\mathcal{L}_{q(\underline{m})}$, and does not satisfy the first order necessary conditions for $x(\underline{\sigma}, \underline{\theta}), x(\bar{\sigma}, \underline{\theta})$ (equations 2.6c and 2.6d respectively).

We can do better by decreasing $x(\bar{\sigma}, \underline{\theta})$, which has a low weight in expected welfare and a high weight in the incentive compatibility constraint, and increasing $x(\underline{\sigma}, \underline{\theta})$, which has a relatively high weight in expected welfare and a low weight in the incentive compatibility constraint. This perturbation would leave the sum of the first three lines of (2.7) strictly greater than zero, such that for sufficiently low audit costs κ, additional audits would always be welfare enhancing. ■

Theorem 2.2 shows that standard debt can be efficient under imperfect audits. The remainder of this section explores the global efficiency of standard debt, and the quantitative relevance. As we will see, an important determinant of the efficiency of standard debt will be whether the entrepreneur has access to a leverage margin—enabling them to scale up and down the size of the project ex ante.

Under perfect audits, Proposition 2.2 showed that when audit costs are low, constant technological returns to scale would result in unbounded leverage and entrepreneurial consumption. When type-I audit errors occur with positive probability, that result no longer holds. To see this, note that the incentive compatibility constraint (2.3) ensures that for some $\sigma \in \{\emptyset, \underline{\sigma}, \bar{\sigma}\}$, $x(\bar{m}, \emptyset, \bar{\theta}) - x(\underline{m}, \sigma, \underline{\theta}) \geq (\alpha + b)(\bar{\theta} - \underline{\theta})$. Combining this with assumption 2 ensures that as b increases, consumption risk must be increasing and consumption in some state $x(\underline{m}, \sigma, \underline{\theta})$ must tend toward zero.

Entrepreneurs will not choose contracts with consumption bundles too close to zero, where their marginal utility of consumption tends to infinity. Entrepreneurs' aversion to low consumption bundles in bad states encourages them to choose contracts with limited leverage, even when their project enjoys constant technological returns to scale. This argument is formalised in the following Proposition.

Proposition 2.3 *When type-I errors occur with positive probability ($\eta(\underline{\theta}) > 0$), positive entrepreneurial consumption in all states requires that high type consumption, leverage and the probability of audit satisfy the following inequalities*

a. high type consumption, $x(\bar{m}, \emptyset, \bar{\theta}) > (\alpha + b)(\bar{\theta} - \underline{\theta})$,

b. leverage, $\dfrac{\alpha + b}{\alpha} < \dfrac{\rho}{\rho - \underline{\theta} + \underline{\pi} q \kappa}$, *and* *c. the probability of audit,* $q < \min\left(1, \frac{1}{\underline{\pi}\kappa}\frac{\alpha\rho}{\alpha+b}\left[1 - \frac{\alpha+b}{\alpha}\cdot\frac{\rho-\underline{\theta}}{\rho}\right]\right)$.

Proof. The first part of Proposition 2.3 is a direct weakening of equation 2.3, which is presented in a form such that all consumption bundles contained in the constraint are earned with positive unconditional probability in contracts with auditing, and are therefore positive by the assumption specified in the Proposition.

$$x(\bar{m}, \emptyset, \bar{\theta}) > (\alpha + b)(\bar{\theta} - \underline{\theta}) \tag{2.8}$$

Substituting equation 2.8 and the budget constraints (2.1) into the participation constraint (2.4) with the assumption that entrepreneurial consumption is positive in every state yields the following inequality:

$$(\alpha + b)(\bar{\pi}\bar{\theta} + \underline{\pi}\underline{\theta} - \rho - \underline{\pi}q\kappa) + \alpha\rho > \bar{\pi}(\alpha + b)(\bar{\theta} - \underline{\theta}) \tag{2.9}$$

which can be rearranged to confirm parts (b) and (c) of the Proposition. ■

We proceed allowing borrowing b to be chosen freely, under the assumption of constant technological returns to scale. This does not mean that firms enjoy constant returns to scale. Firm size is endogenously bounded above according to Proposition 2.3. It does mean that the only source of decreasing returns to scale is the information asymmetry between the entrepreneur and external finance providers.

Figure 2.1 presents evidence of the quantitative relevance of standard debt in our framework for a sample parameterisation. Along the horizontal axis, the risk of the entrepreneur's project $(\bar{\theta} - \underline{\theta})$ is increasing, holding expected returns $(\mathbb{E}(\theta))$ constant. The vertical axis plots audit costs as a share of total assets under management. Two features of the simulation are striking: First, standard debt $(q^* = 1)$ is very prevalent. Very low project risk or high audit costs are required for standard debt to be inefficient. Second, stochastic audit regimes $(0 < q^* < 1)$ are rare. Indeed, when risk is low, efficient contracts 'jump' from standard debt $(q^* = 1)$ to non-contingent debt contracts $(q^* = 0)$. In the model, there is no cost associated with writing a 'complex' contract with stochastic audit regimes, as are optimal under the perfect audits framework. Yet, entrepreneurs tend to prefer 'simple', non-contingent or standard debt contracts.

When contracts are constrained by the upper and lower bound on audits ($q^* = 0$ *or* 1), local analysis of the entrepreneur's problem is relatively straightforward, yielding closed form solutions under constant relative risk aversion when the likelihood of type-II audit error is zero ($\eta(\bar{\theta}) = 0$):[11]

Proposition 2.4 *When the likelihood of type-I and type-II errors are positive and zero respectively ($\eta(\underline{\theta}) > 0, \eta(\bar{\theta}) = 0$) and preferences exhibit constant relative risk aversion $U(x) = x^{1-\gamma}/(1-\gamma)$, leverage, consumption allocations and shadow prices of standard debt contracts and non-contingent debt contracts can be represented by closed-form expressions in terms of exogenous parameters.*

The proof of Proposition 2.4 is given in appendix 2.B. Displayed below are solutions to standard debt contracts under logarithmic utility ($U(x) = \log x$).

$$x(\bar{m}, \emptyset, \bar{\theta}) = \alpha\rho \, \frac{1}{1-\zeta}, \qquad x(\underline{m}, \underline{\sigma}, \underline{\theta}) = \alpha\rho, \qquad x(\underline{m}, \bar{\sigma}, \underline{\theta}) = \alpha\rho \, \frac{\underline{\pi}\eta(\underline{\theta})}{\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta})},$$

$$b = \frac{\alpha\rho}{\bar{\theta} - \underline{\theta}} \, \frac{\zeta}{1-\zeta} \left(\frac{\bar{\pi} + \underline{\pi}\eta(\underline{\theta})}{\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta})} \right) - \alpha, \qquad \text{where} \qquad \zeta = \frac{\mathbb{E}(\theta) - \rho - \underline{\pi}\kappa}{\bar{\pi}(\bar{\theta} - \underline{\theta})}. \tag{2.10}$$

From the solutions presented in (2.10), we can derive measures of leverage and loan coupon rates, which are more easily observed than entrepreneurs' consumption allocations in practise. Leverage, l, as measured by the total assets managed by the entrepreneur over their initial net worth can be described as follows:

$$l = \frac{\alpha + b}{\alpha} = \frac{\rho}{\bar{\theta} - \underline{\theta}} \, \frac{\zeta}{1-\zeta} \left(\frac{\bar{\pi} + \underline{\pi}\eta(\underline{\theta})}{\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta})} \right). \tag{2.11}$$

We can determine the net interest (coupon) rate on loans, r, by subtracting one from the ratio of the repayment following high reports $z(\bar{m}, \emptyset)$ and the initial amount borrowed b. Combining the budget constraints with (2.10) yields

$$r = \frac{z(\bar{m}, \emptyset)}{b} - 1 = (\rho - 1) + \rho\zeta \, \frac{(\bar{\theta} - \rho)(\bar{\pi} + \underline{\pi}\eta(\underline{\theta})) - (\bar{\theta} - \underline{\theta})(\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta}))}{\rho\zeta(\bar{\pi} + \underline{\pi}\eta(\underline{\theta})) - (1-\zeta)(\bar{\theta} - \underline{\theta})(\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta}))}. \tag{2.12}$$

The first term on the right hand side, $(\rho - 1)$, is the opportunity cost of funds for the financial intermediary, expressed as a net interest rate. The second term captures the interest rate credit spread, as measured by the difference between the loan interest (coupon) rate and the financial intermediary's opportunity cost of funds. Section 2.6 presents example comparative statics for leverage ratios and the loan interest rate.

The solutions obtained by Proposition 2.4 are local, though the entrepreneur's problem exhibits 'jumps' between locally efficient contracts. Fact 2.1 states that for a standard debt

[11] Type-II errors, while perhaps more familiar than type-I errors, have little effect on the nature of efficient contracts if they occur with a low probability. The implications of type-II errors are investigated in appendix 2.D.

contract to be globally efficient, it must be both locally efficient, and superior to any non-contingent contract. While these two necessary conditions do not rule out an alternative globally efficient contract, we have not been able to find a numerical example where the two necessary conditions expressed in fact 2.1 are satisfied, and standard debt contracts are not globally efficient.

Fact 2.1 *Let Γ be a globally efficient contract, and $q(\Gamma) = 1$. (a) The Kuhn-Tucker multiplier on the constraint $q \leq 1$ must be positive ($\nu_1 > 0$), and (b) expected utility under Γ must exceed the maximum utility attainable conditional upon private information ($\mathbb{E}(U|\Gamma) \geq \mathbb{E}(U|(\Gamma^*|q = 0))$).*

When utility is CRRA and type-II errors do not occur ($\eta(\bar{\theta}) = 0$), then by Proposition 2.4 we can check fact 2.1 part (b) directly from the closed-form solutions to standard and non-contingent debt contracts provided in appendix 2.B. When utility is logarithmic, we can also directly check fact 2.1 part (a) by the following result:

Proposition 2.5 *When utility is logarithmic, and type-II errors do not occur ($\eta(\bar{\theta}) = 0$), the Kuhn-Tucker multipliers on the upper and lower bounds for the audit probability, $0 \leq q \leq 1$, can be described by closed-form expressions.*

A derivation of Proposition 2.5 is provided in Appendix 2.B.3.

We now return to reconsider the bang-bang feature of efficient contracts observed in figure 2.1, which we formalise by Proposition 2.6.

Proposition 2.6 *When audits are imperfect, there exist parameter specifications which permit both non-contingent ($q = 0$) and standard ($q = 1$) debt contracts as locally efficient contracts.*

When project risk and audit costs are low, efficient contracts appear to jump from non-contingent debt to standard debt. Figure 2.2 plots an example of this bang-bang behaviour. In figure 2.2, the determination of efficient contracts is deconstructed by leverage and audit strategy for one example parameter specification. The horizontal axis plots levels of borrowing. The lower panel plots the efficient probability of audit, conditional upon borrowing, and the upper panel plots the attainable expected welfare conditional upon borrowing. The solid line plots expected welfare attainable with a non-contingent contract ($q = 0$), the dashed line allows the audit strategy to be chosen optimally.

The efficiency of standard debt is sensitive to the assumption that the entrepreneur can determine the scale of the project. When leverage is low, total risk is low, and the gains from insurance provided by auditing are low. On the other hand, when leverage is high, Proposition 2.3 showed that auditing will push the minimum consumption allocation closer to zero. Audits are only useful to the extent that the entrepreneur can absorb type-I errors in low states.

Appendix 2.C solves an example where non-contingent and standard debt contracts are locally efficient. Under the non-contingent debt contract, the marginal resource cost of additional audits exceeds the gains obtained from the audit signal. Under the standard debt contract, the marginal resource gain from reducing the audit probability is smaller than the cost of foregoing the information and incentive gains obtained via the marginal audit.

Leverage is higher under the non-contingent contract than under the efficient standard debt contract, and therefore the marginal resource cost of audits is higher than under the standard debt contract. The marginal benefit from information obtained in additional audits is actually identical under the two locally efficient contracts considered. Under the non-contingent contract, the difference in expected marginal utility across project outcomes is high, suggesting that the gains from insurance should be higher than under the efficient standard debt contract. However, low consumption of low type entrepreneurs also makes type-I errors particularly costly, preventing significant penalties in auditing contracts, and reducing the benefits obtained by auditing.

2.6 Comparative statics

The preceding paragraphs explained how small parameter changes—for example a small increase in project risk—can cause the efficient contract to jump from a high leverage, low risk premium, non-contingent contract, to a low leverage, high risk premium and highly contingent standard debt contract.

Within parameter neighbourhoods where defautable debt contracts are efficient, we can use the solutions obtained in appendix 2.B to analyse local perturbations to expected returns, risk, audit costs and audit quality. Consider the following parameterisation: The probability of default is $\underline{\pi} = 1/10$; conditional upon realisation of the low state, the audit signal returns a high state with vanishing probability $\eta(\underline{\theta}) \to 0^+$; the gross opportunity cost of funds $\rho = 21/20$, equivalent to a 5% interest rate; the expected gross return on projects $\mathbb{E}(\theta) = 6/5$; the coefficient of relative risk aversion $\gamma = 1$; audit costs as a share of the initial assets devoted to the project are $\kappa = 9/80$, and in low states, the project returns $\underline{\theta} = 33/40$. Subsequently, the high type return is $\bar{\theta} = \frac{1}{\bar{\pi}}[\mathbb{E}(\theta) - \underline{\pi\theta}] = 149/120$; project risk is $(\bar{\theta} - \underline{\theta}) = 5/12$; and $\zeta = [\mathbb{E}(\theta) - \rho - \underline{\pi}\kappa]/[\bar{\pi}(\bar{\theta} - \underline{\theta})] = 37/100$.

By equation 2.11, leverage (l) is equal to

$$l = \frac{\rho}{\bar{\theta} - \underline{\theta}} \frac{\zeta}{1 - \zeta} \left(\frac{\bar{\pi} + \underline{\pi}\eta(\underline{\theta})}{\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta})} \right) = 4.$$

By equation 2.12, the loan interest (coupon) rate is

$$r = (\rho - 1) + \rho\zeta \frac{(\mathbb{E}(\theta) + \underline{\pi}(\bar{\theta} - \underline{\theta}) - \rho)(\bar{\pi} + \underline{\pi}\eta(\underline{\theta})) - (\bar{\theta} - \underline{\theta})(\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta}))}{\rho\zeta(\bar{\pi} + \underline{\pi}\eta(\underline{\theta})) - (1 - \zeta)(\bar{\theta} - \underline{\theta})(\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta}))} = 10\%,$$

We can take derivatives of the log of the leverage ratio to find the semi-elasticities of leverage with respect to expected returns, the intermediary's opportunity cost, risk, audit costs and type-I errors:

$$\frac{d \log l}{d(\mathbb{E}(\theta))} = 4.23, \qquad \frac{d \log l}{d\rho} = -3.28, \qquad \frac{d \log l}{d(\bar{\theta} - \underline{\theta})} = -3.81,$$

$$\frac{d \log l}{d\kappa} = -0.423 \qquad \text{and} \qquad \frac{d \log l}{d\eta(\underline{\theta})} = -0.189$$

respectively, reported to 3 significant figures.[12]

We can also determine the sensitivity of efficient loan interest rates to underlying parameters:

$$\frac{dr}{d(\mathbb{E}(\theta))} = -0.219, \qquad \frac{dr}{d\rho} = 1.20, \qquad \frac{dr}{d(\bar{\theta} - \underline{\theta})} = 0.197,$$

$$\frac{dr}{d\kappa} = 0.155 \qquad \text{and} \qquad \frac{dr}{d\eta(\underline{\theta})} = -0.0357.$$

Whilst these calculations are specific to our example, they do provide insights into the

[12]For example, an increase in expected returns by 0.01 causes an increase in the optimal leverage ratio by 4.23 percent. Note that rather than reporting semi-elasticities for returns in each state $(\bar{\theta}, \underline{\theta})$, we have reported responses to expected returns $\mathbb{E}(\theta)$ and risk $(\bar{\theta} - \underline{\theta})$, which we find to be more useful for intuition.

tradeoffs more generally faced by entrepreneurs.

An increase in expected project returns $\mathbb{E}(\theta)$ increases leverage and, perhaps surprisingly, decreases loan interest rates. The prospect of higher returns encourages entrepreneurs to increase leverage, but they are limited in doing so due to the presence of type-I errors. A decrease in interest payments in high states r leaves the entrepreneur enough funds in low states to repay loans in full following type-I errors. In order to satisfy the financial intermediary's participation constraint, providing an expected return of ρ on loans b net of auditing costs, the entrepreneur must absorb more project risk. Low state repayments following verified reports ($z(\underline{m}, \underline{\sigma})/b$) are increased. Decreases in project risk ($\bar{\theta} - \underline{\theta}$) and audit costs ($\kappa$) have a similar effect on leverage and interest rates as increases in expected returns $\mathbb{E}(\theta)$.

Following a decrease in the financial intermediary's opportunity cost of funds, loan interest rates fall by an even greater amount. In other words, the spread between loan interest rates and the opportunity cost of funds is increasing in the opportunity cost of funds. First, a decrease in ρ allows the entrepreneur to make lower repayments in all states, while meeting the intermediary's participation constraint. Second, when ρ is low, the entrepreneur enjoys more of the gains from increased leverage. As in the case of an increase in expected returns $\mathbb{E}(\theta)$, increases in leverage require the entrepreneur to further lower interest repayments such that full repayments are possible even following type-I errors. To compensate, repayments following verified low reports ($z(\underline{m}, \underline{\sigma})/b$) must increase.

The only variable which moves leverage and loan interest rates in the same direction is audit quality, as measured by the conditional probability of type-I error ($\eta(\underline{\theta})$). An increase in the probability of error encourages entrepreneurs to increase consumption in the unlikely event of a type-I error. This adjustment is achieved first by decreasing leverage, which at any given interest rate increases the amount of resources remaining following type-I errors, and secondly by decreasing the loan interest rate r, which further reduces repayments following type-I errors. This adjustment requires an increase in repayments in low states ($z(\underline{m}, \underline{\sigma})/b$) to compensate the financial intermediary.

2.7 A FOUR-STATE EXAMPLE

In this section we extend the model by increasing the number of states from two to four. Unfortunately, that makes the model analytically intractable.[13] The purpose of this extension is twofold. First, it is clearly of interest to investigate how general our analytical results are likely to be, concerning the desirability of standard debt. Second, and related, it is of interest to compare the imperfect audits case to that of the perfect audits case of Mookherjee and Png (1989).

Typically, standard debt is defined as a contract where reports below some cutoff are audited with certainty. See, for example Townsend (1979). Audits above the cutoff are not audited. Thus, we state:

Definition 2.8 *In a model with $n > 2$ possible states ($\theta \in (\theta_1, \theta_2, ..., \theta_n)$ and $\theta_i < \theta_{i+1}$), a contract is a standard debt contract if and only if*

$$\exists\, K \in \{1, 2, ..., n\} \quad s.t. \quad q(m_j) = 1 \quad \forall j < K, \quad q(m_j) = 0 \quad \textit{otherwise.} \tag{2.13}$$

Note that incentive compatibility requires that all repayments following reports above the cutoff must be identical

$$z(m_i, \emptyset) = z(m_j, \emptyset) \quad \forall i, j \geq K. \tag{2.14}$$

Definition 2.8 generalises definition 2.3(b) used to analyse the two-state model in earlier sections. An important feature of standard debt is that when income is sufficiently high, repayments are not sensitive to income—the borrower need not ever pay more than the coupon plus the principal. This is formalised in equation 2.14, and is not observable in the two-state model, where there is only one 'high' state. In this section with a multiple states model it turns out that imperfect audits do indeed motivate that feature of standard debt.

Figure 2.3 presents a numerically-solved example of locally efficient contracts under perfect and imperfect audits. In order to compare the incentive regimes under the two environments, borrowing b is set exogenously. The upper panel presents the probability

[13]Even with just four states, the problem contains 24 choice variables. We have been unable to find superior contracts to the examples presented below, but we do not prove that the contracts presented are globally efficient.

distribution from which states are drawn. The second panel presents the expected repayment conditional upon the true state being equal to θ. The contract with signal errors (marked by $\times$) features constant repayments across the three higher states, and a reduced expected repayment in the lowest state. The perfect audits contract (marked by $+$) exhibits sharply increasing repayments across states—similar to an equity contract with variable dividends. The third panel presents the expected utility of the borrower, conditional upon the realised state. The contract with signal errors exhibits significant sensitivity between expected utility and project outcomes, across all states. The contract with perfect audits exhibits increasing expected utility across states, although the sensitivity of utility across states is very low. The fourth panel presents the auditing regime under each contract. The contract with signal errors resembles a standard debt contract: reports of the lowest state are followed by certain audits ($q(m_1) = 1$). Reports of any higher states are not audited ($q(m_2) = q(m_3) = q(m_4) = 0$). Under perfect audits, audits are conducted with low probability across all of the three lower states.

2.8 DISCUSSION

Standard debt contracts can be the optimal form of external finance contracts when contract enforcement is uncertain due to noisy audit signals. Supporting truth-telling under a stochastic audit strategies requires large penalties. When there is no guarantee that these penalties are fairly applied, these contracts will not be acceptable to risk averse entrepreneurs. The resulting efficient contracts will audit consequent only on low reports, but will likely audit low reports with certainty. As a result, only small penalties are required to ensure truth-telling in equilibrium. In fact, the penalty following a disputed report in an optimal debt contract is typically very close to fully repaying the debt.

Imperfect verification also implies other interesting properties of optimal contracts. For instance, it means that borrowers can only pass a limited amount of risk on to lenders, regardless of contracted audit strategies. And even when projects enjoy constant returns to scale and audits are relatively inexpensive, firm size and leverage is endogenously limited by the entrepreneur's risk preference.

We end with a final observation. The standard debt contracts derived under imperfect monitoring enjoy an additional benefit—one that we did not formalise. When enforcement is certain, or near certain, incentive compatibility is not sensitive to the risk tolerance of the entrepreneur. That reduces the potential for adverse selection in two forms: First, the preferences of the entrepreneur may be unobservable; and second, the entrepreneur may have access to hidden wealth. The presence of either of these sources of asymmetric information would make it more difficult to employ a stochastic incentive scheme.

2.A PROOF OF PROPOSITION 2.1

Proof.

2. It is established that if it were the case that low type agents were indifferent to reporting high or low type messages, then the contract in place must be weakly inferior than a simple non-contingent debt contract. This argument is made via three perturbations which leave the expected utility of the entrepreneur either unchanged or increased, and relax the intermediary's participation constraint.

   Let the probability of audit following high type reports be positive, and the incentive compatibility constraint be binding for low type entrepreneurs: $q(\bar{m}) > 0$ and

$$\sum_{\sigma \in \Theta \cup \emptyset} \Delta(\sigma(\underline{\theta}, q(\underline{m})))U[(\alpha + b)\underline{\theta} - z(\underline{m}, \sigma(\underline{\theta}, q(\underline{m})))] = \sum_{\sigma \in \Theta \cup \emptyset} \Delta(\sigma(\underline{\theta}, q(\bar{m})))U[(\alpha + b)\underline{\theta} - z(\bar{m}, \sigma(\underline{\theta}, q(\bar{m})))]. \tag{2.15}$$

   Perturbation 1: First replace all $z(\bar{m}; \sigma)$ with $z'(\bar{m})$, such that

$$\sum_{\sigma \in \Theta \cup \emptyset} \Delta(\sigma(\bar{\theta}, q(\bar{m})))U[(\alpha + b)\bar{\theta} - z(\bar{m}, \sigma(\bar{\theta}, q(\bar{m})))] = U[(\alpha + b)\bar{\theta} - z'(\bar{m})].$$

   The perturbation leaves truth-telling high type entrepreneurs indifferent. By Jensen's inequality, that perturbation will relax the intermediary's participation constraint, but could possibly violate the incentive compatibility constraint (2.15). If (2.15) is now not violated, then the proof is complete, and $q(\bar{m})$ can be set to zero, as the information yielded by auditing high type messages is ignored.

   If (2.15) is violated after the perturbation, then

$$U[(\alpha + b)\underline{\theta} - z'(\bar{m})] > \sum_{\sigma \in \Theta \cup \emptyset} \Delta(\sigma(\underline{\theta}, q(\underline{m})))U[(\alpha + b)\underline{\theta} - z(\underline{m}, \sigma(\underline{\theta}, q(\underline{m})))].$$

   Perturbation 2: Now replace $z(\underline{m}, \sigma)$ with $z'(\underline{m})$, such that

$$U[(\alpha + b)\underline{\theta} - z'(\underline{m})] = \sum_{\sigma \in \Theta \cup \emptyset} \Delta(\sigma(\underline{\theta}, q(\underline{m})))U[(\alpha + b)\underline{\theta} - z(\underline{m}, \sigma(\underline{\theta}, q(\underline{m}))).$$

   By Jensen's inequality, that perturbation would also relax the participation constraint. Given that (2.15) is violated, $z'(\underline{m}) > z'(\bar{m})$: Low type entrepreneurs have an incentive to report high type shocks.

   Perturbation 3: Replace $z'(\underline{m}), z'(\bar{m})$ with $z'' = \underline{\pi} z'(\underline{m}) + \bar{\pi} z'(\bar{m})$

Perturbation 3 restores incentive compatibility, by equating repayments across reports and states. The participation constraint is respected, as expected repayments are unchanged. The new contract offers expected utility which is strictly greater than under the original contract, by Jensen's inequality:

$$\bar{\pi}U[(\alpha+b)\bar{\theta}-z'']+\underline{\pi}U[(\alpha+b)\underline{\theta}-z'']>\bar{\pi}U[(\alpha+b)\bar{\theta}-z'(\bar{m})]+\underline{\pi}U[(\alpha+b)\underline{\theta}-z'(\underline{m})].$$

Audits are not required, as repayments are non-contingent.

3(a). For the incentive compatibility constraint (4.10 ; $\theta=\bar{\theta}$) to be satisfied, there must be some consumption bundle $x(\underline{m},\sigma,\underline{\theta})<x(\bar{m},\emptyset,\bar{\theta})$, where $\sigma\in\{\emptyset,\bar{\sigma},\underline{\sigma}\}$, and $x(\underline{m},\sigma,\underline{\theta})$ is a bundle consumed with non-zero unconditional probability $(\Delta(\sigma|\underline{m},q(\underline{m}))>0)$. Let equation (4.10 ; $\theta=\bar{\theta}$) be slack. There must be some $\varepsilon\in(0,\infty)$ such that a perturbation increasing $z(\bar{m},\emptyset)$ by $\frac{\varepsilon}{\Delta(\emptyset|\bar{m},0)}$, and decreasing $z(\underline{m},\sigma)$ by $\frac{\varepsilon}{\Delta(\sigma|\underline{m},q(\underline{m}))}$, which would violate neither the participation nor the incentive compatibility constraints. This perturbation would increase expected utility by Jensen's inequality.

3(b). When audits are perfect, we cannot directly follow the proof of part 3(a)—the incentive compatibility constraint does not directly ensure that consumption x varies across states. First, if x does vary across states and audit signals, then we can follow the same argument as in part 3(a), and perturb toward a contract with less consumption variability. If x is constant across all states, then we could reduce the audit probability q, relaxing the participation constraint without violating the incentive compatibility constraint.

■

2.B Efficient allocations under non-contingent and standard debt contracts

Here we solve for efficient allocations and borrowing when efficient audit strategies are deterministic, ie. when $q=0$ *or* 1. Here, we assume CRRA utility, $U(x)=x^{1-\gamma}/(1-\gamma)$. In the text, we refer to the more tractable case of logarithmic utility, which can be found by setting $\gamma=1$ in any of the solutions contained in this section.

The general problem outlined by equation 2.6 is non-convex, due to the presence of the lottery in the right hand side of the incentive compatibility constraint. When the probability of audit $q(\underline{m})$ is constrained by either its upper or lower bound, ν_0 or $\nu_1 > 0$, and the probability of type-II audit errors is zero $\eta(\bar{\theta}) = 0$, then locally the problem is convex, and we can use the first order approach to find local maxima.

2.B.1 PRIVATE INFORMATION CONTRACTS ($q = 0$)

When audits are not used ($q = 0$), efficient contracts are non-contingent. Repayments are independent of entrepreneurs' reports, and no audit signals are obtained to condition repayments. This certainty of repayment makes the incentive compatibility constraint linear in the choice variables, enabling us to solve the entrepreneur's problem with a Lagrangian:

$$\begin{aligned}\mathcal{L}_0 = \bar{\pi}\, U(x(\bar{m}, \emptyset, \bar{\theta})) + \underline{\pi}\, U(x(\underline{m}, \emptyset, \underline{\theta}))\\ + \lambda[(\alpha + b)(\mathbb{E}(\theta) - \rho) + \alpha\rho - \bar{\pi}x(\bar{m}, \emptyset, \bar{\theta}) - \underline{\pi}x(\underline{m}, \emptyset, \underline{\theta})]\\ + \mu\left[U(x(\bar{m}, \emptyset, \bar{\theta})) - U[x(\underline{m}, \emptyset, \underline{\theta}) + (\alpha + b)(\bar{\theta} - \underline{\theta})]\right].\end{aligned} \tag{2.16}$$

The first order conditions are

$$\begin{array}{lcccl} x(\bar{m}, \emptyset, \bar{\theta}) & : & 0 & = & \bar{\pi}U'(x(\bar{m}, \emptyset, \bar{\theta})) - \bar{\pi}\lambda + \mu\, U'(x(\bar{m}, \emptyset, \bar{\theta})) \\ x(\underline{m}, \emptyset, \underline{\theta}) & : & 0 & = & \underline{\pi}U'(x(\underline{m}, \emptyset, \underline{\theta})) - \underline{\pi}\lambda - \mu\, U'[x(\underline{m}, \emptyset, \underline{\theta}) + (\alpha + b)(\bar{\theta} - \underline{\theta})] \\ b & : & 0 & = & \lambda(\mathbb{E}(\theta) - \rho) - \mu(\bar{\theta} - \underline{\theta})U'[x(\underline{m}, \emptyset, \underline{\theta}) + (\alpha + b)(\bar{\theta} - \underline{\theta})]. \end{array}$$

Substituting the incentive compatibility constraint and utility function into the first order conditions yields

$$\frac{U'(x(\underline{m}, \emptyset, \underline{\theta}))}{U'(x(\bar{m}, \emptyset, \bar{\theta}))} = \frac{\bar{\pi}}{\underline{\pi}} \left(\frac{\bar{\theta} - \rho}{\rho - \underline{\theta}}\right). \tag{2.17}$$

The right hand side of equation 2.17 shows that the ratio of weighted returns in high and low states can be interpreted as the cost of consumption in low states relative to consumption in high states. Equation 2.17 along with the incentive compatibility constraint can be substituted into the participation constraint to solve first for $x(\bar{m}, \emptyset, \bar{\theta})$ and the remaining choice variables:

$$
\begin{aligned}
x(\bar{m},\emptyset,\bar{\theta}) &= \frac{\alpha\rho\,(\bar{\theta}-\underline{\theta})}{\left(\frac{\underline{\pi}(\rho-\underline{\theta})}{\bar{\pi}(\bar{\theta}-\rho)}\right)^{1/\gamma}(\bar{\theta}-\rho)+(\rho-\underline{\theta})} \\
x(\underline{m},\emptyset,\underline{\theta}) &= \frac{\alpha\rho\,(\bar{\theta}-\underline{\theta})}{(\bar{\theta}-\rho)+(\rho-\underline{\theta})\left(\frac{\bar{\pi}(\bar{\theta}-\rho)}{\underline{\pi}(\rho-\underline{\theta})}\right)^{1/\gamma}} \\
b &= \alpha\rho\left[\frac{1-\left(\frac{\underline{\pi}(\rho-\underline{\theta})}{\bar{\pi}(\bar{\theta}-\rho)}\right)^{1/\gamma}}{\left(\frac{\underline{\pi}(\rho-\underline{\theta})}{\bar{\pi}(\bar{\theta}-\rho)}\right)^{1/\gamma}(\bar{\theta}-\rho)+(\rho-\underline{\theta})}\right]-\alpha \\
\lambda &= \left[\frac{\underline{\pi}^{1/\gamma}(\bar{\theta}-\rho)^{\frac{\gamma-1}{\gamma}}+\bar{\pi}^{1/\gamma}(\rho-\underline{\theta})^{\frac{\gamma-1}{\gamma}}}{\alpha\rho\,(\bar{\theta}-\underline{\theta})^{\frac{\gamma-1}{\gamma}}}\right]^{\gamma} \\
\mu &= \frac{\bar{\pi}(\mathbb{E}(\theta)-\rho)}{\rho-\underline{\theta}}.
\end{aligned} \tag{2.18}
$$

2.B.2 ALWAYS AUDIT CONTRACTS $(q=1)$ WITH NO TYPE-II ERRORS $(\eta(\bar{\theta})=0)$

When audits occur with certainty following low type reports, and the audit signal correctly identifies high type entrepreneurs with certainty, then as in the private information case the incentive compatibility constraint becomes linear. Our problem can be expressed by the following Lagrangian

$$
\begin{aligned}
\mathcal{L}_1 = {} & \bar{\pi}\,U(x(\bar{m},\emptyset,\bar{\theta}))+\underline{\pi}(1-\eta(\underline{\theta}))\,U(x(\underline{m},\underline{\sigma},\underline{\theta}))+\underline{\pi}\eta(\underline{\theta})\,U(x(\underline{m},\bar{\sigma},\underline{\theta})) \\
& +\lambda[(\alpha+b)(\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa)+\alpha\rho-\bar{\pi}x(\bar{m},\emptyset,\bar{\theta})-\underline{\pi}(1-\eta(\underline{\theta}))x(\underline{m},\underline{\sigma},\underline{\theta})-\underline{\pi}\eta(\underline{\theta})x(\underline{m},\bar{\sigma},\underline{\theta})] \\
& +\mu\left[U(x(\bar{m},\emptyset,\bar{\theta}))-U[x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})]\right]
\end{aligned} \tag{2.19}
$$

The first order conditions are

$$
\begin{array}{llll}
x(\bar{m},\emptyset,\bar{\theta}) & : & 0 & = \bar{\pi}U'(x(\bar{m},\emptyset,\bar{\theta}))-\bar{\pi}\lambda+\mu\,U'(x(\bar{m},\emptyset,\bar{\theta})) \\
x(\underline{m},\underline{\sigma},\underline{\theta}) & : & 0 & = \underline{\pi}(1-\eta(\underline{\theta}))U'(x(\underline{m},\underline{\sigma},\underline{\theta}))-\underline{\pi}(1-\eta(\underline{\theta}))\lambda \\
x(\underline{m},\bar{\sigma},\underline{\theta}) & : & 0 & = \underline{\pi}\eta(\underline{\theta})U'(x(\underline{m},\bar{\sigma},\underline{\theta}))-\underline{\pi}\eta(\underline{\theta})\lambda-\mu\,U'[x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})] \\
b & : & 0 & = \lambda(\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa)-\mu(\bar{\theta}-\underline{\theta})U'[x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})].
\end{array}
$$

Substituting the incentive compatibility constraint into the first order conditions yields

$$
\begin{aligned}
U'(x(\bar{m},\emptyset,\bar{\theta})) &= \lambda\left(\frac{\bar{\pi}(\bar{\theta}-\underline{\theta})-(\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa)}{\bar{\pi}(\bar{\theta}-\underline{\theta})}\right) \\
U'(x(\underline{m},\underline{\sigma},\underline{\theta})) &= \lambda \\
U'(x(\underline{m},\bar{\sigma},\underline{\theta})) &= \lambda\left(\frac{\underline{\pi}\eta(\underline{\theta})(\bar{\theta}-\underline{\theta})+\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa}{\underline{\pi}\eta(\underline{\theta})(\bar{\theta}-\underline{\theta})}\right),
\end{aligned}
$$

which we can combine with the intermediary's participation constraint and the utility function and solve for consumption allocations:

$$\begin{aligned}
x(\bar{m},\emptyset,\bar{\theta}) &= \alpha\rho\chi\left(\frac{1}{1-\zeta}\right)^{1/\gamma} \\
x(\underline{m},\underline{\sigma},\underline{\theta}) &= \alpha\rho\chi \\
x(\underline{m},\bar{\sigma},\underline{\theta}) &= \alpha\rho\chi\left(\frac{\underline{\pi}\eta(\underline{\theta})}{\bar{\pi}\zeta+\underline{\pi}\eta(\underline{\theta})}\right)^{1/\gamma} \\
b &= \frac{\alpha\rho\chi}{\bar{\theta}-\underline{\theta}}\left(\frac{\zeta}{1-\zeta}\right)^{1/\gamma}\left(\frac{\bar{\pi}+\underline{\pi}\eta(\underline{\theta})}{\bar{\pi}\zeta+\underline{\pi}\eta(\underline{\theta})}\right)^{1/\gamma}-\alpha \\
\lambda &= (\alpha\rho\chi)^{-\gamma} \\
\mu &= \frac{\bar{\pi}\zeta}{1-\zeta}.
\end{aligned} \tag{2.20}$$

where

$$\chi = \frac{1}{\bar{\pi}\,(1-\zeta)^{\frac{\gamma-1}{\gamma}}+\underline{\pi}\eta(\underline{\theta})\left(\frac{\bar{\pi}\zeta+\underline{\pi}\eta(\underline{\theta})}{\underline{\pi}\eta(\underline{\theta})}\right)^{\frac{\gamma-1}{\gamma}}+\underline{\pi}(1-\eta(\underline{\theta}))}, \qquad \text{and}$$

$$\zeta = \frac{\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa}{\bar{\pi}(\bar{\theta}-\underline{\theta})}.$$

2.B.3 Proof of Proposition 2.5

Proof. We can write the first order condition for q as follows:

$$\begin{aligned}
\mathcal{L}_q &= (1-\eta(\underline{\theta}))U(x(\underline{m},\underline{\sigma},\underline{\theta}))+\eta(\underline{\theta})U(x(\underline{m},\bar{\sigma},\underline{\theta}))-U(x(\underline{m},\emptyset,\underline{\theta})) \\
&\quad -\lambda[(\alpha+b)\pi\kappa+(1-\eta(\underline{\theta}))x(\underline{m},\underline{\sigma},\underline{\theta})+\eta(\underline{\theta})x(\underline{m},\bar{\sigma},\underline{\theta})-x(\underline{m},\emptyset,\underline{\theta})] \\
&\quad +\mu[U(x(\underline{m},\emptyset,\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta}))-U(x(\underline{m},\bar{\sigma},\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta}))]+\nu_0-\nu_1.
\end{aligned} \tag{2.21}$$

In order to solve for Kuhn-Tucker multiplier ν_1, we need to solve for shadow allocations which are consumed with probability zero: $x(\underline{m},\emptyset,\underline{\theta})$ in the case of standard debt.

In the limit as $q\to 1^-$, the first order condition for $x(\underline{m},\emptyset,\underline{\theta})$ must hold, even though this allocation is consumed with vanishing probability. The first order condition for $x(\underline{m},\emptyset,\underline{\theta})$ is:

$$x(\underline{m},\emptyset,\underline{\theta}): \quad 0=\underline{\pi}U'(x(\underline{m},\emptyset,\underline{\theta}))-\underline{\pi}\lambda-\mu\,U'[x(\underline{m},\emptyset,\underline{\theta})+(\alpha+b)(\bar{\theta}-\underline{\theta})].$$

Under logarithmic utility, $U'(x)=1/x$. We can solve for $x(\underline{m},\emptyset,\underline{\theta})$ using the solutions obtained in (2.20):

$$x(\underline{m},\emptyset,\underline{\theta}) = -\frac{\alpha\rho}{2}\left[\frac{1}{\underline{\pi}}\frac{\zeta}{1-\zeta} - \frac{\overline{\pi}\eta(\underline{\theta})}{\overline{\pi}\zeta+\underline{\pi}\eta(\underline{\theta})}\right] + \frac{\alpha\rho}{2}\sqrt{\left[\frac{1}{\underline{\pi}}\frac{\zeta}{1-\zeta} - \frac{\overline{\pi}\eta(\underline{\theta})}{\overline{\pi}\zeta+\underline{\pi}\eta(\underline{\theta})}\right]^2 + 4\frac{\zeta}{1-\zeta}\left(\frac{\overline{\pi}+\underline{\pi}\eta(\underline{\theta})}{\overline{\pi}\zeta+\underline{\pi}\eta(\underline{\theta})}\right)}. \tag{2.22}$$

Noting that $\nu_0 = 0$ by complimentary slackness, substitute (2.22) and (2.20) into (2.21) to express the Kuhn-Tucker multiplier ν_1 as a closed-form expression in terms of parameters.

■

2.C Proof of Proposition 2.6

Proof. Consider the case where the entrepreneur enjoys consumption with log utility, type–II errors occur with zero probability ($\eta(\bar{\theta}) = 0$) and type-I errors occur with positive but very low probability, $\eta(\underline{\theta}) \to 0^+$. By Proposition 2.4, consumption following overturned low type reports also tends toward zero ($\lim_{\eta(\underline{\theta})\to 0^+} x(\underline{m},\bar{\sigma},\underline{\theta}) = 0$), and by l'Hôpital's rule, the contribution to ex ante expected utility of the entrepreneur from consumption following overturned reports also tends toward zero, ($\lim_{\eta(\underline{\theta})\to 0^+} \underline{\pi}\eta(\underline{\theta})U(x(\underline{m},\bar{\sigma},\underline{\theta})) = 0$). When type-I errors are extremely rare, errors have little effect on the ex ante welfare of entrepreneurs but still limit repayments drawn from high type entrepreneurs.

For tractability, we consider contracts under the following assumptions: First, type-I errors will occur with very low (positive) probability. Second, utility will be logarithmic over consumption. Third, the two income states will occur with equal probability.

Let y_1 (y_0) be the efficient value of choice variable y in the always audit (private information) contract. Substituting $U(x) = \log x$ and $\underline{\pi} = \bar{\pi} = 1/2$ into the solutions from Appendix 2.B, and taking the limit as $\eta(\underline{\theta}) \to 0^+$, we obtain the following solutions:

$$\frac{\alpha+b_1}{\alpha} = \frac{1}{2}\frac{\rho}{\rho-\underline{\theta}+\underline{\pi}\kappa}, \qquad x_1(\bar{m},\emptyset,\bar{\theta}) = \alpha\rho\frac{1}{2}\frac{\bar{\theta}-\underline{\theta}}{\rho-\underline{\theta}+\underline{\pi}\kappa}, \qquad x_1(\underline{m},\underline{\sigma},\underline{\theta}) = \alpha\rho$$

$$x_1(\underline{m},\emptyset,\underline{\theta}) = \alpha\rho\frac{-(\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa)+\sqrt{(\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa)^2+\frac{1}{2}(\bar{\theta}-\underline{\theta})(\rho-\underline{\theta}+\underline{\pi}\kappa)}}{\rho-\underline{\theta}+\underline{\pi}\kappa}$$

$$x_1(\underline{m},\bar{\sigma},\underline{\theta}) = 0, \qquad \lambda_1 = \frac{1}{\alpha\rho}, \qquad \mu_1 = \frac{1}{2}\frac{\mathbb{E}(\theta)-\rho-\underline{\pi}\kappa}{\rho-\underline{\theta}+\underline{\pi}\kappa}. \tag{2.23}$$

$$\frac{\alpha+b_0}{\alpha} = \frac{\rho(\mathbb{E}(\theta)-\rho)}{(\bar{\theta}-\rho)(\rho-\underline{\theta})}, \qquad x_0(\bar{m};\emptyset;\bar{\theta}) = \alpha\rho\frac{1}{2}\frac{\bar{\theta}-\underline{\theta}}{\rho-\underline{\theta}}, \qquad x_0(\underline{m},\underline{\sigma},\underline{\theta}) = \alpha\rho$$

$$x_0(\underline{m},\emptyset,\underline{\theta}) = \alpha\rho\,\frac{1}{2}\frac{\bar{\theta}-\theta}{\bar{\theta}-\rho}, \qquad x_0(\underline{m},\bar{\sigma},\underline{\theta}) = 0, \qquad \lambda_0 = \frac{1}{\alpha\rho}, \qquad \mu_0 = \frac{1}{2}\frac{\mathbb{E}(\theta)-\rho}{\rho-\underline{\theta}}. \tag{2.24}$$

The bang-bang result occurs when efficient contracts 'jump' between private information and standard debt contracts, where both are local maxima. To focus on these bang-bang results, we first set κ such that the always audit and private information contracts provide equal expected utility ($\mathbb{E}U(x_1) = \mathbb{E}U(x_0)$). Solving for κ yields

$$\underline{\pi}\kappa = 2\,\frac{(\rho-\underline{\theta})(\mathbb{E}(\theta)-\rho)}{\bar{\theta}-\underline{\theta}} \tag{2.25}$$

It is useful to define two new parameters, one representing the excess return in good states and the other the shortfall in bad states. Thus, let $\bar{\phi},\underline{\phi}\in\mathbb{R}^+$, where $\bar{\phi}=\bar{\theta}-\rho$, and $\underline{\phi}=\rho-\underline{\theta}$. All else equal, the entrepreneur would prefer a project with large $\bar{\phi}$, and small $\underline{\phi}$. Note that assumptions 2.1 and 2.2 require that $\underline{\phi}\in(0,\bar{\phi})$. Substituting equation 2.25 into the solutions for the always standard debt contract (2.23), we can re-write allocations as follows:

$$b_1 = \frac{\alpha\rho}{4} : \frac{\bar{\phi}+\underline{\phi}}{\bar{\phi}\underline{\phi}} - \alpha, \qquad x_1(\bar{m},\emptyset,\bar{\theta}) = \frac{\alpha\rho}{4}\frac{(\bar{\phi}+\underline{\phi})^2}{\bar{\phi}\underline{\phi}},$$

$$x_1(\underline{m},\emptyset,\underline{\theta}) = \frac{\alpha\rho}{4}\left[\frac{-(\bar{\phi}-\underline{\phi})^2+\sqrt{(\bar{\phi}-\underline{\phi})^4+4\bar{\phi}\underline{\phi}(\bar{\phi}+\underline{\phi})^2}}{\bar{\phi}\underline{\phi}}\right], \qquad \mu_1 = \frac{1}{8}\,\frac{(\bar{\phi}-\underline{\phi})^2}{\bar{\phi}\underline{\phi}} \tag{2.26}$$

Now, consider the trade-off characterised by the first order condition for auditing q at the always audit contract. After substituting equations 2.25 and 2.26 into the first order condition for q (equation 2.21), we obtain

$$\begin{aligned}\mathcal{L}_q = {}& \frac{1}{2}\log\left[\frac{4\bar{\phi}\underline{\phi}}{-(\bar{\phi}-\underline{\phi})^2+\sqrt{(\bar{\phi}-\underline{\phi})^4+4\bar{\phi}\underline{\phi}(\bar{\phi}+\underline{\phi})^2}}\right] \\ & -\frac{1}{2}\left[1-\frac{-(\bar{\phi}-\underline{\phi})^2+\sqrt{(\bar{\phi}-\underline{\phi})^4+4\bar{\phi}\underline{\phi}(\bar{\phi}+\underline{\phi})^2}}{4\bar{\phi}\underline{\phi}}\right] - \frac{1}{4}\,\frac{(\bar{\phi}-\underline{\phi})}{\bar{\phi}} \\ & +\frac{1}{8}\,\frac{(\bar{\phi}-\underline{\phi})^2}{\bar{\phi}\underline{\phi}}\log\left(1+\frac{-(\bar{\phi}-\underline{\phi})^2+\sqrt{(\bar{\phi}-\underline{\phi})^4+4\bar{\phi}\underline{\phi}(\bar{\phi}+\underline{\phi})^2}}{(\bar{\phi}+\underline{\phi})^2}\right) - \nu_1.\end{aligned} \tag{2.27}$$

The first term on the right hand side of equation 2.27 is the welfare gain attained through auditing by verifying low type agents' reports. In this example, agents clearly prefer to be audited, $x(\underline{m},\underline{\sigma},\underline{\theta}) > x(\underline{m},\emptyset,\underline{\theta})$ for all values of $\bar{\phi},\underline{\phi}$. The second term captures the resource cost of this increase in consumption for low type agents whose reports are verified, and the third term represents the extra resource costs expended by the intermediary in conducting more audits. The fourth term captures welfare gains attained by relaxing the incentive compatibility constraint: auditing with a higher probability directly increases the likelihood that misreporting high type entrepreneurs will be punished. The final term on the right hand side is the Lagrange multiplier capturing the shadow cost of the natural upper bound of one attached to the audit probability.

Let $\underline{\phi} \to 0^+$. By l'Hôpital's rule,

$$\lim_{\underline{\phi}\to 0^+} \frac{-(\bar{\phi}-\underline{\phi})^2 + \sqrt{(\bar{\phi}-\underline{\phi})^4 + 4\bar{\phi}\underline{\phi}(\bar{\phi}+\underline{\phi})^2}}{4\bar{\phi}\underline{\phi}} = \frac{1}{2}, \qquad \text{and}$$

$$\lim_{\underline{\phi}\to 0^+} \frac{1}{\underline{\phi}} \log\left(1 + \frac{-(\bar{\phi}-\underline{\phi})^2 + \sqrt{(\bar{\phi}-\underline{\phi})^4 + 4\bar{\phi}\underline{\phi}(\bar{\phi}+\underline{\phi})^2}}{(\bar{\phi}+\underline{\phi})^2}\right) = \frac{2}{\bar{\phi}}.$$

Substituting these results into 2.27 while retaining the same ordering of terms yields

$$\lim_{\underline{\phi}\to 0^+} \mathcal{L}_q(q=1) = \frac{1}{2}\log 2 - \frac{1}{4} - \frac{1}{4} + \frac{1}{4} - \nu_1. \tag{2.28}$$

The Kuhn-Tucker multiplier ν_1 is positive. At the margin, the benefits of additional audits outweigh the costs. The first two terms show that the utility benefits accruing to low type entrepreneurs from verification of their reports exceeds the resource cost associated with awarding more low type entrepreneurs with the post-verification consumption bundle. The resource cost of audits is the product of the Lagrange multiplier on the resource constraint, total assets devoted to the project and audit costs. Here, as the downside shortfall $\underline{\phi}$ approaches zero, borrowing and assets devoted to the project are unbounded above. When downside risk is low, the benefits of auditing are small, and indeed the audit cost which equates the expected welfare of always audit and private information contracts is vanishing $\underline{\pi}\kappa \to 0^+$. The resource cost of the marginal audit is 1/4, which in this case is equal and opposite to the benefit attained from the marginal audit by relaxing the incentive compatibility constraint.

For the same case, letting $\underline{\phi} \to 0^+$, now consider the corresponding private information contract. The first order condition for q can be written as follows:

$$\mathcal{L}_q = \frac{1}{2}\log\frac{2\bar{\phi}}{\bar{\phi}+\underline{\phi}} - \frac{1}{2}\left[1 - \frac{1}{2}\frac{\bar{\phi}+\underline{\phi}}{\bar{\phi}}\right] - \frac{1}{2}\frac{(\bar{\phi}-\underline{\phi})^2}{\bar{\phi}(\bar{\phi}+\underline{\phi})} + \frac{1}{4}\frac{\bar{\phi}-\underline{\phi}}{\underline{\phi}}\log\left(\frac{\underline{\phi}}{\bar{\phi}-\underline{\phi}} + 1\right) + \nu_0.$$

Taking the limit as $\underline{\phi} \to 0^+$ yields

$$\lim_{\underline{\phi}\to 0^+} \mathcal{L}_q(q=0) = \frac{1}{2}\log 2 - \frac{1}{4} - \frac{1}{2} + \frac{1}{4} + \nu_0, \tag{2.29}$$

The first term captures the direct welfare benefit from verifying entrepreneur reports, and providing them with the consumption bundle $x_0(\underline{m}, \underline{\sigma}, \underline{\theta}) > x_0(\underline{m}, \emptyset, \underline{\theta})$. This benefit, and the resource cost associated with it and captured in the second term, are identical to those in the always audit contract (2.28). This is due to the fact that these consumption bundles are identical in both contracts for this limiting case: low type agents whose reports are verified ($x(\underline{m}, \underline{\sigma}, \underline{\theta})$) consume $\alpha\rho$ and low type agents whose reports are unverified ($x(\underline{m}, \emptyset, \underline{\theta})$) consume $\alpha\rho/2$ in both contracts. As in the always audit contract (2.28), the fourth term capturing the relaxation in the incentive compatibility constraint is equal to $1/4$.

The third term, capturing the cost of the marginal audit, is greater in magnitude than under the always audit contract. Audit costs as a fraction of assets employed in the project are constant across contracts by assumption, yet leverage in the private information contract is greater than in the always audit contract. We can see this by taking the limit of the ratio of assets devoted to the project in the two contracts: $\lim_{\underline{\phi}\to 0} \dfrac{\alpha + b_0}{\alpha + b_1} = 2$.

■

2.D STANDARD DEBT CONTRACTS WITH TYPE-II ERRORS

Assume that the probability of Type-II errors $\eta(\bar{\theta})$ following the audit of a high type entrepreneur is low, and that the optimal contract is standard debt ($q^* = 1$). We can find approximate closed form solutions to optimal contracts using a first order Taylor expansion of the Incentive Compatibility Constraint around $\eta(\bar{\theta}) = 0$. Denote the efficient contract at $\eta(\bar{\theta}) = 0$ with all other parameters constant by Γ_0, with associated public actions and allocations labelled $b_0, q_0 = 1, x_0, z_0$.

$$
\begin{aligned}
U(x_0(\bar{m}, \emptyset, \bar{\theta})) \geq\ & U(x_0(\underline{m}, \bar{\sigma}, \bar{\theta})) + U'(x_0(\underline{m}, \bar{\sigma}, \bar{\theta}))[x(\underline{m}, \bar{\sigma}, \bar{\theta}) - x_0(\underline{m}, \bar{\sigma}, \bar{\theta})] \\
& - U'(x_0(\bar{m}, \emptyset, \bar{\theta}))[x(\bar{m}, \emptyset, \bar{\theta}) - x_0(\bar{m}, \emptyset, \bar{\theta})] \\
& + \eta(\bar{\theta})[U(x_0(\underline{m}, \underline{\sigma}, \bar{\theta})) - U(x_0(\underline{m}, \bar{\sigma}, \bar{\theta}))]
\end{aligned}
$$

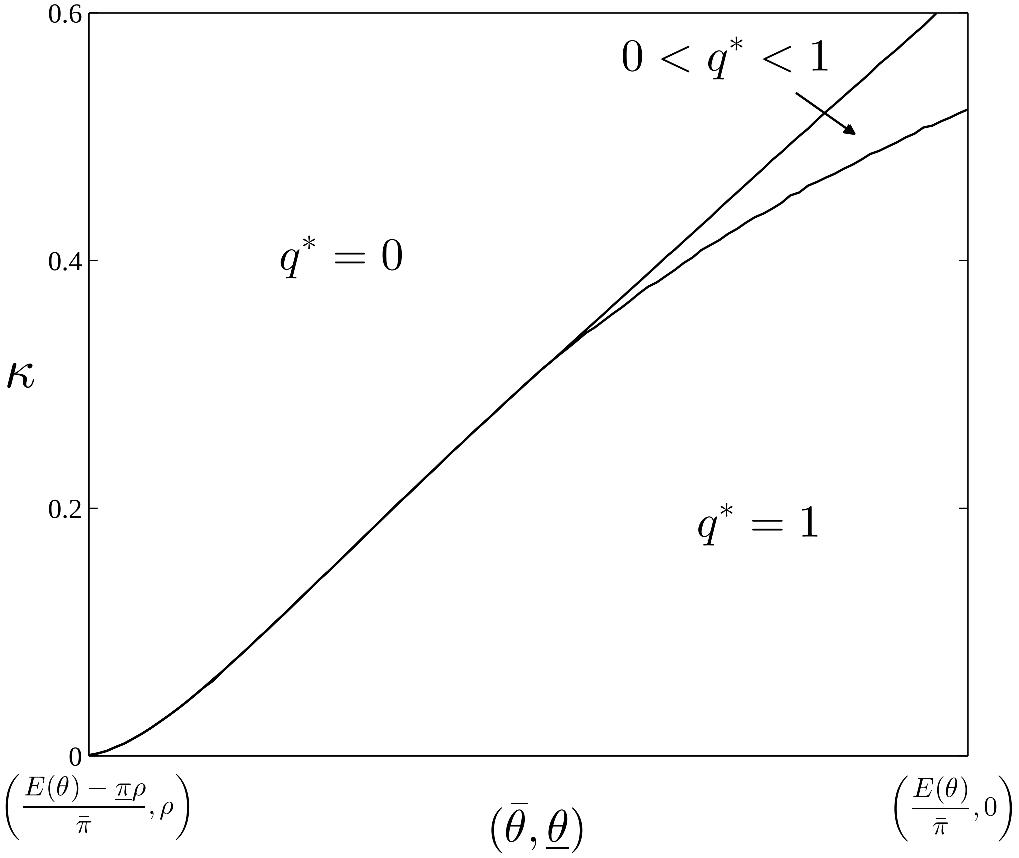
Note that $x_0(\bar{m}, \emptyset, \bar{\theta}) = x_0(\underline{m}, \bar{\sigma}, \bar{\theta})$, which combined with the budget constraints allows us to simplify the above expression as follows:

$$
z(\underline{m}, \bar{\sigma}) - z(\bar{m}, \emptyset) \geq \frac{\eta(\bar{\theta})}{U'(x_0(\underline{m}, \bar{\sigma}, \bar{\theta}))}[U(x_0(\underline{m}, \underline{\sigma}, \bar{\theta})) - U(x_0(\underline{m}, \bar{\sigma}, \bar{\theta}))]. \tag{2.30}
$$

When utility is log, we can solve directly using Proposition 2.4. After rearranging and simplifying, our first order approximation of the ICC can be written as a linear expression in terms of contracted repayments:

$$
z(\underline{m}, \bar{\sigma}) - z(\bar{m}, \emptyset) \geq \eta(\bar{\theta})\ \frac{\alpha\rho}{1-\zeta} \log\left[1 + \frac{\bar{\pi}\zeta(1-\bar{\pi}\zeta)}{\bar{\pi}\zeta + \underline{\pi}\eta(\underline{\theta})}\right] \tag{2.31}
$$

Figure 2.1: Efficient contracts, project risk and audit costs. $(U(x) = \sqrt{x}, \rho = 1, \mathbb{E}(\theta) = 1.2, \bar{\pi} = 0.9, \eta(\underline{\theta}) = 0.01)$.



This linear approximation to the ICC permits closed form approximations to allocations and leverage for efficient contracts, which are not presented here.

One interpretation of equation 2.31 is that it specifies a a small non-refundable fee paid by all entrepreneurs who declare a low type return. Entrepreneurs whose reports are overturned by the audit signal would be required to repay the full contracted repayment $z(\bar{m}, \emptyset)$ in addition to the small fee. When $\eta(\bar{\theta})$ is small, this fee is negligible.

2.E Figures

Figure 2.2: The determination of optimal contracts when there are multiple local maxima. $(U(x) = \log x, \rho = 1, \mathbb{E}(\theta) = 1.2, \bar{\pi} = 0.9, (\bar{\theta} - \underline{\theta}) = 0.3, \eta(\underline{\theta}) = 10^{-4}, \alpha = 1, \kappa \approx 0.18)$

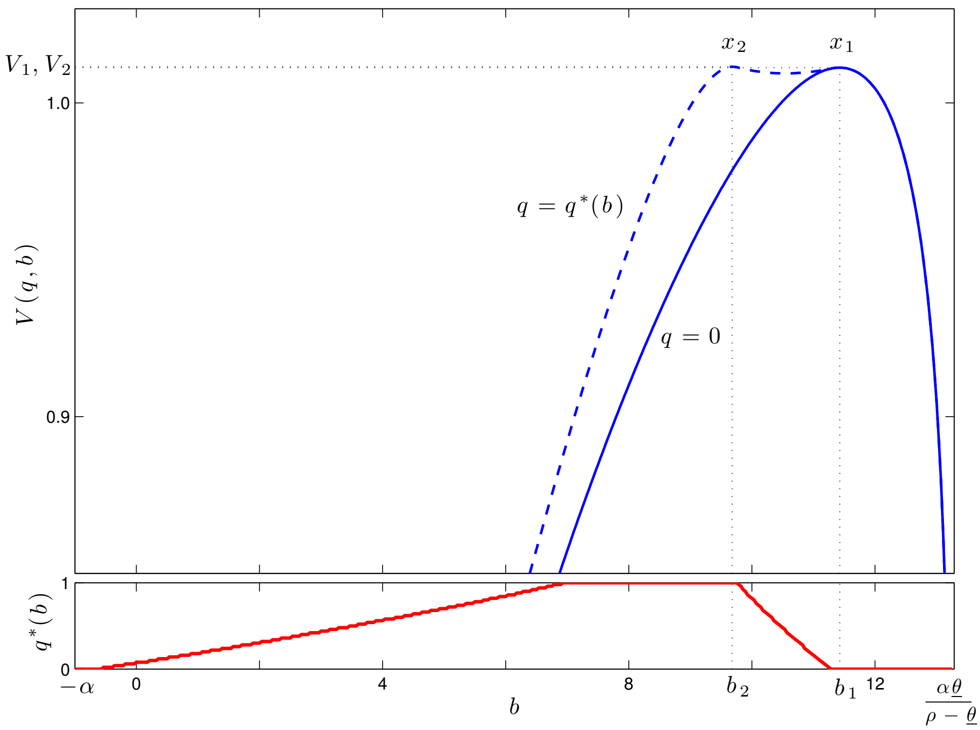
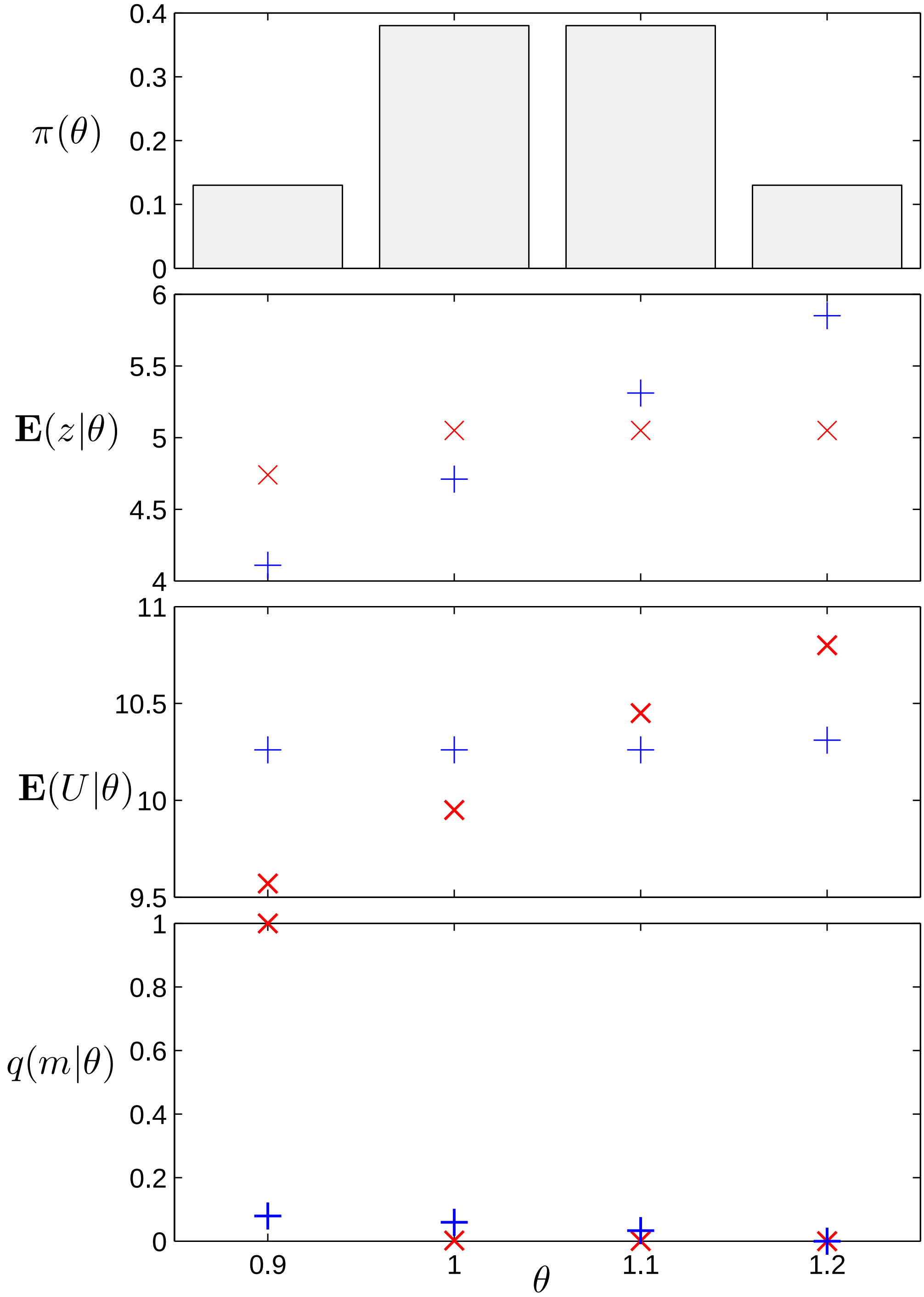


Figure 2.3: A Four-state example. ($\Theta = (0.9, 1.0, 1.1, 1.2), \pi(\Theta) = (1/8, 3/8, 3/8, 1/8), \rho = 1, \alpha = 1, b = 5, \kappa = 0.08, U(x) = x^{1-\gamma}/(1-\gamma), \gamma = 9/10$). Imperfect audits case marked by $\times$, with $P(\sigma = \theta_i|\theta = \theta_j) = 10^{-|i-j|}$ if and only if $i \neq j$. Perfect audits case marked by $+$, with $P(\sigma = \theta_i|\theta = \theta_j) = 0$ if and only if $i \neq j$.



CHAPTER 3

FINANCIAL FRICTIONS AND UNEMPLOYMENT

This chapter is co-authored with Charles Nolan.[1]

This paper presents a DSGE model with risk averse entrepreneurs who must be compensated for bearing productive risk. Increases in leverage or in financial stress increase the risk burden placed on entrepreneurs, and reduce wages and interest rates below their factors' respective marginal products. This increases the volatility of employment, and distorts the link between labour cost growth and inflation pressure. Inequalities between worker households and entrepreneurs amplify and propagate business cycle shocks. Furthermore, the model describes a link between observed macroeconomic trends in the labour share of income and the capital-output ratio and the observed decline in entrepreneurship over recent decades. The model suggest that this decline in entrepreneurship is consistent with trend increases in leverage as well as an increase in the sensitivity of employment and output to financial shocks.

[1]Professor of Economics, University of Glasgow. `Charles.Nolan@glasgow.ac.uk`. The authors would like to thank Christoph Thoenissen and conference and seminar participants at the University of Glasgow Macrotheory Seminar Series and the 2015 Sheffield Workshop in Macroeconomics for helpful comments and discussions. All errors are our own.

INTRODUCTION

This paper introduces a new microfoundation for debt contracts as project finance within a DSGE framework. The key modelling departure from the financial accelerator model of Bernanke, Gertler, and Gilchrist (1999) is the introduction of debt disputes and risk averse entrepreneurs, which combined result in the optimality of standard debt contracts. The main consequence of the introduction of entrepreneur risk aversion is to elicit a precautionary motive which limits the demand for capital and labour factor inputs below their respective marginal revenue products. During periods of financial stress, this precautionary motive increases, reducing the demand for labour further below what would be predicted by shifts in marginal labour productivity. This exacerbates volatility in employment, and amplifies shifts in unemployment resulting from imperfect labour market institutions.

RELATION TO THE LITERATURE

The well studied models described by Bernanke, Gertler, and Gilchrist (1999) and Carlstrom and Fuerst (1997) motivate the use of debt contracts to finance firms' investment projects by appealing to the costly state verification framework. In the costly state verification model, project outcomes are initially privately observed by the borrower alone, but can be detected by the lender using a costly auditing technology. Townsend (1979) and Gale and Hellwig (1985) showed that when stochastic audit regimes are not available, a standard debt contract is an optimal external finance contract in costly state verification settings. Risk sharing and auditing only occurs following bad outcomes, which are interpreted as bankruptcy. Border and Sobel (1987) and Mookherjee and Png (1989) showed with risk neutral and risk averse borrowers respectively that limiting agents to pure audit strategies in this setting was extremely restrictive. Much if not all of the deadweight loss associated with the information asymmetry could be eliminated if the lender can employ stochastic monitoring strategies. With risk averse borrowers, resulting optimal contracts resemble equity, with a high degree of risk sharing and stochastic auditing. Duncan and Nolan (2014) show that the standard debt can be restored as an optimal contract in costly state verification problems if the audit technology provides an imperfect signal and the borrower is risk averse, and this is the model we employ here.

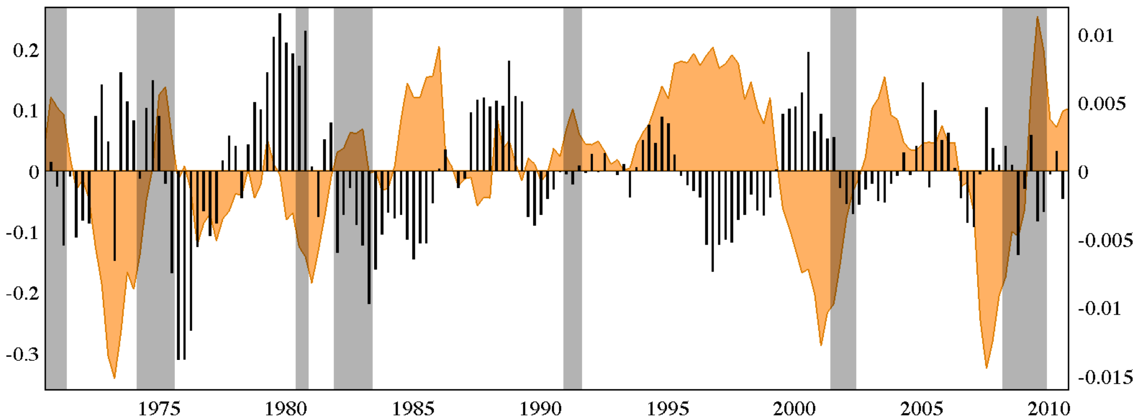
The model we present is Knightian in the sense that entrepreneurs are risk averse and must be compensated for bearing risk even under perfect competition in factor and product markets (see Knight, 1921, part 3, chapter 9). This risk cannot be passed on in full to outside lenders and investors who cannot freely observe the luck enjoyed by firms engaging in risky projects. This luck is private information observed initially only by the entrepreneur. As a result, fixed wages paid to employees and interest paid to creditors will fall short of their respective expected marginal products, the difference compensating the entrepreneur for providing guaranteed fixed wages and partially flexible interest income when revenues are subject to risk.[2]

In terms of the dynamics of the model, it is the wedge between wages and the marginal productivity of labour which differentiates us from Bernanke, Gertler, and Gilchrist (1999) and Kiyotaki and Moore (1997). In Bernanke, Gertler, and Gilchrist (1999), expected bankruptcy costs are increasing in amount of capital employed by the firm, driving a wedge between risk free interest rates and capital's marginal product. In Kiyotaki and Moore (1997), loans funding capital must be backed by sufficient collateral to ensure repayment. This pulls the interest rate below the marginal product of capital, such that total loan repayments are within the stock of pledgeable collateral assets under control. In our model, audit costs increase in the amount of capital employed by the firm as in Bernanke, Gertler, and Gilchrist (1999), but in contrast with that model, our entrepreneurs are risk averse, which limits the extent to which they are willing to provide guaranteed wage and interest income. The wedge between interest rates earned on the representative household's savings and the marginal product of capital is a combination of monitoring costs as in Bernanke, Gertler, and Gilchrist (1999) and a risk premium. The wedge between the marginal product of labour and the wage rate is a risk premium, and the risk premium components of the two factor market wedges move together.

This has important implications for the dynamics of employment. As leverage increases, or following an adverse shock to the efficiency of loan intermediation, the risk borne by entrepreneurs increases, their required compensation for risk increases and both wages and interest rates fall further from their respective marginal products of labour and capital. Chari, Kehoe, and McGrattan (2007), Jermann and Quadrini (2012) and Zanetti

[2]While Knight (1921) also emphasised the role of uncertainty (unquantifiable risks) in limiting risk sharing between entrepreneurs and outside investors/insurers, his theory of entrepreneur compensation and factor price determination relies on the presence of either uninsurable risks or uncertainty. In this paper risks are uninsurable as a result of information asymmetries alone.

Figure 3.1: Equity risk premium over following 8 quarters (annualised, area, LHS) and the labour share of income (deviation from HP-filtered trend, bars, RHS). Shaded areas indicate NBER recessions.



(2015) show empirically that large recessions following financial crises tend to exhibit a large wedge between wages and the marginal product of labour, and our model predicts that financial stress and the labour wedge are tightly linked.

Figure 3.1 presents the forward looking 2-year excess return to equity of a portfolio of 90 day treasury bills, and the deviation of the labour share of national income from Hodrick-Prescott filtered trend. Using the realised return to equity and bond portfolios as a very imperfect proxy for the expected return, the chart shows a clear negative correlation between the equity risk premium and the labour wedge. When wages are relatively low, the premium returned to equityholders over bondholders tends to be high, and vice-versa. This is a key prediction of our model. Firms' willingness to offer partially-guaranteed income to capital through fixed coupon bonds is tightly linked to their willingness to offer guaranteed wage income.[3]

This chapter proceeds to derive the important relationships in our model, starting from the entrepreneurs' problem. Features of the model which are not unique to this paper, for example the representative household, wage and retail price setting and capital goods production are described in Appendix 3.B. The dynamics and key results of this chapter are contained in Section 3.3. In Section 3.4, we consider two extensions of the benchmark model.

First, we show that when the model is subject to New Keynesian retail price setting frictions, the marginal cost of production is equal to the sum of marginal labour costs and

[3] Abo-Zaid (2013) describes a similar correlation between credit spreads and the labour wedge.

the risk premium accruing to entrepreneurs. During periods of financial stress, the risk premium is large, and cost pressure on retail prices exceeds what would be predicted by labour cost growth alone.

Second, we outline the three main candidate financial shocks, which consist of innovations in project risk, entrepreneur risk aversion and the quality of loan monitoring. We show analytically that when the probability of default is low, the effect of project risk shocks relative to risk aversion or loan monitoring shocks is greater.

3.1 THE MODEL

The structure of the economy is similar to the model described by Bernanke, Gertler, and Gilchrist (1999). The economy consists of a representative household and a unit measure of entrepreneurs. Entrepreneurs hire labour and borrow capital from the representative household to combine with their own capital in productive projects.

3.1.1 ENTREPRENEURS

There exists a unit measure of entrepreneurs, indexed by i, who enjoy consumption with logarithmic utility over the composite consumption good,

$$U^e_{ti} = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^{ej} u^e(C^e_{t+ji}), \tag{3.1}$$

where $u^e(C) = \log C$, and $\beta^e < \beta^h$, entrepreneurs are less patient than the representative household. [4] Entrepreneurs undertake projects with binary risky outcomes. The individual output of entrepreneur i exhibits constant returns to scale, and can be expressed as follows:[5]

$$Y_{ti} = \theta_{ti} A_t K^{\alpha}_{t-1i} N^{1-\alpha}_{ti}, \tag{3.2}$$

[4]Entrepreneurs enjoy a greater return on savings than households in the model. Their reduced discount factor is required to ensure that the ratio of entrepreneurial and household wealth is constant in the steady state.

[5]It is important for aggregation that the entrepreneurs' preferences are intertemporally homothetic ($u^e_{ti}(c^e_{ti}) = c^{e(1-\gamma^e)}_{ti}/(1-\gamma^e)$) and that their production technology exhibits constant returns to scale ($Y_{ti} = \theta_{ti} A_t[\alpha K^{\phi}_{ti} + (1-\alpha)N^{\phi}_{ti}]^{1/\phi}$). This means that each entrepreneurs' demand for factors and consumption is proportional to their wealth. Restricting preferences to log utility and technology to Cobb-Douglas is not essential, and Appendix 3.4.2 explores the generalisation of preferences to CRRA utility. For both the steady state and the dynamics, variations in entrepreneur risk aversion resemble variations in project risk.

where θ_{it} is an idiosyncratic shock drawn from $\theta_{ti} \in \{\theta_{1t}, \theta_{2t}\}$ where $\theta_{1t} < \theta_{2t}$. These two states occur with probabilities π_1 and π_2 respectively, and their expectation is equal to one, $\pi_1\theta_{1t} + \pi_2\theta_{2t} = 1$. Throughout this paper, it will be useful to consider the difference between high and low project outcomes $\xi_t = \theta_{2t} - \theta_{1t}$ which we will allow to be time varying. Denote the expectation of output for entrepreneur i in period t conditional upon A_t by $\bar{Y}_{ti} = A_t K_{t-1i}^{\alpha} N_{ti}^{1-\alpha}$. The variable A_t is an aggregate total factor productivity shock. Aggregate shocks are observable at the beginning of the period, the idiosyncratic shock is revealed to the entrepreneur at the end of the period. Capital employed by the entrepreneur is denoted in period t is K_{t-1i}, and N_{ti} is labour hired by the entrepreneur from the household sector.

At the beginning of each period, entrepreneurs borrow capital K_{ti}^{b} from financial intermediaries. Loan contracts specify the interest rate paid in good states, as well as the recovery rate returned to financial intermediaries in bad states. Capital inputs into entrepreneur i's project include the entrepreneur's initial capital holdings and further capital borrowed.

$$K_{ti} = K_{ti}^{e} + K_{ti}^{b}, \tag{3.3}$$

where K_t^e is the capital held by the entrepreneur at the beginning of the period. Entrepreneurs fund consumption and future capital holdings out of the sum of project revenues and current capital holdings, after repaying loans and paying workers' wages,

$$Q_t K_{ti}^{e} + C_{ti}^{e} = \frac{Y_{ti}(\theta_{ti})}{X_t} + Q_t(1-\delta)K_{t-1i}^{e} - K_{t-1i}^{b} r_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti}. \tag{3.4}$$

where δ is the depreciation rate of capital, Q_t is the real price of capital, X_t is the markup of retail prices over the wholesale price of entrepreneurs' project output, meaning that $1/X_t$ is the real price of the entrepreneurs' project output good. The state contingent real rental rate of capital is $r_{ti}(\theta_{ti}, \sigma_{ti})$, and W_t is the real wage rate. We attach the Lagrange multipliers $\lambda_{ti}^{e}(\theta_{ti}, \sigma_{ti})$ to each of the state contingent accumulation constraints. Note that capital rental payments $r_{ti}(\theta_{ti}, \sigma_{ti})$ are contingent on the idiosyncratic shock θ_{ti} as well as any audit signal obtained by the financial intermediary, $\sigma_{ti} \in \{\sigma_1, \sigma_2\}$. Audit signals are distributed as follows: $\mathrm{P}(\sigma_2|\theta_2) = 1, \mathrm{P}(\sigma_2|\theta_1) = \eta, \mathrm{P}(\sigma_1|\theta_1) = 1-\eta$. The unconditional probabilities of the three possible outcomes are as follows: $\mathrm{P}(\theta_1, \sigma_1) = \pi_1(1-\eta), \mathrm{P}(\theta_1, \sigma_2) = \pi_1\eta,$ and $\mathrm{P}(\theta_2, \emptyset) = \pi_2$.

3.1.2 FINANCIAL CONTRACTS

Chapter 2 show that when audit costs are sufficiently low and auditing is imperfect, standard debt contracts with deterministic audit strategies are constrained efficient. Within the class of debt contracts, an important determinant of the nature of optimal repayment schedules is the extent to which lenders can use current information to penalise the entrepreneur for past misreporting. For example, consider an extreme case with a high frequency repeated loans between a borrower and lender. In this situation a high degree of risk sharing can be accomplished even when the borrower has private information to project returns. After some time, any reporting strategy which is systematically dishonest will become apparent to the lender, who can observe that the distribution of reported returns is not converging to the true distribution of project returns. The lender might not be able to detect individual lies, but with time they can easily detect liars.

In practise, there are limits to the extent to which lending relationships are repeated, and how new information can be used to punish historical claims. In this paper, we proceed under the assumption that contract repayments can only be made contingent on information revealed in the current period. Relaxing this assumption would affect repayment schedules, but would not change the qualitative relationships between leverage, marginal rates of substitution across project outcomes and project risk that drive the results of this paper.

Assumption 3.1 *External finance contracts are only contingent on information revealed in the current period.*

This restriction is referred to as the *anonymity* constraint. Once repayments on current period loans are made, entrepreneurs are considered to become anonymous, and their future actions in other markets cannot be used as evidence of past false reports.

We apply Theorem 2.2 from Chapter 2 to motivate standard debt contracts as optimal.

Theorem 2.2 *Let borrowing be taken as given $b = \hat{b}$. When type-I audit errors occur with positive probability ($\eta(\underline{\theta}) > 0$), there exists some strictly positive audit cost $\hat{\kappa}$ such that for all $\kappa < \hat{\kappa}$, standard debt contracts ($q(\underline{m}) = 1$) are efficient.*

As we discuss in Appendix 3.A, it is difficult to pin down audit costs in a way that produces the high credit spreads observed in the data, given the low historical probabilities of corporate default. Microeconomic estimates of direct bankruptcy costs as a share of firms' assets (what would be interpreted as κ in our framework) typically fall between 0.01 and 0.06, which is sufficiently low to be consistent with standard debt contracts being optimal in accordance with Theorem 2.2.[6]

Assumption 3.2 *Audit costs are sufficiently low, such that standard debt contracts are optimal in equilibrium.*

Following Theorem 2.2, contracts are subject to two constraints. First, repayments following overturned low reports must exceed those following high reports,

$$r(\theta_1, \sigma_2) \geq r(\theta_2, \emptyset). \tag{3.5}$$

Equation 3.5 is the incentive compatibility constraint, and we attach to it the Lagrange multiplier μ. Second, expected loan repayments must exceed the sum of the financial intermediaries' deposit interest rate and expected audit costs,

$$\sum_{(\theta_{ti}, \sigma_{ti})} \mathrm{P}(\theta_{ti}, \sigma_{ti}) \hat{r}_t(\theta_{ti}, \sigma_{ti}) K^b_{t-1i} \geq r^b_t K^b_{t-1i} + \pi_1 \kappa K_{t-1i}. \tag{3.6}$$

where r^b_t is the financial intermediary's opportunity cost of funds. Equation 3.6 describes the financial intermediaries' participation constraint, to which we'll attach the Lagrange multiplier ν. Both the incentive compatibility and participation constraints will be binding under efficient contracts.

[6]These estimates are drawn from Warner (1977), Weiss (1990) and Altman (1984).

3.1.3 FIRST ORDER NECESSARY CONDITIONS

Now that we have defined the entrepreneurs' problem, we can take first order necessary conditions:

$$N_{ti}: \; 0 = \mathbb{E}_t \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}) \left[\frac{Y_{Nti}(\theta_{ti})}{X_t} - W_t \right], \tag{3.7}$$

$$K^b_{ti}: \; 0 = \mathbb{E}_t \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}) \left[\frac{Y_{Kti}(\theta_{ti})}{X_t} - r_{ti}(\theta_{ti}, \sigma_{ti}) \right] + \nu_{ti} \left[\mathbb{E}_t \hat{r}_t(\theta_{ti}, \sigma_{ti}) - r^b_t - \pi_1 \kappa \right] \tag{3.8}$$

$$C^e_{ti}(\theta_{ti}, \sigma_{ti}): \; 0 = u^{e\prime}(C^e_{ti}(\theta_{ti}, \sigma_{ti})) - \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}), \tag{3.9}$$

$$\begin{aligned} K^e_{ti}(\theta_{ti}, \sigma_{ti}): \; 0 = & -Q_t \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}) \\ & + \beta^e \mathbb{E}_t \left[\lambda^e_{t+1i}(\theta_{t+1i}, \sigma_{t+1i}) \left(\frac{Y_{Kt+1i}}{X_{t+1}} + Q_{t+1}(1-\delta) \right) - \nu_{t+1i} \pi_l \kappa \right] \end{aligned} \tag{3.10}$$

$$\hat{r}_t(\theta_1, \sigma_1): \; 0 = -\mathrm{P}(\theta_1, \sigma_1) \lambda^e_{ti}(\theta_1, \sigma_1) K^b_{t-1i} + \nu_{ti} \mathrm{P}(\theta_1, \sigma_1) K^b_{t-1i} \tag{3.11}$$

$$\hat{r}_t(\theta_1, \sigma_2): \; 0 = -\mathrm{P}(\theta_1, \sigma_2) \lambda^e_{ti}(\theta_1, \sigma_2) K^b_{t-1i} + \nu_{ti} \mathrm{P}(\theta_1, \sigma_2) K^b_{t-1i} + \mu_{ti} \tag{3.12}$$

$$\hat{r}_t(\theta_2, \emptyset): \; 0 = -\mathrm{P}(\theta_2, \emptyset) \lambda^e_{ti}(\theta_2, \emptyset) K^b_{t-1i} + \nu_{ti} \mathrm{P}(\theta_2, \emptyset) K^b_{t-1i} - \mu_{ti} \tag{3.13}$$

Without loss of generality, $Y_{jti}(\theta_{ti})$ denotes the derivative of output with respect to factor j for entrepreneur i in period t given idiosyncratic shock realisation θ_{ti}. Also, let $\bar{Y}_{jti}$ denote the expectation of the derivative of output with respect to factor j for entrepreneur i in period t over idiosyncratic shock realisations θ_{ti}

3.1.4 RISK ACROSS STATES

Equations 3.11, 3.12 and 3.13 describe how the entrepreneurs' marginal utility (captured by $\lambda^e_{ti}(\theta, \sigma)$) varies across states. Entrepreneurs can vary loan repayment rates across states $r(\theta, \sigma)$ in order to attempt to reduce variations in λ^e_{ti} across states. Entrepreneurs' ability to reduce variations in λ^e_{ti} across states is limited by the entrepreneurs' incentive compatibility constraint (3.5). The incentive compatibility constraint is binding under an efficient contract ($\mu_{ti} > 0$) resulting in varying marginal utilities across idiosyncratic states $\lambda^e_{ti}(\theta_1, \sigma_2) > \lambda^e_{ti}(\theta_1, \sigma_1) > \lambda^e_{ti}(\theta_2, \emptyset)$. Combining equations 3.11, 3.12 and 3.13 yields

$$\nu_{ti} = \lambda^e_{ti}(\theta_1, \sigma_1), \tag{3.14}$$

$$\mu_{ti} = \mathrm{P}(\theta_1, \sigma_2) K^b_{ti} (\lambda^e_{ti}(\theta_1, \sigma_2) - \lambda^e_{ti}(\theta_1, \sigma_1)) \quad \text{and} \tag{3.15}$$

$$\lambda^e_{ti}(\theta_1, \sigma_1) = \mathrm{P}(\theta_1, \theta_1) \lambda^e_{ti}(\theta_1, \theta_1) + \mathrm{P}(\theta_1, \theta_2) \lambda^e_{ti}(\theta_1, \theta_2) + \mathrm{P}(\theta_2, \emptyset) \lambda^e_{ti}(\theta_2, \emptyset). \tag{3.16}$$

In addition to the option of reduced repayments following successful audits, entrepreneurs can mitigate project risk by reducing the size of projects, relative to the size which maximises expected profits. This precautionary reduction in the size of projects translates into reductions in the quantities of capital and labour demanded compared with first best efficient allocations.

3.1.5 Demand for labour and capital

Equation 3.7 describes the entrepreneurs' first order necessary condition for labour demanded. Wage contracts are determined at the beginning of the period, prior to the revelation of idiosyncratic shocks. Idiosyncratic project outcomes affect the marginal product of labour, which means that wages are not equal to the marginal product of labour ex post. When hiring labour, entrepreneurs weight deviations between wages and labour's marginal product according to the likelihood of states and the entrepreneurs' state-contingent marginal utility. With a high marginal utility weight on bad outcomes where the wage rate exceeds the marginal product of labour, entrepreneurs exercise precaution in the labour market. The resulting market labour demand leaves a wedge between the average marginal product across firms and the competitive market wage rate.

Entrepreneurs' risk aversion also affects the capital market. Equation 3.8 presents the first order necessary condition for the quantity of capital rented from the household sector. The first term captures deviations between the marginal product of capital and the capital rental rate. As in the labour market, these deviations are weighted by the entrepreneurs' state-contingent marginal utility. The second term accounts for real resource costs associated with auditing.

The capital market distortion resulting from the costs of auditing represents a real resource cost. On the other hand, the distortions arising from entrepreneurs weighting outcomes by their state contingent marginal utilities affects the distribution of income. Entrepreneur risk aversion reduces the income of the representative household, but increases the expected income of entrepreneurs.

3.2 FACTOR PRICES AND LEVERAGE UNDER FLEXIBLE PRICES

It is useful to first consider factor prices and leverage with perfect competition in product markets. This allows us to consider the distortions arising from the financial friction in isolation from distortions arising from nominal rigidities. Under flexible prices and perfect competition in product markets, the gross retail markup of consumption goods over wholesale prices will be equal to one, $X_t = 1$.

Let the labour and capital market wedges be defined as follows,

$$\tau_{Nti} = \frac{\bar{Y}_{Nti} - W_t}{\bar{Y}_{Nti}} \qquad \text{and} \tag{3.17}$$

$$\tau_{Kti} = \frac{\bar{Y}_{Kti} - r_t^b}{\bar{Y}_{Kti}}. \tag{3.18}$$

Combining equations 3.7, 3.8 and 3.16 yields the following optimality conditions:

$$\frac{\lambda_{ti}^e(\theta_1, \theta_2)}{\lambda_{ti}^e(\theta_1, \theta_1)} = 1 + \frac{\tau_{Nti}}{\mathrm{P}(\theta_1, \theta_2)\xi_t}, \tag{3.19}$$

$$\frac{\lambda_{ti}^e(\theta_2, \emptyset)}{\lambda_{ti}^e(\theta_1, \theta_1)} = 1 - \frac{\tau_{Nti}}{\mathrm{P}(\theta_2, \emptyset)\xi_t} \qquad \text{and} \tag{3.20}$$

$$\tau_{Nti} = \tau_{Kti} - \frac{\pi_1 \kappa}{\bar{Y}_{Kti}}. \tag{3.21}$$

Derivations of equations 3.19, 3.20 and 3.21 can be found in Appendix 3.E.1. Equations 3.19 and 3.20 relate the entrepreneurs' marginal rates of substitution for consumption across project outcomes to the labour market wedge. Equation 3.21 relates the labour market wedge to the capital market wedge. For entrepreneurs, it is efficient to reduce both labour and capital demanded in order to mitigate project risk. Equation 3.21 confirms that the labour market wedge (the left hand side, τ_{Nti}) is less than the capital market wedge (τ_{Kti}). The difference between the two wedges results follows as a result of auditing costs, which are increasing in the capital factor but not in the labour factor.

3.2.1 FACTOR INCOME

The previous subsection showed how the entrepreneurs' inability to share idiosyncratic productive risk with households results in positive labour and capital market wedges.

Returns to the representative household's productive factor inputs are lower than their expected marginal products. Aggregating over entrepreneurs, we can combine the definitions of the factor market wedges (3.17, 3.18) with the optimality condition relating the two wedges (3.21) to derive the composition of output in terms of factor income:

$$Y_t = W_t N_t + r_t^b [K_{t-1}^b + K_{t-1}^e] + Y_t \tau_{Nt} + \pi_l \kappa K_{t-1}. \tag{3.22}$$

The first two terms on the right hand side of (3.22) typically form the decomposition of factor income in a frictionless model. When entrepreneurs are compensated for risk, there are two additional terms: the compensation for risk bearing earned by entrepreneurs is captured by $Y_t \tau_{Nt}$. The fourth term, $\pi_l \kappa K_{t-1}$, captures the resource costs of audits.

3.2.2 Entrepreneurs' savings behaviour

Entrepreneurs' preferences are intertemporally homothetic, and their technology is scalable. Consequently, entrepreneurs' actions in terms of consumption, labour and capital hired are equal as a share of capital brought into the current period. Additionally, efficient loan coupon rates r are identical across entrepreneurs. Taken together, we can describe the aggregate behaviour of the population of entrepreneurs as a function of the mean wealth of entrepreneurs. Mean preserving fluctuations in the ex ante distribution of wealth across entrepreneurs do not affect market prices or aggregate quantities traded.

Also note that under logarithmic utility, where the income and substitution effects of interest rates on savings cancel, the efficient savings decision of entrepreneurs which uniquely satisfies equation 3.10 is

$$C_{ti}^e = \frac{1 - \beta^e}{\beta^e} Q_t K_{ti}^e. \tag{3.23}$$

A derivation of equation 3.23 is found in Appendix 3.E.2.

3.2.3 GREAT RATIOS

We can re-write equation 3.21 in two alternative ways, in order to find the efficient labour-capital and output-capital ratios in terms of factor prices and parameters:

$$\frac{N_{ti}}{K_{t-1i}} = \frac{1-\alpha}{\alpha}\frac{r_t^b + \pi_l \kappa}{W_t} \qquad \text{and} \tag{3.24}$$

$$\frac{\bar{Y}_{ti}}{K_{t-1i}} = Z_t \left[\frac{1-\alpha}{\alpha}\frac{r_t^b + \pi_l \kappa}{W_t}\right]^{1-\alpha}. \tag{3.25}$$

3.2.4 SOLVING FOR LEVERAGE

Given that C_{ti}^e is equal to a fixed proportion of K_{t+1i}^e regardless of idiosyncratic state (equation 3.23), and that $C_{ti}^e = 1/\lambda_{ti}^e$ by equation 3.9, we can write down the ratios $\lambda_{ti}^e(\theta_1, \theta_2)/\lambda_{ti}^e(\theta_1, \theta_1)$ and $\lambda_{ti}^e(\theta_2, \emptyset)/\lambda_{ti}^e(\theta_1, \theta_1)$ in terms of the accumulation constraints 3.4:

$$\frac{\lambda_{ti}^e(\theta_1, \theta_2)}{\lambda_{ti}^e(\theta_1, \theta_1)} = \frac{Y_{ti}(\theta_1) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b r_{ti}(\theta_1, \sigma_1) - W_t N_{ti}}{Y_{ti}(\theta_1) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b r_{ti}(\theta_1, \sigma_2) - W_t N_{ti}} \qquad \text{and} \tag{3.26}$$

$$\frac{\lambda_{ti}^e(\theta_2, \emptyset)}{\lambda_{ti}^e(\theta_1, \theta_1)} = \frac{Y_{ti}(\theta_1) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b r_{ti}(\theta_1, \sigma_1) - W_t N_{ti}}{Y_{ti}(\theta_2) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b r_{ti}(\theta_2, \emptyset) - W_t N_{ti}}. \tag{3.27}$$

Combining (3.26,3.27,3.19,3.20), we can first solve for the efficient amount of risk sharing obtained by entrepreneurs. Entrepreneurs can set capital rental repayment rates on a contingent basis, enabling partial risk sharing.

$$K_{t-1i}^b[r_{ti}(\theta_2, \emptyset) - r_{ti}(\theta_1, \sigma_1)] = \bar{Y}_{ti}\left[\frac{\pi_2 \xi_t - \tau_{Nti}}{\pi_2 + \pi_1 \eta}\right] \tag{3.28}$$

Equation 3.28 shows that risk sharing through differentiated repayment rates across states is limited as a share of project risk $\mathbb{E}Y_{ti}\xi_t$. This means that for each entrepreneur, an increase in output through hiring more labour and renting more capital will increase the risk borne by that entrepreneur.

It will be helpful to derive a measure of leverage, and to show how leverage relates to the labour wedge, τ_N. From the entrepreneur's perspective, an increase in labour or capital hired from the household sector both increase the risk of projects, and the expected factor payments due at the end of projects. It makes sense therefore to include labour payments in our measure of leverage. One useful measure of leverage is the following:

$$L_{ti} = \frac{\bar{Y}_{ti}}{R_t Q_{t-1} K^e_{t-1i}} \tag{3.29}$$

Project output $\bar{Y}_{ti}$ is the expectation of the total income generated by the project, and $R_t Q_{t-1} K^e_{t-1i}$ is the net worth of entrepreneur i, expressed in terms of their opportunity cost, which was to redeem their capital holdings for deposits at the end of the period $t-1$.

After some rearranging, substituting equation 3.28 into equation 3.26 yields

$$L_{ti} = \frac{(\pi_2 + \pi_1 \eta)\tau_{Nti}}{[\pi_2 \xi_t - \tau_{Nti}][\pi_1 \eta \xi_t + \tau_{Nti}]}, \tag{3.30}$$

which can be rearranged to yield an approximate solution for τ_{Nti},[7]

$$\tau_{Nti} \approx \pi_2 \left[\xi_t - \frac{1}{L_{ti}}\right] + \pi_1 \eta \left[\frac{L_t^2 \xi_t^2 + 1}{L_{ti}\left[L_{ti}\xi_t - 1\right]}\right] + \mathcal{O}([\pi_1 \eta]^2).$$

Derivations of equations 3.28 and 3.30 are contained in Appendix 3.E.2. Consider equation 3.30. The left hand side is our production or income based measure of leverage. The right hand side is increasing in τ_{Nti}, indicating that all else equal, an increase in leverage means an increase in the labour market wedge (by equation 3.21, this also translates into an increase in the capital market wedge). This is because an increase in leverage requires the entrepreneur to accept a greater share of productive risk. Entrepreneurs hire factors until their expected marginal product, weighted by their marginal rates of substitution across states, is equal to their prices, in this case wages and interest rates. When leverage is high, the entrepreneur bears more risk. The marginal rates of substitution between worse and better project outcomes increase, increasing the wedge between the risk adjusted expected marginal product of factors, which determines demand for factors in our model, and the risk neutral expectation of factor marginal products, which would equal factor prices in a perfect information environment.

[7]While serving as a useful approximation for some applications, it is important to note that the behaviour of this function diverges significantly from the true solution when τ_N approaches zero. In particular, unlike equation 3.30, the approximation does not restrict τ_N to positive values.

In terms of business cycle dynamics, the positive relationship between leverage and factor market wedges expressed in equation 3.30 marks the main departure of our model from the related literature.

Equation 3.30 makes two clear predictions about business cycle dynamics. First, real wages and real interest rates will respond sluggishly to total factor productivity shocks. Consider a shock to total factor productivity. This will increase output, and the leverage ratio. According to equation 3.30, this increase in the leverage ratio will also increase the factor market wedge, captured by τ_N. In other words, real wages and real interest rates will respond sluggishly to movements in marginal products driven by fluctuations in total factor productivity.

Second, real wages and interest rates will diverge from their marginal products in response to risk shocks (movements in ξ_t). For any given leverage ratio, an increase in risk will mean a greater factor market wedge τ_N, and consequently a decrease in real wages and real interest rates. When labour supply is elastic, or the wage setting process exhibits additional rigidities, then employment and output will decrease in response to positive shocks to project risk.

An alternative measure of leverage is the ratio of capital holdings between the two populations, K^h/K^e. This measure is closer to that considered in related literature, but in the context of our model is less useful. Importantly, it does not capture wage bill obligations which are determined at the beginning of the period and cannot be renegotiated in the case of a bad project outcome. Combining equations 3.6, 3.28 and 3.30 yields

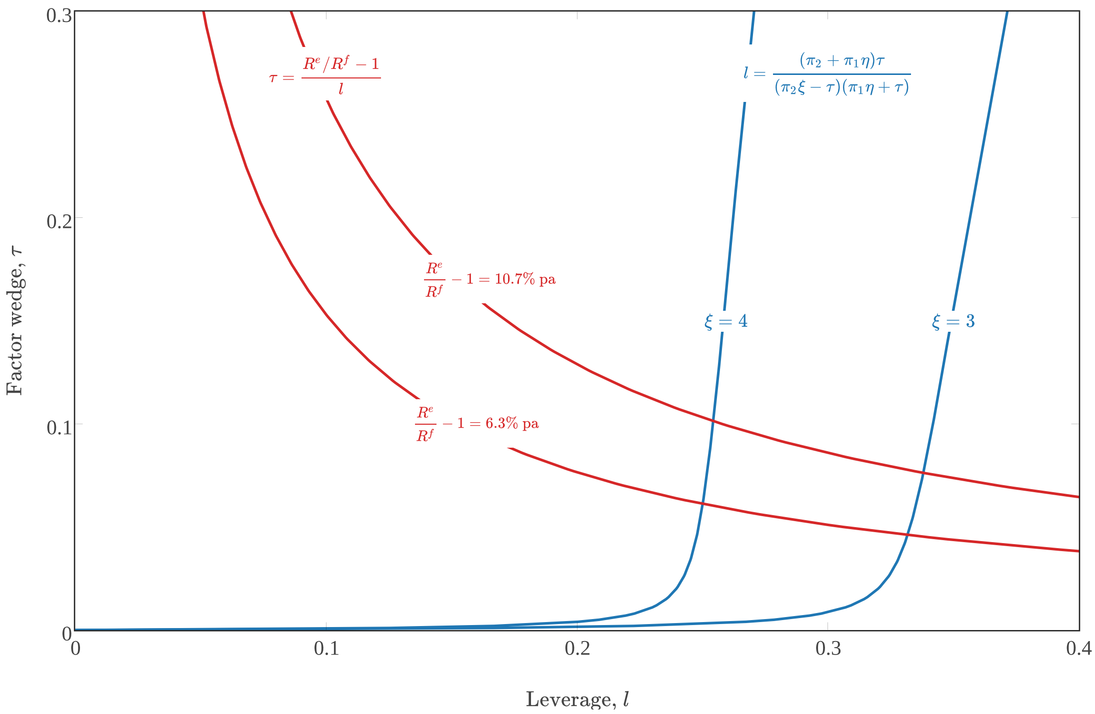
$$\frac{K^b_{t-1i}}{K^e_{t-1i}}([r_{ti}(\theta_2,\emptyset)-r^b_t]-\pi_1\kappa)=R_tQ_{t-1}\left[\frac{\pi_1(1-\eta)\tau_{Nti}}{\pi_1\eta\xi_t+\tau_{Nti}}\right]+\pi_1\kappa \tag{3.31}$$

Note that $[r_{ti}(\theta_2,\emptyset)-r^b_t]$ can be interpreted as the interest rate risk premium on loans, the difference between the loan coupon rate and the deposit rate.

3.2.5 Equilibrium in the capital market

Figure 3.2 presents the determination of equilibrium leverage and factor market wedges. The blue, upward sloping schedules depict equation 3.30, the increasing relationship between leverage and risk, which draws the entrepreneurs' idiosyncratic marginal rates of substitution away from unity, resulting in the factor market wedge τ. This schedule is determined by parameters taken as exogenous in our analysis including project risk,

Figure 3.2: Equilibrium in the capital market.[a]



[a]Unless otherwise stated, all parameter values are taken from our benchmark parameterisation.

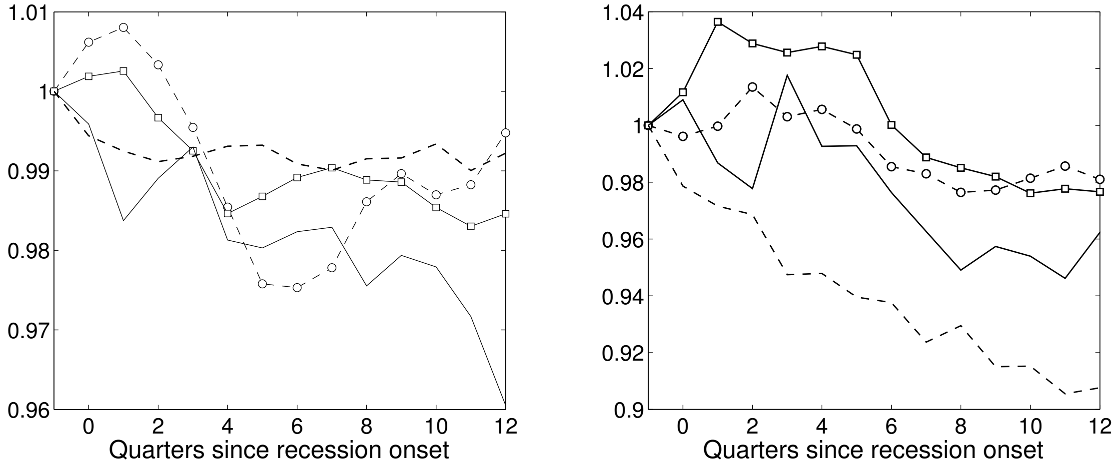
the probability of bad idiosyncratic outcomes, the probability of audit error and the entrepreneurs' tolerance for risk (an increase in any of which would result in a leftward shift of the blue schedule.

The downward sloping red schedules plot the relationship between the factor market wedge, leverage and the equity risk premium. While endogenous in the short run, the long run equity risk premium is tied down by the discount rates of the worker households and entrepreneurs. Without loss of generality, when $\beta^h R^f > 1$, the wealth of worker households will tend to be increasing. When $\beta^e R^e > 1$, the (mean) wealth of entrepreneurs will tend to be increasing.

3.3 Model dynamics

Figure 3.3, reproduced from the Introduction for convenience, describes two measures of the behaviour of wages to the marginal product of labour for the United States recessions beginning in 1974Q1, 1981Q4, 2001Q2 and 2008Q1. For both the measure derived from national income accounts (left hand panel) and the measure derived from firm surveys

Figure 3.3: Labour income and productivity in US NBER recessions (Recession starting 1974Q1 dashed o, 1981Q4 solid □, 2001Q2 dashed, 2008Q1 solid). Left hand panel: Labour share of National Income (BEA NIPA Table 2.1). Right hand panel: Nonfarm real wages relative to labour productivity (BLS PRS85006102, PRS85006092. Nominal wages deflated by the consumer price index, St. Louis Federal Reserve FRED Database CPIAUCSL). Both series normalised to one at the onset of the respective recessions.

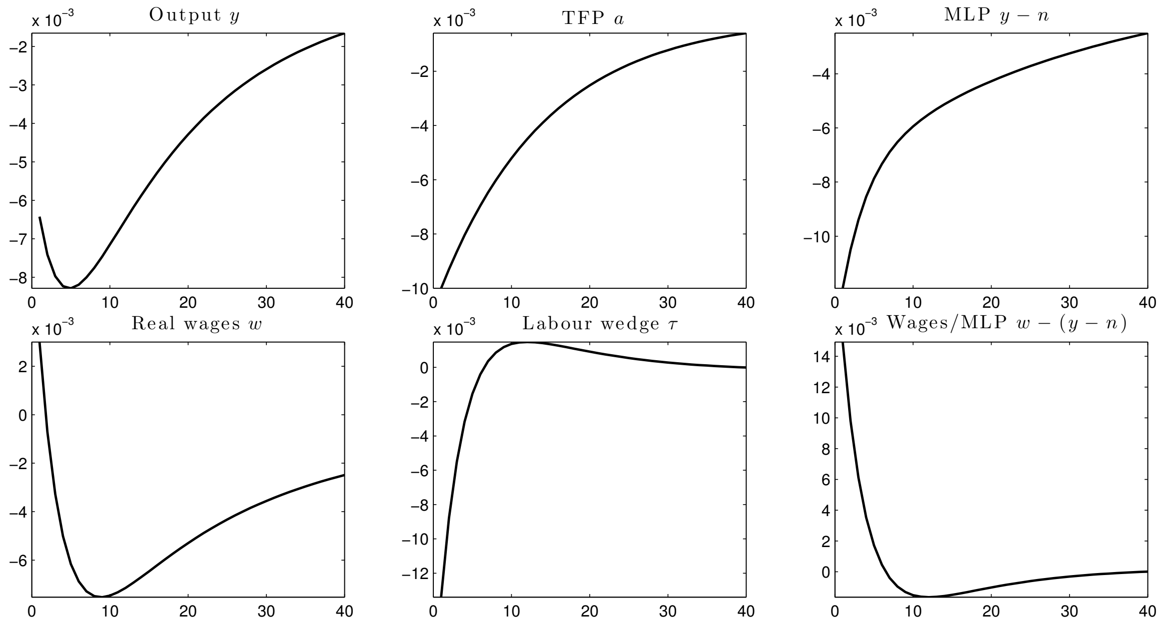


(right hand measure), wages tend to increase relative to marginal labour productivity at the onset of the 1974 and 1981 recessions before falling below marginal labour productivity over the medium run. For the more recent recessions starting in 2001 and 2008, real wages tend to lag marginal labour productivity growth from the onset of the recession and for the whole sample period.

Qualitatively, these trends can be reproduced within the flexible price version of our model. Figures 3.4 and 3.4 display the dynamics of selected variables related to the labour wedge, following total factor productivity (TFP) and risk shocks respectively. Full impulse responses are presented in Appendix 3.F. Following the negative total factor productivity shock, output and leverage falls, resulting in a short run decrease in the labour wedge. At the onset of the shock, real wages actually rise initially, and over the short run real wage falls lag the fall in labour productivity. Over the medium run, wages fall below the path of marginal labour productivity. The lower right hand panel presents the dynamics of real wages relative to labour productivity, and it is this panel that within our model can be interpreted as equivalent to the panels presented in 3.3. The relative paths of real wages and labour productivity resemble those observed in the 1974 and 1981 recessions, initially rising before falling over the medium run.

Following the risk shock, the project risk absorbed by the entrepreneurs increases sharply, and this manifests itself through a sharp increase in the labour wedge. Factor

Figure 3.4: Dynamics of important labour market variables. Negative total factor productivity (TFP) shock (-1%). Log-linearised model. Periods measured in quarters.

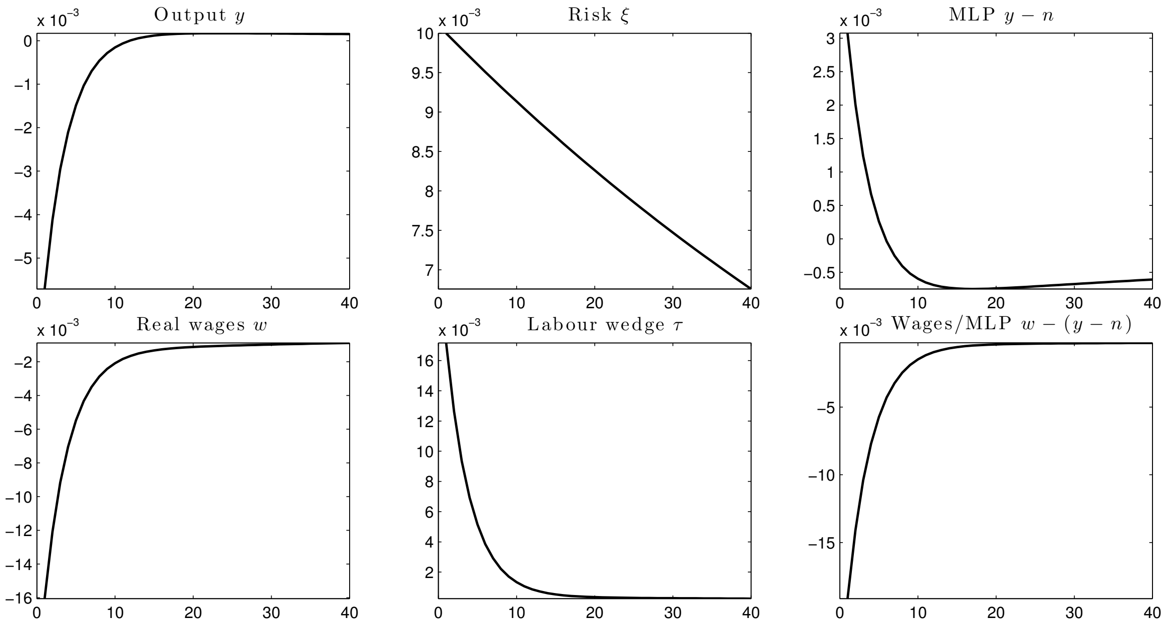


prices including wages fall, although marginal labour productivity actually increases as employment falls in the short run while the capital stock is inelastic. Taken together, the diverging paths of real wages and labour productivity mean that the fall in wages relative to the path of labour productivity is large, approximately three times the magnitude of the fall in output under our parameterisation.

Figures 3.6 and 3.7 present impulse selection for a broader set of model variables under flexible prices and sticky retail prices respectively, where the sticky price model is closed by a Taylor rule responding solely to contemporaneous inflation. The full parameterisations are available in Appendix 3.A. The broad qualitative dynamics of the model are similar under flexible and sticky price settings, with the main distinction being the now humped-shaped and persistent response to the risk shock under sticky prices.

Under the Taylor-type monetary policy rule with which we close the model, the real interest rate response to shocks is sluggish compared with the rapid responses exhibited under flexible prices. In the case of the positive TFP shock, the sluggish increase in the real interest rate means that entrepreneurs' (for whom the real interest rate is an expense) accumulate capital quickly in response to the shock, but their initial leverage is high (the real interest rate is also a determinant of their net worth—it reflects their opportunity cost

Figure 3.5: Dynamics of important labour market variables. Risk shock (1%). Log-linearised model. Periods measured in quarters.



of projects). Taken together, the path of the factor wedge is more volatile than in the flexible price counterfactual: it rises sharply at the onset of the shock, then falls sharply soon after. Consequently, the paths of wages, hours worked, output and household consumption are also more volatile than under the flexible price counterfactual.

Turning to the (recessionary) risk shock, again it is the dampened initial response of the real interest rate to the shock that propagates the effects of the shock. At the onset of the shock, the relatively high real interest rate keeps leverage lower than under the flexible price counterfactual, dampening the initial response of the factor wedge. But the relatively high real interest rate reduces the equity premium, slowing the accumulation of entrepreneurial capital, fundamental to the adjustment process.

Following the risk shock, output returns to near its steady state level within 20 quarters. But this relatively quick return of output to trend masks the adjustment processes that are still occurring. After 40 quarters, there remain large differences in the inequality of consumption and capital holdings across the household and entrepreneur groups. Perhaps it is these persistent and sluggish trends in consumption and wealth inequality that are part of the reason why financial crises tend to provoke concerns about the *fairness* of the sharing of business cycle risks.

3.4 Extensions

The main message of this paper is that during periods of financial stress, the demand for labour will fall sharply. This reduces the risk absorbed by firm insiders, mitigating the effects of the financial stress. How this reduction in the demand for labour is reflected by increases in unemployment and decreases in wages will be determined in part by workers' preferences and wage and price setting frictions.

Within this section, we outline some of the implications of the model for business cycle dynamics. First, we show that within a New Keynesian setting, the marginal cost of production will include a risk premium, which increases during times of financial stress. This means that following a negative financial shock, cost pressure on inflation will be greater than that predicted by trends in wages and labour productivity. Second, we outline three candidate financial shocks (risk, risk tolerance and loan monitoring) and show how they result in similar business cycle dynamics.

3.4.1 Wage growth and retail price inflation

Within a New Keynesian setting, retailers purchase the entrepreneurs' wholesale output good at relative price $1/X_t$, and sell differentiated retail goods to entrepreneurs and households. From the retailers' perspective, the marginal cost of production is $1/X_t$. When retailers are restricted in their ability to adjust prices in response to fluctuations in marginal costs, retail price inflation in the current period will be a function of current and expected future fluctuations in marginal costs. In this Chapter, we motivate New Keynesian price setting rigidities through Rotemberg adjustment costs, and the resulting New Keynesian Phillips' curve takes the following form

$$(1-\epsilon)+\frac{\epsilon}{X_t}-\phi^{\Pi}\Pi_t(\Pi_t-1)+\phi^{\Pi}\beta\mathbb{E}_t\left[\left(\frac{\lambda_{t+1}^h}{\lambda_t^h}\right)\frac{Y_{t+1}}{Y_t}\Pi_{t+1}(\Pi_{t+1}-1)\right]=0, \quad (3.32)$$

where $\epsilon > 0$ is the demand elasticity of substitution between differentiated varieties, $\phi^{\Pi} > 0$ is the adjustment cost parameter and Π is the gross inflation rate. A full derivation is included in Appendix 3.C. Current inflation will be high if current marginal costs or expected inflation are high.

We can derive marginal costs as a function of labour costs and leverage from the entrepreneurs' problem. When product market nominal rigidities are present, leverage across entrepreneurs becomes

$$L_t = \frac{Y_t}{X_t R_t Q_{t-1} K^e_{t-1}}, \tag{3.33}$$

where $\phi_x > 0, \hat{x}_t$ is the log difference in retail markup over wholesale prices from the steady state.

Leverage and labour costs are related according to[8]

$$\frac{1}{X_t} = \frac{\pi_2}{1-\alpha}\frac{W_t N_t}{Y_t}\left(\xi_t - L_t^{-1} + 1/\pi_2\right) + O(\pi_1 \eta). \tag{3.34}$$

Taking a log-linear approximation around the steady state we obtain[9]

$$-\hat{x}_t = \hat{w}_t + \hat{n}_t - \hat{y}_t + \phi_l \hat{l}_t + \phi_\xi \hat{\xi}_t + O(\pi_1 \eta), \tag{3.35}$$

where all variables are expressed in terms of log deviations from their steady state values. The left hand side of equation 3.35 captures marginal costs. Within the benchmark New Keynesian model, or in standard financial frictions models including Bernanke, Gertler, and Gilchrist (1999), marginal costs consist solely of marginal labour costs, and are equal to the first three terms on the right hand side $\hat{w}_t + \hat{n}_t - \hat{y}_t$, which capture the real wage and the marginal productivity of labour. The terms $\phi_l \hat{l}_t + \phi_\xi \hat{\xi}_t$ capture the risk premium earned by entrepreneurs who absorb some of the additional productive risk associated with an increase in firm size. Increases in leverage $\hat{l}_t$ and project risk $\hat{\xi}_t$ increase project risk and the risk premium added to marginal labour costs to obtain the marginal cost of production.

This means that the relationship between labour marginal costs and total marginal costs will be time varying, and importantly will respond to financial stress. During periods of high financial stress, cost pressure on inflation will be high relative to prevailing wage and productivity growth.

[8] A derivation is contained in Appendix 3.C.

[9] where

$$\phi_l = \frac{\pi_2}{1-\alpha}\frac{X_{ss} W_{ss} N_{ss}}{Y_{ss} L_{ss}}, \qquad \phi_\xi = \frac{\pi_2}{1-\alpha}\frac{X_{ss} W_{ss} N_{ss} \xi_{ss}}{Y_{ss}}.$$

Note that here we've continued to restrict financial shocks to risk shocks, as in Christiano, Motto, and Rostagno (2014). Note however that other financial shocks including fluctuations in entrepreneurs' risk tolerance or in audit quality would have similar first order effects as fluctuations in project risk $\hat{\xi}_t$. Section 3.4.2 considers the relative effects of these three potential financial shocks.

3.4.2 Financial shocks

Several recent studies have pointed toward financial shocks affecting the intermediation of capital loans as an important determinant of business cycle dynamics (see Benk, Gillman, and Kejak, 2005, Nolan and Thoenissen, 2009, Fornari and Stracca, 2012, Jermann and Quadrini, 2012 and Christiano, Motto, and Rostagno, 2014). While financial shocks in our model differ from the standard financial accelerator model, the primary difference is the inclusion of a direct effect on the labour market wedge, which is an important feature of the model described by Jermann and Quadrini (2012).[10]

There remains some debate about the interpretation of these financial shocks. Christiano, Motto, and Rostagno (2014) find in an estimated model based on Bernanke, Gertler, and Gilchrist (1999) that shocks to the dispersion of firms' project outcomes (risk shocks) are a suitable financial shock and can explain a large share of business cycle volatility in US data. Pinter, Theodoridis, and Yates (2013) cast some doubt on this finding, arguing that risk shocks as identified by market expectations of stock price volatility cannot account for a large share of business cycle volatility.

For the bulk of the analysis within this paper, we've assumed that the 'financial shock' takes the form of an increase in idiosyncratic project risk across firms. Within our model, the effects of a risk shock on the dynamics of macroeconomic aggregates are isomorphic to the effects of a shock to loan monitoring quality, and are similar to the effects of a shock to entrepreneurs' risk tolerance.[11] All three potential shocks enter the model through the relationship between leverage and the labour market wedge outlined by equation 3.30.[12]

[10]It is important to note that within the model described by Jermann and Quadrini (2012), the financial friction is motivated by inefficient tax design, rather than by appeal to informational frictions.

[11]A shock to entrepreneurs' risk tolerance would also affect the entrepreneurs' savings responses to shocks, but for the purposes of this comparison we will focus on the relationship between our selected financial shocks, leverage and the labour market wedge outlined by equation 3.30. Numerical simulations suggest that the effects on entrepreneurs' savings of fluctuations in risk tolerance are small.

[12]The three shocks have differential effects on interest rate spreads, but within the model their effects on macroeconomic aggregates can be isolated from their effects on interest rate spreads.

We'll compare our three financial shocks within the flexible price model ($X_t = 1$) by taking first order approximations of equation 3.30 in terms of log deviations from steady state of leverage $\hat{l}_t$, project risk ξ_t and the entrepreneurs' coefficient of relative risk aversion $[\widehat{\log \gamma^e}]_t$, and absolute deviations in terms of the labour wedge $\hat{\tau}_{Nt}$ and type-I audit error $\hat{\eta}_t$. We will take this approximation around the initial values $\eta \to 0^+$ and $\gamma^e = 1$.

First we need to work through the entrepreneurs' problem again, this time allowing γ^e to take any positive value. Equation 3.30 is replaced by

$$L_{ti}^{-1} = \frac{[\pi_2\xi_t - \tau_{Nt}]\left[\dfrac{(\pi_2+\pi_1\eta)\tau_{Nt}}{\pi_1\eta[\pi_2\xi_t - \tau_{Nt}]} + 1\right]^{\frac{1}{\gamma^e}} + \pi_1(1-\eta)\xi_t\left[\dfrac{\tau_{Nt}}{\pi_1\eta\xi_t} + 1\right]^{\frac{1}{\gamma^e}} + \pi_1\eta\xi_t + \tau_{Nt}}{\left[\dfrac{(\pi_2+\pi_1\eta)\tau_{Nt}}{\pi_1\eta[\pi_2\xi_t - \tau_{Nt}]} + 1\right]^{\frac{1}{\gamma^e}} - 1}. \tag{3.36}$$

Now, assuming η_t and γ_t^e to be time-varying around steady state values of $\eta_{ss} \to 0^+$ and $\gamma_{ss}^e = 1$, we can derive the following expression in log linear and linear differences from the steady state,

$$\varphi_\tau\hat{\tau}_t = \hat{l}_t + \underbrace{\varphi_\xi\hat{\xi}_t + \varphi_\gamma[\widehat{\log \gamma^e}]_t + \varphi_\eta\hat{\eta}_t}_{\text{financial shocks}} \tag{3.37}$$

where

$$\varphi_\tau = \xi_{ss}L_{ss} - 1, \quad \varphi_\xi = \xi_{ss}L_{ss}, \quad \varphi_\gamma = \pi_1\log(\xi_{ss}L_{ss}), \quad \varphi_\eta = \frac{\pi_1}{\pi_2(\xi_{ss}L_{ss} - 1)}.$$

For any value of leverage, the factor market wedge τ_t is increasing in project risk ξ_t, relative risk aversion γ_t^e and audit errors η_t. When the probability of a bad project outcome π_1 is low, small fluctuations in project risk ξ have a large impact on financial stress relative to fluctuations in relative risk aversion γ^e or audit quality η. When the probability of the bad project outcome π_1 is high, this means that the likelihood of bankruptcy is greater, and the costs of poor monitoring quality will be larger along with the effects of fluctuations in risk tolerance.

3.A Parameterisation

The worker households' preferences over consumption and labour are described by $u(C, N) = C^{1-\sigma}/(1-\sigma) - N^{1+\psi}/(1+\psi)$, with $\sigma = 2$ and $\psi = 1.5$. The difference between the worker households' and entrepreneurs' intertemporal elasticities of substitution $(1/\sigma)$ are important for our results. Recall that the entrepreneurs' intertemporal elasticity of substitution is equal to one. Typically, when worker households' intertemporal elasticities of substitution are low, the equilibrium insurance flows toward worker households will be greater in recessions. Consequently, the amplification of shocks will be larger.

The worker households' quarterly discount factor is $\beta^h = 0.995$, and constant steady state wealth shares require that the entrepreneurs' quarterly discount factor is $\beta^e = 0.975$. In the steady state, the risk free interest rate ($R = 1/\beta^h$), and the average return to entrepreneurs' equity is ($R^e = 1/\beta^e$).[13] The probability of a bad project outcome is $\pi_1 = 0.00415$, which corresponds to an annual probability of default of 1.66%, the average historical annual default probability of credit rated US firms (Schuermann and Hanson, 2004). and the probability of a type-I error resulting from an audit is $\eta = 0.1$. Small changes in this value have no effect on our main results, so long as η remains strictly positive. The resource cost of auditing (bankruptcy costs), expressed as a share of capital invested in the project is $\kappa = 0.3$, which is high in comparison with microeconomic studies that have found direct bankruptcy costs of between 1% and 6% of firms' assets (see for example Warner, 1977, Weiss, 1990 and Altman, 1984). This is quite a common weakness of costly state verification models. It is difficult to obtain the high interest rate spreads observed in the data with both realistically low default rates and bankruptcy costs. Across parameterisations matching any two of these three variables with available data has little effect on the dynamics of the model. Given that credit spreads resemble an insurance premium in our model, it is possibly the case that greater entrepreneur risk aversion would help in obtaining the high and volatile credit spreads seen in the data while still retaining realistically low default probabilities and bankruptcy costs. That being said, we would expect entrepreneurs to self select as being relatively risk tolerant, so it is unlikely that greater risk aversion is the best way to bring the model closer to the financial data.

[13] It would be possible to write up the model with a common discount factor across worker households and entrepreneurs, but, given entrepreneurs' relatively high expected return to capital, there would need to be some process governing the exit or death of entrepreneurs in the steady state in order to ensure a stable steady state distribution of capital wealth between worker households and entrepreneurs.

Table 3.1: Properties of the deterministic steady state

	STEADY STATE VALUE	
	Data	Model
GREAT RATIOS		
Capital-output†, K/Y	2.74	2.99
Consumption-output, C/Y	0.65	0.75
Hh. consumption share, C^h/C	NA	0.83
Investment-output, I/Y	0.16	0.24
Ent. capital share, K^e/K	0.42	0.42
Labour share, NW/Y	0.56	0.58
FINANCIAL VARIABLES		
Real interest rate†, $R^f - 1$	1.35%	2.01%
Return on equity†, $R^e - 1$	9.43%	10.61%
Credit spread†, $r(\theta_2, \emptyset) - r^b$	2.04%	1.91%

† These figures have been converted to annualised values.

The idiosyncratic risk coefficient is 5, which is sufficiently large that in low states θ_1, individual output is negative in the sense that more capital is destroyed than output produced. The steady state factor wedge is quite sensitive to the entrepreneurs' discount factor (the lower the discount factor, the less savings the entrepreneurs will accrue) and idiosyncratic risk. Under our parameterisation, the steady state factor wedge is τ_N is equal to 10.25%.

The Cobb-Douglas weight on capital is $\alpha = 0.35$. The depreciation rate of capital is $\delta = 0.02$ per quarter. The investment adjustment cost parameter is $\phi = 4$. The persistence of both risk and total factor productivity shocks are set at $\rho_A = 0.93, \rho_\xi = 0.99$, the standard deviation and magnitude of impulse response for both shocks are set at $\sigma_A = 0.01, \sigma_\xi = 0.01$.

3.A.1 DATA

All macroeconomic variables are taken from the St Louis Federal Reserve FRED database, except where otherwise stated. FRED unique identifies are provided in brackets.

The risk free real interest rate is calculated from the Effective Federal Funds rate (FEDFUNDS).[14] The average return on equity is the annualised percentage change in

[14] Using the Effective Federal Funds rate (FEDFUNDS) results in an average real interest rate of 1.35%, which compares with 1.00% derived from the 90 day treasury bill rate (TB3MS).

the Russell 3000 Total Market Return Index (RU3000TR_PC1).[15] Both interest rates are converted to real returns by subtracting CPI inflation (CPIAUCSL_PC1). The average annualised credit spread is taken from Gilchrist and Zakrajsek (2012). The average capital leverage ratio is taken from Kalemli-Ozcan et al. (2012) and is set to the upper end of the range of financial leverage (Assets / Equity) values they find for US listed non-financial firms of 2.4.[16]

For the sticky price model, the elasticity of substitution between differentiated varieties of the retail good is $\epsilon = 11$ and the Rotemberg adjustment cost parameter is $\phi^{\Pi} = 116$, standard values in the literature. The Taylor rule coefficient on inflation is 1.5.

Note that in our model, the consumption and investment shares of output add to one. This is not true in the data, which includes Government Spending.

3.B The rest of the model

3.B.1 Households

There exists a representative household enjoying consumption c and supplying labour n in each period. Households maximise

$$\mathcal{U}_t^h = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j U^h(C_{t+j}, N_{t+j}), \tag{3.38}$$

where $U_1^h, -U_{11}^h, U_2^h, U_{22}^h > 0$.

At the start of the period, the representative household holds deposits D_{t-1} which are claims issued by the financial intermediaries. At the end of the period, the financial intermediaries repay the household with interest on their deposits $R_t D_{t-1}$, and wage income $W_t N_t$. The representative household uses these deposit holdings to purchase consumption

[15]For comparison, the average real return on equity for United States Banks (USROE) over the time period is equal to 8.65%, which compares with the average of 9.43% calculated from the Russell 3000 Index.

[16]Estimates on this measure vary dramatically, with McGrattan and Prescott (2005) finding an average Assets / Equity ratio of 1.2. This makes quite a big difference to the dynamics of the model, but the main points made within this paper still hold under a wide range of steady state leverage ratios.

goods C_t, and the remaining deposits are carried over to the following period D_t,

$$D_t + C_t = R_t D_{t-1} + W_t N_t. \tag{3.39}$$

The first order conditions of the representative household can be described as follows:

$$-\frac{U_2^h(C_t, N_t)}{W_t} = U_1^h(C_t, N_t) \tag{3.40}$$

$$U_1^h(C_t, N_t) = \beta \mathbb{E}_t R_{t+1} U_1^h(C_{t+1}, N_{t+1}). \tag{3.41}$$

3.B.2 Financial intermediaries

Financial intermediaries are perfectly competitive and risk neutral, earning zero profits in equilibrium. Financial intermediaries provide a deposit asset to households, and intermediate payments from entrepreneurs to households. Between periods, financial intermediaries hold the durable capital good. Within periods, these capital holdings are lent to entrepreneurs.

Loan contracts specify interest rates following successful projects, and recovery rates following unsuccessful projects (bankruptcies). As the probability of project success is variable, intermediaries' diversified portfolios are subject to aggregate risks. This risk is passed on to the household sector depositors.

At the end of each period, all deposits are backed by capital holdings

$$D_t = Q_t K_t^b \tag{3.42}$$

where K_t^f is the amount of capital held by financial intermediaries at the end of period t, and Q_t is the cost of capital in period t. All loan earnings are passed on directly to the representative household through interest on deposits.

$$D_{t-1} R_t = Q_t (1 - \delta) K_{t-1}^b + r_t^b K_{t-1}^b \tag{3.43}$$

The first term on the right hand side of equation 3.43 is the value of bank capital holdings in the current period after depreciation δ. The second term is the capital rental income earned through loans to entrepreneurs, net of monitoring costs.

3.B.3 CAPITAL PRODUCERS

Competitive capital producers (indexed by j) combine the combine the consumption good with existing capital to produce new capital goods. Firm j can produce I_{jt} units of the investment good for total cost

$$I_{tj} + \Phi\left(\frac{I_{tj}}{K_{t-1j}}\right) K_{t-1j}, \tag{3.44}$$

where

$$\Phi\left(\frac{I}{K}\right) = \frac{\phi}{2}\left(\frac{I}{K} - \delta\right)^2.$$

In competitive equilibrium, the cost of capital can be described as follows,

$$Q_t = 1 + \Phi'\left(\frac{I_t}{K_{t-1}}\right) = 1 + \phi\left(\frac{I_t}{K_{t-1}} - \delta\right). \tag{3.45}$$

The final condition we require ensures market clearing in the goods market:

$$Y_t = C_t + C_t^e + Q_t I_t + \pi_1 \kappa K_{t-1} \tag{3.46}$$

3.C THE NEW KEYNESIAN MODEL

In this appendix we outline the key features of the New Keynesian extension to the model. Monopolistically competitive retailers indexed by j purchase wholesale output of entrepreneurs and produce differentiated products without cost. Households' consumption bundle includes individual consumption goods indexed by j:

$$C_t = \left[\int_0^1 C_t(j)^{1-\frac{1}{\epsilon}} di\right]^{\frac{\epsilon}{1-\epsilon}}. \tag{3.47}$$

The representative household also receives any profits χ accruing to the retailers, in which they own a diversified portfolio of shares:

$$D_t + C_t = R_t D_{t-1} + W_t N_t + \chi_t. \tag{3.48}$$

Entrepreneurs's consumption bundles and the investment good bundle are also adapted in a similar way:

$$C_t^e = \left[\int_0^1 C_t^e(j)^{1-\frac{1}{\epsilon}} di\right]^{\frac{\epsilon}{1-\epsilon}} \tag{3.49}$$

$$I_t = \left[\int_0^1 I_t(j)^{1-\frac{1}{\epsilon}} di\right]^{\frac{\epsilon}{1-\epsilon}} \tag{3.50}$$

Let X_t be the gross markup of retail prices over the price of wholesale goods produced by entrepreneurs. The entrepreneurs' capital accumulation constraint becomes

$$Q_t K_{ti}^e + C_{ti}^e = \frac{1}{X_t} Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b r_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti}, \tag{3.51}$$

where $Q, W, r, 1/X, \kappa$ remain prices expressed in real terms, relative to the price level of consumption goods P. Note that the labour and capital market wedges must be re-defined in terms of marginal revenue products as follows:

$$\tau_{Nti} = \frac{\bar{Y}_{Nti} - X_t W_t}{\bar{Y}_{Nti}}. \tag{3.52}$$

In addition, the leverage ratio becomes

$$L_{ti} = \frac{\bar{Y}_{ti}/X_t}{R_t Q_{t-1} K_{t-1i}^e} \tag{3.53}$$

Working again through the entrepreneurs' problem, we can re-write equation 3.30 in terms of the retail markup over wholesale prices X_t:

$$X_t = \frac{1-\alpha}{\pi_2} \frac{\bar{Y}_{ti}}{W_t N_{ti}} \big/ \left(\xi_t - L_{ti}^{-1} + 1/\pi_2\right) + O(\pi_1 \eta) \tag{3.54}$$

Note that $(1-\alpha)\dfrac{\bar{Y}_{ti}}{W_t N_{ti}}$ is the traditional markup in the benchmark New Keynesian model with Cobb-Douglas production. What equation 3.54 shows is that for any marginal labour cost of production, higher leverage L means a smaller markup. This is because higher leverage increases the risk borne by the entrepreneur. The total marginal cost of production includes compensation to the entrepreneur for production risk.

The aggregate accumulation constraint of entrepreneurs becomes

$$Q_t K_t^e + C_t^e = \frac{1}{X_t} Y_{ti}(\theta_t) + Q_t(1-\delta)K_{t-1}^e - r^b K_{t-1}^l - \pi_l \kappa K_{t-1} - W_t N_t, \tag{3.55}$$

Individual retailers may reset their price with probability $1-\rho$ in each period. Retailers are completely owned by the worker households, who hold diversified portfolios of shares in retailers. It follows that retailers discount expected future cashflows according to the worker households' common stochastic discount factor.

Retailers' nominal price setting decisions are subject to a quadratic adjustment cost following Rotemberg (1982). The adjustment costs incurred by firm j in period t are given by

$$\frac{\phi^\Pi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 Y_t. \tag{3.56}$$

Retailer j determines a state contingent path for prices $\{P_t(j)\}_{t=0}^\infty$ that maximises the discounted value of nominal profits χ:

$$\max_{\{P_t(j)\}_{t=0}^\infty} \mathbb{E}_t \sum_{s=0}^\infty [\mathrm{sdf}]_{t,t+s}^h \chi_{t+s}, \tag{3.57}$$

where

$$\chi_t = P_t(j)Y_t(j) - [\mathrm{mc}]_t Y_t(j) P_t - \frac{\phi^\Pi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 Y_t P_t, \tag{3.58}$$

And $[\mathrm{sdf}]_{t,t+s}^h$ is the stochastic discount factor applied by households in period t for income received in period $t+s$.

The resulting New Keynesian Phillips Curve in symmetric equilibrium is described as follows:

$$(1-\epsilon) + \epsilon[\mathrm{mc}]_t - \phi^\Pi \Pi_t(\Pi_t - 1) + \phi^\Pi \beta \mathbb{E}_t\left[\left(\frac{\lambda_{t+1}^h}{\lambda_t^h}\right)\frac{Y_{t+1}}{Y_t}\Pi_{t+1}(\Pi_{t+1}-1)\right] = 0. \tag{3.59}$$

We close the model with a Taylor rule that responds solely to current inflation. Let R_t^n denote the gross nominal interest rate on household deposits D. The Fisher relation is described as follows,

$$R_t = R_t^n / \Pi_t. \tag{3.60}$$

The monetary rule can be written the following way,

$$R_t^n \beta = \psi_\Pi \Pi_t. \tag{3.61}$$

3.D EQUILIBRIUM CONDITIONS

3.D.1 THE FLEXIBLE PRICE MODEL

Definition 3.1 *A* decentralised competitive equilibrium *specifies the set of state-contingent paths for* $\{Y_t, K_t, K_t^e, K_t^b, D_t, N_t, C_t, C_t^e, I_t, R_t, Q_t, r_t^b, r_t(\theta_2, \emptyset), W_t, \tau_{Nt}\}_{t=0}^{\infty}$ *which satisfies the following system of equations:*

WORKER HOUSEHOLDS

$$D_t + C_t = R_t D_{t-1} + W_t N_t \tag{3.39}$$

$$-\frac{U_2(C_t, N_t)}{W_t} = U_1(C_t, N_t) \tag{3.40}$$

$$U_1(C_t, N_t) = \beta \mathbb{E}_t R_{t+1} U_1(C_{t+1}, N_{t+1}) \tag{3.41}$$

FINANCIAL INTERMEDIARIES

$$D_t = Q_t K_t^b \tag{3.42}$$

$$D_{t-1} R_t = Q_t (1-\delta) K_{t-1}^b + r_t^b K_{t-1}^b \tag{3.43}$$

ENTREPRENEURS

$$Y_t = Z_t K_{t-1}^{\alpha} N_t^{1-\alpha} \tag{3.2}$$

$$K_t = K_t^b + K_t^e \tag{3.3}$$

$$Q_t K_t^e + C_t^e = Y_t + Q_t(1-\delta)K_{t-1}^e - r_t^b K_{t-1}^b - \pi_1 \kappa K_{t-1} - W_t N_t \tag{3.4}$$

$$\tau_{Nt} = \frac{(1-\alpha)Y_t - W_t N_t}{(1-\alpha)Y_t} \tag{3.17}$$

$$C_t^e = \frac{1-\beta^e}{\beta^e} Q_t K_t^e \tag{3.23}$$

$$\frac{N_t}{K_{t-1}} = \frac{1-\alpha}{\alpha} \frac{r_t^b + \pi_l \kappa}{W_t} \tag{3.24}$$

$$L_t = \frac{Y_t}{R_t Q_{t-1} K_{t-1}^e} \tag{3.29}$$

$$L_t = \frac{(\pi_2 + \pi_1 \eta)\tau_{Nt}}{[\pi_2 \xi_t - \tau_{Nt}][\pi_1 \eta \xi_t + \tau_{Nt}]} \tag{3.30}$$

$$\frac{K_{t-1}^b}{K_{t-1}^e} = \frac{R_t Q_{t-1} \left[\dfrac{\pi_1 (1-\eta)\tau_{Nti}}{\pi_1 \eta \xi_t + \tau_{Nti}} \right] + \pi_1 \kappa}{[r_{ti}(\theta_2, \emptyset) - r_t^b] - \pi_1 \kappa} \tag{3.31}$$

Capital producers and aggregate demand

$$Q_t = 1 + \phi \left(\frac{I_t}{K_t} - \delta \right) \tag{3.45}$$

$$Y_t = C_t + C_t^e + Q_t I_t + \pi_1 \kappa K_{t-1} \tag{3.46}$$

3.D.2 The sticky price model

Definition 3.2 *A* decentralised competitive equilibrium *specifies the set of state-contingent paths for* $\{Y_t, K_t, K_t^e, K_t^b, D_t, N_t, C_t, C_t^e, I_t, R_t, Q_t, r_t^b, r_t(\theta_2, \emptyset), W_t, \tau_{Nt}, X_t, \Pi_t, R_t^n\}_{t=0}^{\infty}$ *which satisfies the following system of equations:*

Worker households

$$D_t + C_t = R_t D_{t-1} + W_t N_t + \chi_t \tag{3.48}$$

$$-\frac{U_2(C_t, N_t)}{W_t} = U_1(C_t, N_t) \tag{3.40}$$

$$U_1(C_t, N_t) = \beta \mathbb{E}_t R_{t+1} U_1(C_{t+1}, N_{t+1}) \tag{3.41}$$

FINANCIAL INTERMEDIARIES

$$D_t = Q_t K_t^b \tag{3.42}$$

$$D_{t-1} R_t = Q_t (1-\delta) K_{t-1}^b + r_t^b K_{t-1}^b \tag{3.43}$$

ENTREPRENEURS

$$Y_t = Z_t K_{t-1}^{\alpha} N_t^{1-\alpha} \tag{3.2}$$

$$K_t = K_t^b + K_t^e \tag{3.3}$$

$$Q_t K_t^e + C_t^e = \frac{Y_t}{X_t} + Q_t(1-\delta) K_{t-1}^e - r_t^b K_{t-1}^b - \pi_1 \kappa K_{t-1} - W_t N_t \tag{3.55}$$

$$\tau_{Nt} = \frac{(1-\alpha) Y_t - X_t W_t N_t}{(1-\alpha) Y_t} \tag{3.53}$$

$$C_t^e = \frac{1-\beta^e}{\beta^e} Q_t K_t^e \tag{3.23}$$

$$\frac{N_t}{K_{t-1}} = \frac{1-\alpha}{\alpha} \frac{r_t^b + \pi_l \kappa}{W_t} \tag{3.24}$$

$$L_t = \frac{Y_t / X_t}{R_t Q_{t-1} K_{t-1}^e} \tag{3.33}$$

$$L_t = \frac{(\pi_2 + \pi_1 \eta) \tau_{Nt}}{[\pi_2 \xi_t - \tau_{Nt}][\pi_1 \eta \xi_t + \tau_{Nt}]} \tag{3.30}$$

$$\frac{K_{t-1}^b}{K_{t-1}^e} = \frac{R_t Q_{t-1} \left[\dfrac{\pi_1 (1-\eta) \tau_{Nti}}{\pi_1 \eta \xi_t + \tau_{Nti}} \right] + \pi_1 \kappa}{[r_{ti}(\theta_2, \emptyset) - r_t^b] - \pi_1 \kappa} \tag{3.31}$$

CAPITAL PRODUCERS AND AGGREGATE DEMAND

$$Q_t = 1 + \phi \left(\frac{I_t}{K_t} - \delta \right) \tag{3.45}$$

$$Y_t = C_t + C_t^e + Q_t I_t + \pi_1 \kappa K_{t-1} \tag{3.46}$$

$$(1-\epsilon) + \epsilon [\mathrm{mc}]_t - \phi^{\Pi} \Pi_t (\Pi_t - 1) + \phi^{\Pi} \beta \mathbb{E}_t \left[\left(\frac{\lambda_{t+1}^h}{\lambda_t^h} \right) \frac{Y_{t+1}}{Y_t} \Pi_{t+1} (\Pi_{t+1} - 1) \right] = 0. \tag{3.59}$$

$$R_t = R_t^n / \Pi_t \tag{3.60}$$

$$R_t^n \beta = \psi_\Pi \Pi_t, \tag{3.61}$$

3.E Useful derivations

3.E.1 Derivation of equations (3.19 , 3.20, 3.21)

Expanding the entrepreneurs' first order conditions for labour demand (3.7) yields

$$\begin{aligned}
0 &= \mathrm{P}(\theta_1,\sigma_1)\lambda_{ti}^e(\theta_1,\sigma_1)Y_{Nti}(\theta_1) + \mathrm{P}(\theta_1,\sigma_1)\lambda_{ti}^e(\theta_1,\sigma_2)Y_{Nti}(\theta_1) + \mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset)Y_{Nti}(\theta_2)\\
&\quad - \mathbb{E}_t\lambda_{ti}^e(\theta_{ti},\sigma_{ti})W_t\\
&= \mathrm{P}(\theta_1,\sigma_1)\lambda_{ti}^e(\theta_1,\sigma_1)Y_{Nti}(\theta_1) + \mathrm{P}(\theta_1,\sigma_1)\lambda_{ti}^e(\theta_1,\sigma_2)Y_{Nti}(\theta_1) + \mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset)Y_{Nti}(\theta_2)\\
&\quad + \mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset)Y_{Nti}(\theta_1) - \mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset)Y_{Nti}(\theta_1) - \mathbb{E}_t\lambda_{ti}^e(\theta_{ti},\sigma_{ti})W_t\\
&= \mathbb{E}_t\lambda_{ti}^e(\theta_{ti},\sigma_{ti})\theta_1 Y_{Nti} + \mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset)Y_{Nti}(\theta_2-\theta_1) - \mathbb{E}_t\lambda_{ti}^e(\theta_{ti},\sigma_{ti})W_t\\
&= \lambda_{ti}^e(\theta_1,\sigma_1)(\theta_1 Y_{Nti} - W_t) + \mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset)Y_{Nti}\xi_t\\
&= -(W_t - \theta_1 Y_{Nti}) + \mathrm{P}(\theta_2,\emptyset)\frac{\lambda_{ti}^e(\theta_2,\emptyset)}{\lambda_{ti}^e(\theta_1,\sigma_1)}Y_{Nti}\xi_t
\end{aligned}$$

$$\begin{aligned}
\frac{\lambda_{ti}^e(\theta_2,\emptyset)}{\lambda_{ti}^e(\theta_1,\sigma_1)} &= \frac{1}{\mathrm{P}(\theta_2,\emptyset)\xi_t}\frac{W_t-\theta_1 Y_{Nti}}{Y_{Nti}}\\
&= \frac{1}{\mathrm{P}(\theta_2,\emptyset)\xi_t}\frac{W_t-[\theta_1+\mathrm{P}(\theta_2,\emptyset)\xi_t-\mathrm{P}(\theta_2,\emptyset)\xi_t]Y_{Nti}}{Y_{Nti}}\\
&= \frac{1}{\mathrm{P}(\theta_2,\emptyset)\xi_t}\frac{W_t-[1-\mathrm{P}(\theta_2,\emptyset)\xi_t]Y_{Nti}}{Y_{Nti}}\\
\frac{\lambda_{ti}^e(\theta_2,\emptyset)}{\lambda_{ti}^e(\theta_1,\sigma_1)} &= 1-\frac{1}{\mathrm{P}(\theta_2,\emptyset)\xi_t}\frac{Y_{Nti}-W_t}{Y_{Nti}} = (3.20\ \text{RHS}).
\end{aligned}$$

Now, re-write (3.16) as follows:

$$\mathrm{P}(\theta_2,\emptyset)\lambda_{ti}^e(\theta_2,\emptyset) = [\mathrm{P}(\theta_2,\emptyset)+\mathrm{P}(\theta_1,\theta_2)]\lambda_{ti}^e(\theta_1,\sigma_1) - \mathrm{P}(\theta_1,\theta_2)\lambda_{ti}^e(\theta_1,\theta_2),$$

and divide through by $\lambda_{ti}^e(\theta_1,\sigma_1)$:

$$\mathrm{P}(\theta_2,\emptyset)\frac{\lambda_{ti}^e(\theta_2,\emptyset)}{\lambda_{ti}^e(\theta_1,\sigma_1)} = [\mathrm{P}(\theta_2,\emptyset)+\mathrm{P}(\theta_1,\theta_2)] - \mathrm{P}(\theta_1,\theta_2)\frac{\lambda_{ti}^e(\theta_1,\theta_2)}{\lambda_{ti}^e(\theta_1,\sigma_1)},$$

Rearranging and substituting in (3.20) yields

$$\begin{aligned}
\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\sigma_1)} &= \frac{\mathrm{P}(\theta_2,\emptyset)+\mathrm{P}(\theta_1,\theta_2)}{\mathrm{P}(\theta_1,\theta_2)} - \frac{\mathrm{P}(\theta_2,\emptyset)}{\mathrm{P}(\theta_1,\theta_2)}\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)} \\
&= \frac{\mathrm{P}(\theta_2,\emptyset)+\mathrm{P}(\theta_1,\theta_2)}{\mathrm{P}(\theta_1,\theta_2)} - \frac{\mathrm{P}(\theta_2,\emptyset)}{\mathrm{P}(\theta_1,\theta_2)}\left[1-\frac{\tau_{Nti}}{\mathrm{P}(\theta_2,\emptyset)\xi_t}\right] \\
&= \frac{\mathrm{P}(\theta_2,\emptyset)+\mathrm{P}(\theta_1,\theta_2)}{\mathrm{P}(\theta_1,\theta_2)} - \frac{\mathrm{P}(\theta_2,\emptyset)}{\mathrm{P}(\theta_1,\theta_2)} + \frac{\mathrm{P}(\theta_2,\emptyset)}{\mathrm{P}(\theta_1,\theta_2)}\frac{\tau_{Nti}}{\mathrm{P}(\theta_2,\emptyset)\xi_t} \\
\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\sigma_1)} &= 1+\frac{\tau_{Nti}}{\mathrm{P}(\theta_1,\theta_2)\xi_t} = (3.19\ \text{RHS}).
\end{aligned}$$

Now we consider equation (3.21). Substituting (3.14) into (3.8) yields

$$\begin{aligned}
0 &= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})[Y_{Kti}(\theta_{ti}) - r_{ti}(\theta_{ti},\sigma_{ti})] + \lambda^e_{ti}(\theta_1,\sigma_1)\left[\mathbb{E}_t\hat{r}_t(\theta_{ti},\sigma_{ti}) - r^d_t - \pi_1\kappa\right] \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] + \lambda^e_{ti}(\theta_1,\sigma_1)\left[\mathbb{E}_t\hat{r}_t(\theta_{ti},\sigma_{ti})\right] \\
&\quad - \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})r_{ti}(\theta_{ti},\sigma_{ti}) \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\quad + \lambda^e_{ti}(\theta_1,\sigma_1)[\mathrm{P}(\theta_1,\theta_2)r_{ti}(\theta_1,\theta_2)+\mathrm{P}(\theta_2,\emptyset)r_{ti}(\theta_2,\emptyset)] \\
&\quad - [\mathrm{P}(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2)r_{ti}(\theta_1,\theta_2)+\mathrm{P}(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)r_{ti}(\theta_2,\emptyset)] \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\quad + \lambda^e_{ti}(\theta_1,\sigma_1)[\mathrm{P}(\theta_1,\theta_2)+\mathrm{P}(\theta_2,\emptyset)]r_{ti}(\theta_2,\emptyset) \\
&\quad - [\mathrm{P}(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2)+\mathrm{P}(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)]r_{ti}(\theta_1,\theta_2) \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\quad + \{[\mathrm{P}(\theta_1,\theta_2)+\mathrm{P}(\theta_2,\emptyset)]\lambda^e_{ti}(\theta_1,\sigma_1) - \mathrm{P}(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2) - \mathrm{P}(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)\}r_{ti}(\theta_1,\theta_2) \\
&= \lambda^e_{ti}(\theta_1,\sigma_1)\theta_1 Y_{Kti} + \mathrm{P}(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Kti}\xi_t - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\qquad \text{(by (3.16))} \\
&= \theta_1 Y_{Kti} + \mathrm{P}(\theta_2,\emptyset)\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)}Y_{Kti}\xi_t - \left[r^d_t+\pi_1\kappa\right] \\
&= \theta_1 Y_{Kti} + \left[1-\frac{\tau_{Nti}}{\mathrm{P}(\theta_2,\emptyset)\xi_t}\right]\mathrm{P}(\theta_2,\emptyset)Y_{Kti}\xi_t - \left[r^d_t+\pi_1\kappa\right] \qquad \text{(by (3.20))}
\end{aligned}$$

$$\tau_{Nti} = \frac{Y_{Kti}\left[\theta_1 + \mathrm{P}(\theta_2, \emptyset)\xi_t\right] - r_t^d - \pi_1\kappa}{Y_{Kti}}$$

$$\tau_{Nti} = \frac{Y_{Kti} - r_t^d - \pi_1\kappa}{Y_{Kti}} = (3.21\ \text{RHS})$$

3.E.2 DERIVATION OF EQUATION 3.23

First, note that the entrepreneurs' preferences exhibit intertemporal homotheticity, and that projects are constant-returns-to-scale. The combination of these facts results in the consequence that both the choices of each entrepreneur and their returns and risks are scalable in K^e_{ti}. For simplicity, we'll re-write the accumulation constraint as

$$Q_t K^e_{t+1i} = R^e_{ti} K^e_{ti} - C^e_{ti},$$

where

$$R^e_{ti} = \frac{Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K^e_{ti} - K^l_{ti} r_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti}}{K^e_{ti}}$$

in order to capture the scalability of the entrepreneur's problem. R^e_{ti} is an idiosyncratic shock realised before the consumption decision is taken. In order to describe the entrepreneur's consumption decision, we'll re-write their problem as a Bellman equation formulated after the realisation of R^e_{ti}. We'll denote $W_{ti} = R^e_{ti} K^e_{ti}$, which is the resources the entrepreneur has available when they make their consumption decision in period t.

$$V(W_{ti}) = \max_{C^e_{ti}} \log C^e_{ti} + \beta^e \mathbb{E}_t V(W_{t+1i}),$$

subject to

$$W_{t+1i} = \frac{R^e_{t+1i}}{Q_t}(W_{ti} - C^e_{ti}).$$

Which we can re-write as

$$V(W_{ti}) = \max_{C^e_{ti}} \log C^e_{ti} + \beta^e \mathbb{E}_t V\left(\frac{R^e_{t+1i}}{Q_t}(W_{ti} - C^e_{ti})\right).$$

Dropping the subscript i, the first order condition for C^e is

$$\frac{1}{C^e_t} = \beta^e \mathbb{E} \frac{R^e_{t+1}}{Q_t} V'\left(\frac{R^e_{t+1i}}{Q_t}(W_{ti} - C^e_{ti})\right)$$

We proceed by guessing a particular functional form for V, and verifying that this functional form satisfies the conditions above.

Let

$$\hat{V}(W) = \frac{1}{1-\beta^e}\log W + k,$$

where k is some constant.

Assuming, that $\hat{V}$ is the correct value function, we substitute it into the first order condition to solve for consumption:

$$\frac{1}{C_t^e} = \frac{\beta^e}{1-\beta^e}\frac{1}{(W_t - C_t^e)}$$

which rearranges to yield

$$C_t^e = (1-\beta^e)W_{ti}$$

We now verify our guess value function by substituting it into the entrepreneur's Bellman equation,

$$V(W_{ti}) = \max_{C_{ti}^e} \log C_{ti}^e + \beta^e \mathbb{E}_t \frac{1}{1-\beta^e}\log\left(\frac{R_{t+1i}^e}{Q_t}(W_{ti} - C_{ti}^e)\right) + \beta^e k.$$

And substitute in our optimal consumption decision,

$$\begin{aligned}
V(W_{ti}) &= \log[(1-\beta^e)W_{ti}] + \beta^e \mathbb{E}_t \frac{1}{1-\beta^e}\log\left(\frac{R_{t+1i}^e}{Q_t}(W_{ti} - (1-\beta^e)W_{ti})\right) + \beta^e k \\
&= \log W_{ti} + \log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}(\beta^e W_{ti})\right) + \beta^e k \\
&= \log W_{ti} + \log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\log W_{ti} + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}\beta^e\right) + \beta^e k \\
&= \frac{1}{1-\beta^e}\log W_{ti} + \log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}\beta^e\right) + \beta^e k \\
V(W_{ti}) &= \frac{1}{1-\beta^e}\log W_{ti} + k,
\end{aligned}$$

where

$$k = \frac{1}{1-\beta^e}\left[\log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}\beta^e\right)\right].$$

This confirms that our guess value function $\hat{V}$ satisfies the entrepreneur's problem, and that $C_t^e = (1-\beta^e)W_{ti}$ is the optimal consumption decision for our entrepreneur.

We can now substitute this back into the entrepreneurs' accumulation constraint (3.4) to obtain

$$C^e_{ti} = (1-\beta^e)[Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K^e_{ti} - K^l_{ti}r_{ti}(\theta_{ti},\sigma_{ti}) - W_tN_{ti}].$$

It follows that

$$Q_tK^e_{t+1i} = \beta^e[Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K^e_{ti} - K^l_{ti}r_{ti}(\theta_{ti},\sigma_{ti}) - W_tN_{ti}],$$

and

$$C^e_{ti} = \frac{1-\beta^e}{\beta^e}Q_tK^e_{t+1i}. \tag{3.23}$$

DERIVATION OF EQUATIONS 3.28, 3.30 AND 3.31

First, substitute the national income equation (3.22) into the marginal rates of substitution conditions (3.26) and (3.26) to eliminate W_tN_t,

$$\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)} = \frac{(\theta_1-1+\tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}r_{ti}(\theta_1,\sigma_1) + r^dK^l_t + \pi_l\kappa K_t}{(\theta_1-1+\tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}r_{ti}(\theta_2,\emptyset) + r^dK^l_t + \pi_l\kappa K_t} \tag{3.62}$$

$$\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\theta_1)} = \frac{(\theta_1-1+\tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}r_{ti}(\theta_1,\sigma_1) + r^dK^l_t + \pi_l\kappa K_t}{(\theta_2-1+\tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}r_{ti}(\theta_2,\emptyset) + r^dK^l_t + \pi_l\kappa K_t} \tag{3.63}$$

Note that $\pi_1\theta_1 + \pi_2\theta_2 = 1$, which implies $\theta_1 = 1 - \pi_2\xi_t$, and $\theta_2 = 1 + \pi_1\xi_t$.

Rather than working with specific interest rates $r(\cdot)$ and project returns θ, it will be helpful to re-write these conditions in terms of project risk $\xi_t = \theta_2 - \theta_1$, and risk sharing $[r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1)]$. From the financial intermediaries' participation constraint we have

$$K^l_{ti}[(\pi_2+\pi_1\eta)r_{ti}(\theta_2,\emptyset) + \pi_1(1-\eta)r_{ti}(\theta_1,\sigma_1)] = r^dK^l_t + \pi_l\kappa K_t$$

which allows us to re-write $r_{ti}(\theta_2,\emptyset), r_{ti}(\theta_1,\sigma_1)$ in terms of required returns and risk:

$$r_{ti}(\theta_2,\emptyset) = r^d + \pi_l\kappa\frac{K_t}{K^l_{ti}} + \pi_1(1-\eta)(r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1))$$

$$r_{ti}(\theta_1,\sigma_1) = r^d + \pi_l\kappa\frac{K_t}{K^l_{ti}} - (\pi_2+\pi_1\eta)(r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1))$$

which we can substitute back into (3.62,3.63) and rearrange to obtain

$$R_t Q_{t-1}\frac{K^e_{ti}}{K^l_{ti}} = (\pi_2\xi_t - \tau_{Nt})\frac{\bar{Y}_t}{K^l_{ti}} - \left[\frac{1}{1-\lambda^e_{ti}(\theta_1,\theta_2)/\lambda^e_{ti}(\theta_1,\theta_1)} - \pi_1(1-\eta)\right](r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1))$$

$$\begin{aligned} R_t Q_{t-1}\frac{K^e_{ti}}{K^l_{ti}} &= (\pi_2\xi_t - \tau_{Nt})\frac{\bar{Y}_t}{K^l_{ti}} + \frac{\lambda^e_{ti}(\theta_2,\emptyset)/\lambda^e_{ti}(\theta_1,\theta_1)}{1-\lambda^e_{ti}(\theta_2,\emptyset)/\lambda^e_{ti}(\theta_1,\theta_1)}\frac{\bar{Y}_t}{K^l_{ti}}\xi_t \\ &\quad - \left[\frac{1}{1-\lambda^e_{ti}(\theta_2,\emptyset)/\lambda^e_{ti}(\theta_1,\theta_1)} - \pi_1(1-\eta)\right](r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1)) \end{aligned}$$

And we equate the right hand sides to solve

$$\frac{\bar{Y}_t\xi_t}{K^l_{ti}(r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1))} = \left[\frac{\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)}/\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\theta_1)} - 1}{\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)} - 1}\right]$$

Note that $\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)} > 1$, and $\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\theta_1)} < 1$, making the right hand side strictly greater than 1. The left hand side is the ratio of productive risk to the possible amount of risk sharing following loan restructuring. The equation shows how an increase in the ratio of productive risk to risk sharing increases the entrepreneurs' marginal rates of substitution for consumption across idiosyncratic states. These marginal rates of substitution in turn determine entrepreneurs' precautionary reductions in wage and capital hiring compared with the first best efficient levels, through the wedges specified in equations (3.19) and (3.20). Substituting these factor wedges in place of the marginal rates of substitution yields

$$K^b_{t-1i}[r_{ti}(\theta_2,\emptyset) - r_{ti}(\theta_1,\sigma_1)] = \bar{Y}_{ti}\left[\frac{\pi_2\xi_t - \tau_{Nti}}{\pi_2+\pi_1\eta}\right] \tag{3.28}$$

Now we can use this solution to solve for the the leverage ratio

$$\frac{\bar{Y}_{ti}}{R_t Q_{t-1}K^e_{t-1i}} = \frac{(\pi_2+\pi_1\eta)\tau_{Nti}}{[\pi_2\xi_t - \tau_{Nti}][\pi_1\eta\xi_t + \tau_{Nti}]}, \qquad \text{(= (3.30 RHS))}$$

and also the ratio of household's to entrepreneurs' capital in terms of the loan coupon interest rate spread $[r_{ti}(\theta_2,\emptyset) - r^b_t]$.

$$\frac{K^b_{t-1i}}{K^e_{t-1i}}([r_{ti}(\theta_2,\emptyset) - r^b_t] - \pi_1\kappa) = R_t Q_{t-1}\left[\frac{\pi_1(1-\eta)\tau_{Nti}}{\pi_1\eta\xi_t + \tau_{Nti}}\right] + \pi_1\kappa \tag{3.31}$$

3.F Model dynamics

Figure 3.6: Log-linearised model dynamics: Total factor productivity shock. Flexible prices (black,– –), sticky prices (red,—).

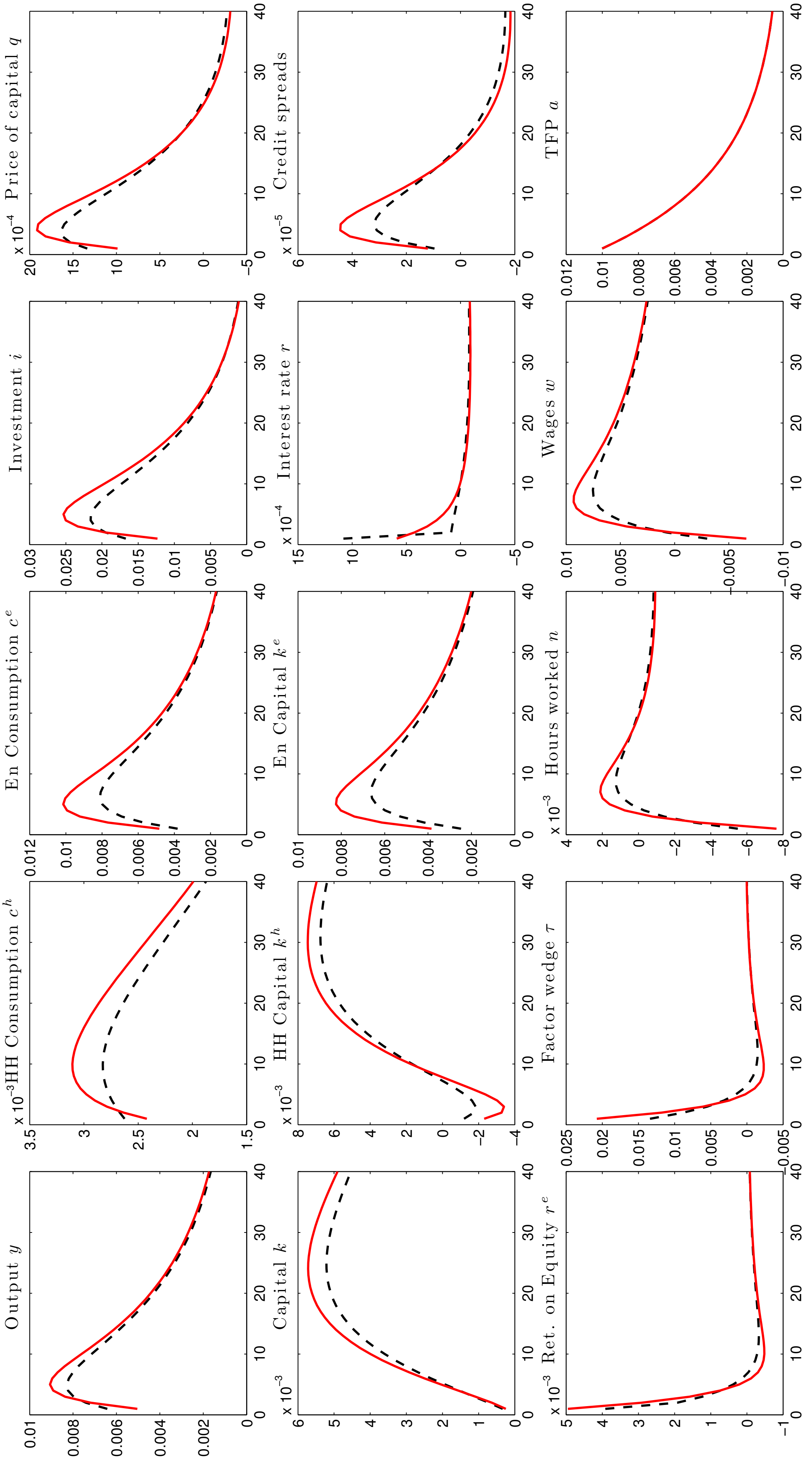
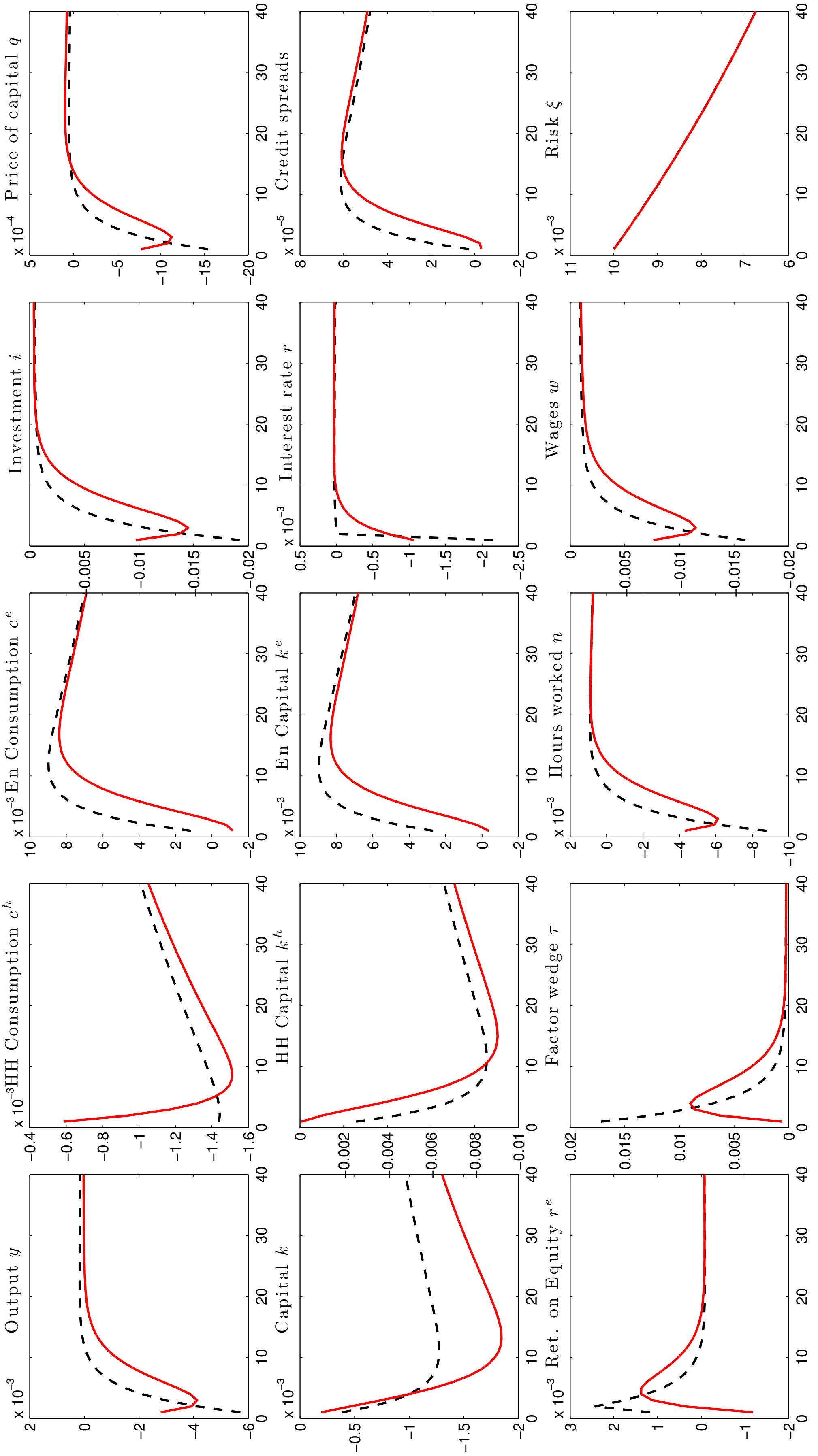


Figure 3.7: Log-linearised model dynamics: Risk shock. Flexible prices (black,– –), sticky prices (red,—).



CHAPTER 4

SYSTEMIC RISK MARKETS AND THE FINANCIAL ACCELERATOR

This chapter is co-authored with Charles Nolan.[1]

This paper considers the importance of insurance markets for commonly observed systemic risks in a setting where entrepreneur borrowers' individual specific risks are private information. Within a new framework allowing risk averse entrepreneur borrowers in a costly state verification setting, the key finding of this paper is that when systemic risk markets are open, the decentralised competitive equilibrium is typically constrained inefficient, with excessive volatility in leverage. This contrasts with earlier studies with risk neutral entrepreneurs which have suggested that trade in systemic risk markets would tend to dampen the financial accelerator and collateral amplification mechanisms.

[1]Professor of Economics, University of Glasgow. Charles.Nolan@glasgow.ac.uk. The authors would like to thank Costas Azariades, Martin Eichenbaum, Campbell Leith and Bruce Mizrach in addition to conference and seminar participants at the 2015 Society for Nonlinear Dynamics Annual Conference, the 2015 Centre for Economic Growth and Policy Conference and the 2015 Sheffield Workshop in Macroeconomics for helpful comments and discussions. All errors are our own.

INTRODUCTION

Equity earns high returns in bad times—after its price has fallen. Firms *should* buy insurance against downturns and financial stress. This would transfer wealth to them precisely when it is likely to earn a high return. This is a well-known prediction of the two workhorse models of financial macroeconomics described by Bernanke, Gertler, and Gilchrist (1999) and Kiyotaki and Moore (1997). Carlstrom, Fuerst, and Paustian (2014), Krishnamurthy (2003) and Nikolov (2014) show in these models respectively that this privately optimal insurance trade should eliminate the financial accelerator and credit cycle mechanisms which amplify and propagate business cycle shocks within these models.

But we don't appear to see this trade in practise. Confronting the macroeconomic data, business cycle models that include financial sector amplification seem to perform better empirically.[2] From the microeconomic perspective, the more typical story we hear from the period preceding the Global Financial Crisis of 2007-08 is that financial and non-financial firms (with the notable example of AIG) had overly exposed themselves to the state of the US economy, housing market and financial system, effectively *selling* insurance against systemic risk.

In this paper, we show that this observation is consistent with a broad class of financial macroeconomic models, of which the financial accelerator and collateral amplification models described by Bernanke, Gertler, and Gilchrist (1999) and Kiyotaki and Moore (1997) are special cases. The combination of both risk averse firm insiders and monotonic insider compensation schemes limits the extent to which competitive trade in systemic risk insurance results in the sacrifice of consumption insurance in return for the stabilisation of factor price wedges that amplify business cycle volatility in the model.

We then present a costly state verification model with risk averse entrepreneurs and optimal debt contracts.[3] In our model, competitive insurance allocations result in full consumption insurance, at the cost of high volatility in factor market distortions. We show that the resulting competitive equilibrium allocations are constrained inefficient,

[2] See for example Christiano, Motto, and Rostagno (2014), Nolan and Thoenissen (2009), Jermann and Quadrini (2012) and Del Negro, Giannoni, and Schorfheide (2014).

[3] More specifically, we allow for stochastic auditing schemes as in Border and Sobel (1987) and Mookherjee and Png (1989), while still recovering standard debt contracts with deterministic audit schemes as optimal.

and further that closing the market for systemic risk insurance could result in a Pareto welfare gain.

THE EXTERNALITY ASSOCIATED WITH SYSTEMIC RISK INSURANCE

At the heart of all popular models in financial macroeconomics is an equilibrium condition relating wedges between factor marginal products and factor prices τ_t to the financial leverage of entrepreneurs l_t and some measure of financial stress ξ_t,

$$\tau_t = \mathcal{T}(l_t, \xi_t), \qquad \mathcal{T}_1, \mathcal{T}_2 > 0.$$[4]

At business cycle frequencies, leverage dynamics can be volatile, and are heavily influenced by the extent of trade in systemic risk (or alternatively by government enforced transfers) between worker households and entrepreneurs.

Consider a systemic risk insurance contract between worker household A and an individual entrepreneur. Who insures whom against the various sources of systemic risk is ambiguous at this stage. On the one hand, risk averse worker j will likely suffer a greater increase in consumption marginal utility during a downturn, suggesting that systemic risk insurance should result in transfers from the entrepreneur to worker household j during a downturn. On the other hand, in the downturn the return to entrepreneurs wealth will be relatively high. This is equivalent to saying that the price of the entrepreneurs' consumption relative to worker household consumption is low, and suggests suggesting that in equilibrium it may be privately optimal to write a contract that transfers wealth from the worker household j to the entrepreneur during downturns.

Any transfer from the individual entrepreneur to worker household j will increase the leverage of the entrepreneur. This increase in leverage will increase the factor market wedges, reducing the return to the capital holdings and the wage offered for the labour endowment of other worker households. In the absence of Coasean bargains between worker households, or Kilenthong and Townsend (2014) mechanisms,[5] the privately opti-

[4]In the models described by Bernanke, Gertler, and Gilchrist (1999) and Kiyotaki and Moore (1997), the wedge τ_t^k is the difference between the marginal product of capital and interest rates earned on household savings. In our model (and also in the model described by Jermann and Quadrini, 2012) financial stress also affects labour markets through τ_t^n, a wedge between real wages and the marginal product of labour. In all of these models, the factor market wedge(s) will be larger when leverage is high or during periods of financial stress.

[5]Kilenthong and Townsend (2014) show that in generalised incomplete market settings, constrained

mal systemic risk insurance trade between worker household j and the entrepreneur may impose a net social cost.

The models described by Bernanke, Gertler, and Gilchrist (1999) and Kiyotaki and Moore (1997) predict that systemic risk insurance trade typically resulted in transfers from worker households to entrepreneurs during downturns. This has the effect of dampening volatility in the factor market wedges. We show that this result does not generalise. In our model, systemic risk insurance will often result in transfers from entrepreneurs to workers during downturns, exacerbating volatility in the factor market wedges.

Implications for stabilisation policy

The emerging empirical consensus is that financial frictions play an important rule in amplifying and propagating business cycles. In light of the stark results of Carlstrom, Fuerst, and Paustian (2014), Krishnamurthy (2003) and Nikolov (2014), suggesting that trade in systemic risk markets should eliminate financial amplification, one plausible interpretation of the empirical evidence is that systemic risk markets are closed.

If systemic risk markets are indeed closed, then there may be a role for policy in replicating the flows of funds between worker households and entrepreneurs that would have occurred if systemic risk markets were open. The policy interventions studied by Christiano and Ikeda (2011) resemble the insurance payments implicit in Carlstrom, Fuerst, and Paustian (2014), Krishnamurthy (2003) and Nikolov (2014). The studies of Koenig (2011) and Sheedy (2014) essentially attempt to restore these same real transfers through nominal output targeting monetary rules that allow high inflation during downturns.

But this transfer policy would be ineffective if systemic risk markets were actually open. We have no microeconomic evidence or theoretical justification underlying the presumption that systemic risk markets are closed. We cannot appeal to information asymmetries. Further, we know that in some cases, systemic risk markets are clearly open. Commodity futures markets are essentially systemic risk markets, and these markets are widely used by firms exposed to commodity price risk.

---

efficient allocations are obtainable through decentralised market trade if markets can be organised in such a way that the rights to participate in an individual market can be traded ex ante. Given this and the ability to write long term contracts and use stochastic mechanisms, constrained efficiency can be restored without government intervention.

In our model, the introduction of systemic risk markets does not eliminate the financial amplification mechanism. Consequently, the predictions of our model do not necessarily favour the presumption that systemic risk markets are closed. There is a potential role for policy to play in dampening the volatility of distortions arising from the financial friction.

ROADMAP

Section 1 describes how systemic risk markets address the twin goals of consumption insurance and stabilisation of factor market wedges.Section 2 introduces the model, focusing on the entrepreneurs' problem which is new to the literature. Section 3 presents a comparison between equilibrium allocations under the assumptions of open and closed systemic risk insurance markets.

4.1 INSURANCE AND STABILISATION

With unrestricted trade in securities contingent on commonly observed states revealed at the beginning of period t, the competitive equilibrium evolution of the marginal values of wealth of worker household j and entrepreneur i are characterised as follows:

$$\frac{\beta^h V^{h\prime}(W_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} = \frac{\beta^e V^{e\prime}(W_t^{ei})}{U^{e\prime}(C_{t-1}^{ei})}$$

With unrestricted trade in insurance, payments following the revelation of the commonly observed state will flow toward agents who would otherwise have a relatively high marginal value of wealth.

For each agent, the marginal value of wealth is the expectation of returns to wealth within the current period and the marginal utility of consumption at the termination of the period. Entrepreneurs' return to wealth R^e is subject to risk within the period while wealth held by the worker households earns the risk free rate R^h. We can re-write the marginal values of worker household j and entrepreneur i as follows

$$V^{h\prime}(W_t^{hj}) = R_t^f U^{h\prime}(C_t^{hj})$$

$$V^{e\prime}(W_t^{ei}) = \mathbb{E}_t[R_t^{ei} U^{e\prime}(C_t^{ei})]$$

In equilibrium, insurance payments will flow toward agents who otherwise would have relatively high marginal values, from agents who would otherwise have relatively low marginal values. What we're interested in is seeing how marginal values and consequently insurance flows are linked to the equity risk premium $\mathbb{E}_t[R_t^{ei}]/R_t^f$, which will tend to be large during periods of financial stress. If equilibrium insurance payments tend to flow toward entrepreneurs when the equity risk premium is high, then this insurance trade will tend to dampen volatility in the distortions associated with the financial friction, and consequently dampen business cycle volatility.

When risk averse entrepreneurs cannot costlessly pass through productive risk, the story changes. The entrepreneurs' marginal value remains

$$V^{e\prime}(W_t^{ei}) = \mathbb{E}_t[R_t^{ei} U^{e\prime}(C_t^{ei})]$$

Equilibrium in the insurance market implies

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} = \frac{\beta^e \mathbb{E}_t[R_t^{ei} U^{e\prime}(C_t^{ei})]}{R_t^f U^{e\prime}(C_{t-1}^{ei})} .$$

If entrepreneurs cannot costlessly pass through idiosyncratic production risk, then entrepreneurs' individual marginal utility of consumption will be low following high returns, and high following low returns. This negative covariance between marginal utility and returns implies that $\mathbb{E}_t[R_t^{ei} U^{e\prime}(C_t^{ei})] < \mathbb{E}_t[R_t^{ei}] \cdot \mathbb{E}_t[U^{e\prime}(C_t^{ei})]$. We can re-write the insurance market equilibrium condition as follows

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} \Bigg/ \frac{\beta^e \mathbb{E}_t[U^{e\prime}(C_t^{ei})]}{U^{e\prime}(C_{t-1}^{ei})} < \frac{\mathbb{E}_t[R_t^{ei}]}{R_t^f}. \tag{4.1}$$

The inequality described by equation 4.1 can be interpreted as a statement about the trade-off faced by the invisible hand as it seeks to achieve two sometimes conflicting goals. The left hand side is the gross deviation from full consumption insurance. The right hand side is the gross distortion in the capital market, and is a reflection of factor market distortions—namely the difference between savers' intertemporal marginal rates of substitution and the marginal product of capital, and in our model, the difference between the worker households' consumption-leisure marginal rates of substitution and their respective marginal labour product. Holding all else constant, a one-off transfer of wealth from worker households to entrepreneurs will raise the left hand side (increasing household marginal utility relative to entrepreneurs') and reduce the left hand side, as the increased entrepreneur wealth reduces leverage and factor market distortions. Under the first best efficient allocations, both left and right hand sides would be equal to one.

Considering specific examples in the following paragraphs will show how information asymmetries and other frictions affecting idiosyncratic risk sharing also have an impact on the sharing of systemic risks. Competitive trade in systemic risk insurance will not allow large deviations from full consumption insurance to counter volatility in the equity risk premium.

EXAMPLE 1: RISK NEUTRAL ENTREPRENEURS

We can think of the model of Bernanke, Gertler, and Gilchrist (1999) as the limit case as entrepreneurs' aversion to risk disappears. This helps us to reveal the intuition behind the result of Carlstrom, Fuerst, and Paustian (2014). With risk neutral entrepreneurs, entrepreneurs' marginal utilities are constant ($= U^{e\prime}$) and their marginal value can be described as follows

$$V^{e\prime}(W_t^{ei}) = U^{e\prime}\mathbb{E}_t[R_t^{ei}].$$

The equilibrium condition resulting from unrestricted trade in common shock insurance can be written as follows:

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} \Big/ \frac{\beta^e U^{e\prime}}{U^{e\prime}} = \frac{\mathbb{E}_t[R_t^{ei}]}{R_t^f} \tag{4.2}$$

Equation 4.2 describes a tight link between deviations from full consumption insurance and the equity risk premium. When the equity risk premium is large, worker households' marginal utilities must also be high relative to the full consumption insurance benchmark, indicating a transfer of resources from worker households to entrepreneurs. This flow of wealth to entrepreneurs during periods of high equity risk premia helps to dampen volatility in the equity risk premium, stabilising factor wedges and dampening business cycle volatility.

EXAMPLE 2: FULL PROJECT RISK SHARING

We can also consider the nature of optimal insurance flows when entrepreneurs are risk averse but their private return to wealth is not subject to idiosyncratic risk.[6] It could be either projects are not subject to idiosyncratic risk, or that this risk can be shared costlessly with financial intermediaries. There may still be financial constraints as a result of a limited commitment friction limiting the amount of external finance in order to ensure the entrepreneur does not wish to walk away from their commitments. Here, the entrepreneurs' marginal value is

$$V^{e\prime}(W_t^{ei}) = R_t^e U^{e\prime}(C_t^{ei}).$$

and equilibrium in the insurance market implies,

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} \Bigg/ \frac{\beta^e U^{e\prime}(C_t^{ei})}{U^{e\prime}(C_{t-1}^{ei})} = \frac{R_t^e}{R_t^f}. \tag{4.3}$$

As in the previous example, high equity risk premia during periods of financial stress are consistent with relatively high worker household marginal utilities and relatively low entrepreneur marginal utilities. This indicates that compared with the full consumption insurance counterfactual, insurance payments tend to flow toward entrepreneurs during periods of financial stress. The insurance trade is responding to the financial friction, and consumption insurance is traded off against fluctuations in the equity risk premium.

EXAMPLE 3: OUR MODEL

Examples 1 and 2 describe very different models, yet result in the consistent prediction that in competitive equilibrium, agents will suffer deviations from consumption risk insurance in order to stabilise costly factor market wedges. Both of these examples can be thought of as limit cases of the general model, and crucially, neither example permits idiosyncratic risk in entrepreneurs' marginal utilities.

In the model we describe in the following sections, entrepreneurs' marginal utility will be subject to idiosyncratic production risk. Entrepreneurs' demand for external fi-

[6] This could be the case if the financial friction was not based on private information but rather on limited commitment or enforceability constraints.

nance loans will be limited by imperfect risk sharing, which will result in a tradeoff between expected consumption and consumption risk. In equilibrium, the entrepreneurs' consumption Euler condition will bind with respect to the risk free interest rate.

It follows that

$$R_t^f \cdot \mathbb{E}_t[U^{e\prime}(C_t^{ei})] = \mathbb{E}_t[R_t^{ei} U^{e\prime}(C_t^{ei})].$$

Equilibrium in the insurance market implies

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} \Big/ \frac{\beta^e \mathbb{E}_t[U^{e\prime}(C_t^{ei})]}{U^{e\prime}(C_{t-1}^{ei})} = 1. \tag{4.4}$$

The competitive equilibrium is consistent with full consumption insurance. Unlike the previous two examples, increases in the equity risk premium do not correspond to a compensating departure from full consumption risk insurance. The insurance market does not respond to factor market distortions.

High returns to entrepreneurs correspond to high entrepreneurial risk, breaking the link between entrepreneurs' marginal values of wealth and the equity risk premium. The prospect of earning high equity returns during periods of financial stress does not necessarily encourage entrepreneurs to purchase the insurance against downturns—insurance which would recapitalise entrepreneurs during periods of financial stress and dampen the financial accelerator mechanism.

4.2 THE MODEL

The model has a similar structure to Bernanke, Gertler, and Gilchrist (1999). There exists a large population of worker households, who enjoy consumption and supply labour. Individual households' consumption, wealth and employment outcomes are perfectly insured in competitive markets within the worker household population. There is also a large population of risk averse entrepreneurs, who hire labour and borrow capital to augment their own capital wealth in productive risky projects. Individual project outcomes are initially private information to entrepreneurs, and this individual specific productive risk can be partially but not completely insured through state contingent capital loan contracts. Specifically, lender financial intermediaries have a costly and imperfect state verification technology, which allows them to investigate individual entrepreneurs' project outcomes following low reported income—which we interpret as default.

The optimal external finance contract

State-contingent loan contracts follow a variant of the costly state verification problem proposed by Townsend (1979), with the addition of possible disputes as described in Chapter 2.[7]

If both principal and agent were risk neutral and the audit technology were perfect, then the optimal external finance contract would inherit some of the properties highlighted in Border and Sobel (1987). In particular, large penalties (for misinforming the principal) or large rebates (for honest reporting) could be combined with a low probability of audit. That way, audit costs are reduced whilst incentive compatibility is ensured. That combination of penalty/reward/audit probability could be applied to any reported income level of the agent delivering, as result, a lot of risk-sharing. Such a contract does not resemble simple debt.

If the agent were risk averse, as in Mookherjee and Png (1989), large penalties following misinformation would remain desirable and high marginal utilities following penalties further provide an even greater incentive for truth telling, even when these penalties applied with low probability. On the other hand, low marginal utility in good states would dampen the positive incentive effects of large rebates following verified honest reports. However, stochastic audits would remain optimal: So long as large penalties can be levied for misinforming the principal, audits can be applied with low probability, reducing costs while maintaining incentives for truth telling even when contracts allow for a high degree of risk sharing. Indeed, optimal consumption profiles and audit probabilities may be non-monotonic in income! Those properties, along with the optimality of stochastic audits, would again mean that debt contracts are not optimal.

Like Mookherjee and Png (1989), we consider an environment where the principal is risk neutral and the agent is risk averse. However, we assume that the audit technology is imperfect; sometimes, the technology will incorrectly identify a low-income agent as a high-income agent. We call this a Type I error. The possibility of such errors further

[7] Chapter 2 showed that borrower risk aversion combined with the introduction of the possibility of disputed claims can result in the optimality of standard debt contracts which are not subject to the critiques of Border and Sobel (1987) and Mookherjee and Png (1989).

erodes the efficacy of penalties exacerbating the misreporting problem. Nevertheless, when the cost of audit is not too high, Chapter 2 showed that low-income agents would always wish to be audited. The possibility of error notwithstanding, being audited helps maintain consumption just when it is especially valuable; were they not audited they would be mistaken for high-income agents. At higher income levels, trying to risk-share requires a high probability of audits and entails costly audits and the prospect of dispute. At that point, it is preferable simply to pay interest plus principal. Such a contract bears close similarity to standard debt.

Much of the derivations in this section follow Chapter 3 closely.

4.2.1 TIMING

Within this paper, we assume that common shocks are revealed at the beginning of each period, and entrepreneurs' projects are undertaken following the revelation of the commonly observed state. This means that there is no overlap between common shock insurance contracts and external finance contracts. Table 4.1 presents a timeline for entrepreneur i in period t, reproduced for convenience from Chapter 3.

Table 4.1: Timeline for entrepreneur i, period t

Nature	Variables Determined	Description
$\mathbf{z}_t$		Common shocks revealed
	K_t^{ei}	Insurance payments made
	$N_t^i, K_t^{bi}, r_t^i(\theta_t^i, \sigma_t^i)$	Leverage, employment determined
θ_t^i		Individual projects completed
σ_t^i		Auditing, factor payments
	$C_t^{ei}, X_t^{ei}(\mathbf{z}_{t+1}), K_t^{pei}$	Consumption, insurance determined
$\mathbf{z}_{t+1}$		Common shocks revealed

4.2.2 ENTREPRENEURS

There exists a unit measure of entrepreneurs, indexed by i, who enjoy consumption with logarithmic utility,

$$\mathcal{U}^e_{ti} = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^{ej} U^e(C^e_{t+ji}), \tag{4.5}$$

where $u^e(C) = \log C$, and $\beta^e < \beta$, entrepreneurs are less patient than households.[8] Entrepreneurs undertake projects with binary risky outcomes. The individual output of entrepreneur i is

$$Y_{ti} = \theta_{ti} A_t K^{\alpha}_{t-1i} N^{1-\alpha}_{ti}, \tag{4.6}$$

where θ_{it} is an idiosyncratic shock drawn from $\theta_{ti} \in \{\theta_{1t}, \theta_{2t}\}$ where $\theta_{1t} < \theta_{2t}$. These two states occur with probabilities π_1 and π_2 respectively, and their expectation is equal to one, $\pi_1\theta_{1t}+\pi_2\theta_{2t} = 1$. Throughout our analysis, it will be useful to consider the difference between high and low project outcomes $\xi_t = \theta_{2t} - \theta_{1t}$ which we will allow to be time varying and following the law of motion $\log \xi_t = \rho_\xi \log \xi_t + \epsilon_{\xi t}$, where $\epsilon_{\xi t}$ is a white noise process with standard deviation σ_ξ. Denote the expectation of output for entrepreneur i in period t conditional upon A_t by $\bar{Y}_{ti} = A_t K^{\alpha}_{t-1i} N^{1-\alpha}_{ti}$. The variable A_t is an aggregate total factor productivity shock following the law of motion $\log A_t = \rho_A \log A_t + \epsilon_{At}$, where ϵ_{At} is a white noise process with standard deviation σ_A. Aggregate shocks are observable at the beginning of the period, whereas the unobservable shock is revealed to the entrepreneur at the end of the period. Capital employed by the entrepreneur is denoted in period t is K_{t-1i}, and N_{ti} is labour hired by the entrepreneur from the household sector.

At the beginning of each period, entrepreneurs borrow capital K^b_{ti} from financial intermediaries. Loan contracts specify the interest rate paid in good states, as well as the recovery rate returned to financial intermediaries in bad states. Capital inputs into entrepreneur i's project include the entrepreneur's initial capital holdings and further capital borrowed.

$$K_{ti} = K^e_{ti} + K^b_{ti}, \tag{4.7}$$

where K^e_{t-1} is the capital held by the entrepreneur at the beginning of period t. Entrepreneurs fund consumption and future capital holdings out of the sum of project rev-

[8]Entrepreneurs enjoy a greater return on savings than households in the model. Their reduced discount factor is required to ensure that the ratio of entrepreneurial and household wealth is constant in the steady state.

enues and current capital holdings, after repaying loans and paying workers' wages,

$$Q_t K^e_{ti} + C^e_{ti} = Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K^e_{t-1i} - K^b_{t-1i}\hat{r}_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti}. \tag{4.8}$$

where δ is the depreciation rate of capital. We attach the Lagrange multipliers $\lambda^e_{ti}(\theta_{ti}, \sigma_{ti})$ to each of the state contingent accumulation constraints. Note that capital rental payments $\hat{r}_{ti}(\theta_{ti}, \sigma_{ti})$ are contingent on the idiosyncratic shock θ_{ti} as well as any audit signal obtained by the financial intermediary, $\sigma_{ti} \in \{\sigma_1, \sigma_2\}$. Audit signals are distributed as follows: $P(\sigma_2|\theta_2) = 1, P(\sigma_2|\theta_1) = \eta, P(\sigma_1|\theta_1) = 1 - \eta$. Consequently, the unconditional probabilities of the three possible outcomes are as follows:

$$P(\theta_1, \sigma_1) = \pi_1(1-\eta), \qquad P(\theta_1, \sigma_2) = \pi_1\eta, \quad \text{and} \quad P(\theta_2, \emptyset) = \pi_2. \tag{4.9}$$

FINANCIAL CONTRACTS

Chapter 2 show that when audit costs are sufficiently low and auditing is imperfect, defaultable debt contracts with deterministic audit strategies are constrained efficient.

Assumption 4.1 *Contracts are only contingent on reports and audit signals within the current period.*

This restriction is referred to as the *anonymity* constraint. Once repayments on current period loans are made, entrepreneurs are considered to become anonymous, and their future actions in other markets cannot be used as evidence of past false reports.

We apply Theorem 2.2 from Chapter 2 to motivate standard debt contracts as optimal.

Theorem 2.2 *Let borrowing be taken as given $b = \hat{b}$. When type-I audit errors occur with positive probability ($\eta(\underline{\theta}) > 0$), there exists some strictly positive audit cost $\hat{\kappa}$ such that for all $\kappa < \hat{\kappa}$, standard debt contracts ($q(\underline{m}) = 1$) are efficient.*

As we discuss in Appendix 4.A, it is difficult to pin down audit costs in a way that produces the high credit spreads observed in the data, given the low historical probabilities of corporate default. Microeconomic estimates of direct bankruptcy costs as a share of firms' assets (what would be interpreted as κ in our framework) typically fall between 0.01 and 0.06, which is sufficiently low to be consistent with standard debt contracts being optimal in accordance with Theorem 2.2.[9]

Assumption 4.2 *Audit costs are sufficiently low, such that standard debt contracts are optimal in equilibrium.*

Following Theorem 2.2, contracts are subject to two constraints. First, repayments following overturned low reports must exceed those following high reports,

$$\hat{r}(\theta_1, \sigma_2) \geq \hat{r}(\theta_2, \emptyset). \tag{4.10}$$

Equation 4.10 is the incentive compatibility constraint, and we attach to it the Lagrange multiplier μ. Second, expected loan repayments must exceed the sum of the financial intermediaries' deposit interest rate and expected audit costs,

$$\sum_{(\theta_{ti}, \sigma_{ti})} P(\theta_{ti}, \sigma_{ti}) \hat{r}_t(\theta_{ti}, \sigma_{ti}) K^b_{t-1i} \geq r^b_t K^b_{t-1i} + \pi_1 \kappa K_{t-1i}. \tag{4.11}$$

Equation 4.11 describes the financial intermediaries' participation constraint, to which we'll attach the Lagrange multiplier ν. Both the incentive compatibility and participation constraints will be binding under efficient contracts.

[9]These estimates are drawn from Warner (1977), Weiss (1990) and Altman (1984).

FIRST ORDER NECESSARY CONDITIONS

Now that we have defined the entrepreneurs' problem, we can take first order necessary conditions:

$$N_{ti}: \; 0 = \mathbb{E}_t \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}) \left[Y_{Nti}(\theta_{ti}) - W_t \right], \tag{4.12}$$

$$\begin{aligned} K^l_{ti}: \; 0 &= \mathbb{E}_t \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}) [Y_{Kti}(\theta_{ti}) - \hat{r}_{ti}(\theta_{ti}, \sigma_{ti})] \\ &\quad + \nu_{ti} \left[\mathbb{E}_t \hat{r}_t(\theta_{ti}, \sigma_{ti}) - r^b_t - \pi_1 \kappa \right] \end{aligned} \tag{4.13}$$

$$C^e_{ti}(\theta_{ti}, \sigma_{ti}): \; 0 = u^{e\prime}(C^e_{ti}(\theta_{ti}, \sigma_{ti})) - \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}), \tag{4.14}$$

$$\begin{aligned} K^e_{t+1i}(\theta_{ti}, \sigma_{ti}): \; 0 &= -Q_t \lambda^e_{ti}(\theta_{ti}, \sigma_{ti}) \\ &\quad + \beta^e \mathbb{E}_{t+} \left[\lambda^e_{t+1i}(\theta_{t+1i}, \sigma_{t+1i}) \left(Y_{Kt+1i} + Q_{t+1}(1-\delta) \right) - \nu_{t+1i} \pi_l \kappa \right] \end{aligned} \tag{4.15}$$

$$\hat{r}_t(\theta_1, \sigma_1): \; 0 = -P(\theta_1, \sigma_1) \lambda^e_{ti}(\theta_1, \sigma_1) K^b_{t-1i} + \nu_{ti} P(\theta_1, \sigma_1) K^b_{t-1i} \tag{4.16}$$

$$\hat{r}_t(\theta_1, \sigma_2): \; 0 = -P(\theta_1, \sigma_2) \lambda^e_{ti}(\theta_1, \sigma_2) K^b_{t-1i} + \nu_{ti} P(\theta_1, \sigma_2) K^b_{t-1i} + \mu_{ti} \tag{4.17}$$

$$\hat{r}_t(\theta_2, \emptyset): \; 0 = -P(\theta_2, \emptyset) \lambda^e_{ti}(\theta_2, \emptyset) K^b_{t-1i} + \nu_{ti} P(\theta_2, \emptyset) K^b_{t-1i} - \mu_{ti} \tag{4.18}$$

Without loss of generality, $Y_{jti}(\theta_{ti})$ denotes the derivative of output with respect to factor j for entrepreneur i in period t given idiosyncratic shock realisation θ_{ti}. Also, let $\bar{Y}_{jti}$ denote the expectation of the derivative of output with respect to factor j for entrepreneur i in period t over idiosyncratic shock realisations θ_{ti}

RISK ACROSS STATES

Equations 4.16, 4.17 and 4.18 describe how the entrepreneurs' marginal utility (captured by $\lambda^e_{ti}(\theta, \sigma)$) varies across states. Entrepreneurs can vary loan repayment rates across states $\hat{r}(\theta, \sigma)$ in order to attempt to reduce variations in λ^e_{ti} across states. Entrepreneurs' ability to reduce variations in λ^e_{ti} across states is limited by the entrepreneurs' incentive compatibility constraint (4.10). The incentive compatibility constraint is binding under an efficient contract ($\mu_{ti} > 0$) resulting in varying marginal utilities across idiosyncratic states $\lambda^e_{ti}(\theta_1, \sigma_2) > \lambda^e_{ti}(\theta_1, \sigma_1) > \lambda^e_{ti}(\theta_2, \emptyset)$. Combining equations 4.16, 4.17 and 4.18 yields

$$\nu_{ti} = \lambda^e_{ti}(\theta_1, \sigma_1), \tag{4.19}$$

$$\mu_{ti} = P(\theta_1, \sigma_2) K^b_{ti} (\lambda^e_{ti}(\theta_1, \sigma_2) - \lambda^e_{ti}(\theta_1, \sigma_1)) \quad \text{and} \tag{4.20}$$

$$\lambda^e_{ti}(\theta_1, \sigma_1) = P(\theta_1, \theta_1)\lambda^e_{ti}(\theta_1, \theta_1) + P(\theta_1, \theta_2)\lambda^e_{ti}(\theta_1, \theta_2) + P(\theta_2, \emptyset)\lambda^e_{ti}(\theta_2, \emptyset). \tag{4.21}$$

In addition to the option of reduced repayments following successful audits, entrepreneurs can mitigate project risk by reducing the size of projects, relative to the size which maximises expected profits. This precautionary reduction in the size of projects translates into reductions in the quantities of capital and labour demanded compared with first best efficient allocations.

Labour and capital market wedges

Equation 4.12 describes the entrepreneurs' first order necessary condition for labour demanded. The entrepreneurs weight deviations between the ex post marginal product of labour and wages more highly in bad states, when their marginal utility is high. Averaging over entrepreneurs, the expected marginal product of labour does not equal the wage rate, even in the absence of aggregate risk. The labour market wedge is sensitive to both the idiosyncratic distribution of project outcomes $(\Theta, \pi(\Theta))$, as well as financial variables including the accuracy of audit signals (η). The same holds for the capital market.

It is through these time varying factor market wedges that shocks are amplified in our model, relative to the dynamics of the first best efficient allocations. The insurance flows we study in this chapter have a direct bearing on the behaviour of these factor market wedges: a flow of insurance transfers toward entrepreneurs will increase the entrepreneurs' wealth and decrease their borrowing, reducing the variation of entrepreneurs' marginal utilities across states and consequently reducing the factor market wedges of inefficiency.

Let the labour and capital market wedges be defined as follows,

$$\tau_{Nti} := \frac{\bar{Y}_{Nti} - W_t}{\bar{Y}_{Nti}} \qquad \text{and} \qquad \tau_{Kti} := \frac{\bar{Y}_{Kti} - r^b_t}{\bar{Y}_{Kti}}.$$

Combining equations 4.12, 4.13 and 4.21 yields the following optimality conditions:

$$\frac{\lambda^e_{ti}(\theta_1, \theta_2)}{\lambda^e_{ti}(\theta_1, \theta_1)} = 1 + \frac{\tau_{Nti}}{P(\theta_1, \theta_2)\xi_t}, \tag{4.22}$$

$$\frac{\lambda^e_{ti}(\theta_2, \emptyset)}{\lambda^e_{ti}(\theta_1, \theta_1)} = 1 - \frac{\tau_{Nti}}{P(\theta_2, \emptyset)\xi_t} \qquad \text{and} \tag{4.23}$$

$$\tau_{Nti} = \tau_{Kti} - \frac{\pi_1 \kappa}{\bar{Y}_{Kti}}. \tag{4.24}$$

Derivations of equations 4.22, 4.23 and 4.24 can be found in Appendix 4.D.1. Equations 4.22 and 4.23 relate the entrepreneurs' marginal rates of substitution for consumption across project outcomes to the labour market wedge. Equation 4.24 relates the labour market wedge to the capital market wedge. For entrepreneurs, it is efficient to reduce both labour and capital demanded in order to mitigate project risk. Equation 4.24 confirms that the labour market wedge (the left hand side, τ_{Nti}) is less than the capital market wedge (τ_{Kti}). The difference between the two wedges results follows as a result of auditing costs, which are increasing in the capital factor but not in the labour factor.

ENTREPRENEURS' SAVINGS BEHAVIOUR

Entrepreneurs' preferences are intertemporally homothetic, and their technology is scalable. This means that entrepreneurs' actions in terms of consumption, labour and capital hired are equal as a share of capital brought into the current period. Taken together, we can describe the aggregate behaviour of the population of entrepreneurs as a function of the mean wealth of entrepreneurs. Mean preserving fluctuations in the ex ante distribution of wealth across entrepreneurs do not affect market prices or aggregate quantities traded.

Also note that under logarithmic utility, where the income and substitution effects of interest rates on savings cancel, the efficient savings decision of entrepreneurs which uniquely satisfies equation 4.15 is

$$C^e_{ti} = \frac{1-\beta^e}{\beta^e} Q_t K^e_{ti}. \tag{4.25}$$

A derivation of equation 4.25 is found in Appendix 4.D.2.

SOLVING FOR LEVERAGE

Given that C^e_{ti} is equal to a fixed proportion of K^e_{t+1i} regardless of idiosyncratic state (equation 4.25), and that $C^e_{ti} = 1/\lambda^e_{ti}$ by equation 4.14, we can write down the ratios $\lambda^e_{ti}(\theta_1, \theta_2)/\lambda^e_{ti}(\theta_1, \theta_1)$ and $\lambda^e_{ti}(\theta_2, \emptyset)/\lambda^e_{ti}(\theta_1, \theta_1)$ in terms of the accumulation constraints

4.8:

$$\frac{\lambda_{ti}^e(\theta_1,\theta_2)}{\lambda_{ti}^e(\theta_1,\theta_1)} = \frac{Y_{ti}(\theta_1) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b\hat{r}_{ti}(\theta_1,\sigma_1) - W_tN_{ti}}{Y_{ti}(\theta_1) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b\hat{r}_{ti}(\theta_1,\sigma_2) - W_tN_{ti}} \quad \text{and} \tag{4.26}$$

$$\frac{\lambda_{ti}^e(\theta_2,\emptyset)}{\lambda_{ti}^e(\theta_1,\theta_1)} = \frac{Y_{ti}(\theta_1) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b\hat{r}_{ti}(\theta_1,\sigma_1) - W_tN_{ti}}{Y_{ti}(\theta_2) + Q_t(1-\delta)K_{t-1i}^e - K_{t-1i}^b\hat{r}_{ti}(\theta_2,\emptyset) - W_tN_{ti}}. \tag{4.27}$$

Combining (4.26,4.27,4.22,4.23), we can first solve for the efficient amount of risk sharing obtained by entrepreneurs. Entrepreneurs can set capital rental repayment rates on a contingent basis, enabling partial risk sharing.

$$K_{t-1i}^b[\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1)] = \bar{Y}_{ti}\left[\frac{\pi_2\xi_t - \tau_{Nti}}{\pi_2 + \pi_1\eta}\right] \tag{4.28}$$

Equation 4.28 shows that risk sharing through differentiated repayment rates across states is limited as a share of project risk $\mathbb{E}Y_{ti}\xi_t$. This means that for each entrepreneur, an increase in output through hiring more labour and renting more capital will increase the risk borne by that entrepreneur.

It will be helpful to derive a measure of leverage, and to show how leverage relates to the labour wedge, τ_N. From the entrepreneur's perspective, an increase in labour or capital hired from the household sector both increase the risk of projects, and the expected factor payments due at the end of projects. It makes sense therefore to include labour payments in our measure of leverage. One useful measure of leverage is the following:

$$L_{ti} = \frac{\bar{Y}_{ti}}{R_tQ_{t-1}K_{t-1i}^e} \tag{4.29}$$

Project output $\bar{Y}_{ti}$ is the expectation of the total income generated by the project, and $R_tQ_{t-1}K_{t-1i}^e$ is the net worth of entrepreneur i, expressed in terms of their opportunity cost, which was to redeem their capital holdings for deposits at the end of the period $t-1$.

After some rearranging, substituting equation 4.28 into equation 4.26 yields

$$L_{ti} = \frac{(\pi_2 + \pi_1\eta)\tau_{Nti}}{[\pi_2\xi_t - \tau_{Nti}][\pi_1\eta\xi_t + \tau_{Nti}]}. \tag{4.30}$$

Derivations of equations 4.28 and 4.30 are contained in Appendix 4.D.2. Consider equation 4.30. The left hand side is our production or income based measure of leverage. The right hand side is increasing in τ_{Nti}, indicating that all else equal, an increase in leverage means an increase in the labour market wedge (by equation 4.24, this also translates into an increase in the capital market wedge). This is because an increase in leverage requires the entrepreneur to accept a greater share of productive risk. Entrepreneurs hire factors until their expected marginal product, weighted by their marginal rates of substitution across states, is equal to their prices, in this case wages and interest rates. When leverage is high, the entrepreneur bears more risk. The marginal rates of substitution between worse and better project outcomes increase, increasing the wedge between the risk adjusted expected marginal product of factors, which determines demand for factors in our model, and the risk neutral expectation of factor marginal products, which would equal factor prices in a perfect information environment.

An alternative measure of leverage is the ratio of capital holdings between the two populations, K^h/K^e. This measure is closer to that considered in related literature, but in the context of our model is less useful. Importantly, it does not capture wage bill obligations which are determined at the beginning of the period and cannot be renegotiated in the case of a bad project outcome. Combining equations 4.11, 4.28 and 4.30 yields

$$\frac{K^b_{t-1i}}{K^e_{t-1i}}([\hat{r}_{ti}(\theta_2,\emptyset) - r^b_t] - \pi_1\kappa) = R_t Q_{t-1}\left[\frac{\pi_1(1-\eta)\tau_{Nti}}{\pi_1\eta\xi_t + \tau_{Nti}}\right] + \pi_1\kappa \tag{4.31}$$

Note that $[\hat{r}_{ti}(\theta_2,\emptyset) - r^b_t]$ can be interpreted as the interest rate risk premium on loans, the difference between the loan coupon rate and the deposit rate.

4.3 SYSTEMIC RISK MARKETS

4.3.1 EQUILIBRIUM

Entrepreneurs' consumption in equilibrium is strictly positive in all states (enforced by restrictions on preferences to the CRRA class of utility functions). When determining privately optimal leverage, each entrepreneur will always have access to additional loans (or savings) paying (or yielding) the risk free interest rate. In equilibrium, the net marginal benefit of additional loans or deposits will be zero, and the entrepreneurs' consumption Euler condition must bind with respect to the risk free interest rate.

It follows that

$$R^f_t \cdot \mathbb{E}_t[U^{e\prime}(C^{ei}_t)] = \mathbb{E}_t[R^{ei}_t U^{e\prime}(C^{ei}_t)].$$

Equilibrium in the insurance market implies

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} = \frac{\beta^e \mathbb{E}_t[U^{e\prime}(C_t^{ei})]}{U^{e\prime}(C_{t-1}^{ei})}. \tag{4.32}$$

Equilibrium insurance entails full consumption insurance, with the ratio of ex post intertemporal marginal rates of substitution for worker households and entrepreneurs equal to the steady state ratio β^e/β^h. The flows of insurance payments respond to consumption across the two groups, but not explicitly to the equity risk premium. This is because entrepreneurs discount the high returns associated with the high equity risk premium in accordance with the risk of those returns. As such, high expected returns to equity to not provide the same increase in entrepreneurs' marginal value as in the aforementioned models. In this way, the prospect of earning high equity returns during periods of financial stress does not necessarily encourage entrepreneurs to purchase the insurance against downturns—insurance which would recapitalise entrepreneurs during periods of financial stress and dampen the financial accelerator mechanism.

Comparison with closed common shock insurance markets

When common shock insurance markets are closed, we can combine the Euler conditions of the worker households and entrepreneurs to derive the evolution of marginal utilities of consumption as a deviation from the full common shock insurance benchmark:

$$\frac{\beta^h U^{h\prime}(C_t^{hj})}{U^{h\prime}(C_{t-1}^{hj})} = \frac{\beta^e \mathbb{E}_t[U^{e\prime}(C_t^{ei})]}{U^{e\prime}(C_{t-1}^{ei})} \cdot \left[\frac{R_t^f U^{h\prime}(C_t^{hj})}{\mathbb{E}_{t-1}[R_t^f U^{h\prime}(C_t^{hj})]} \frac{\mathbb{E}_{t-1}[R_t^f U^{e\prime}(C_t^{ei})]}{R_t^f \mathbb{E}_t U^{e\prime}(C_t^{ei})}\right]. \tag{4.33}$$

The second term on the right hand side is the ratio of the relative deviations in marginal value from the expectation formed in the previous period. Consider a recessionary shock. If the innovation in worker household marginal utilities are greater than the innovation in entrepreneur marginal utilities, then this ratio will exceed 1, and there will be a deviation from the full consumption insurance outcome, with the entrepreneurs retaining more wealth in the downturn than in the full common shock insurance counterfactual.

The innovation in worker household marginal utilities is likely to be larger in magnitude than those of entrepreneurs in response to shocks in our model for two reasons. The first is that by assumption, have a lower elasticity of intertemporal substitution than entrepreneurs—and consequently a greater desire to smooth consumption.[10] The second

[10]A lower elasticity of intertemporal substitution is a corrollary of a higher degree of consumption rel-

is that in our model, wage income is particularly volatile as the financial friction affecting firms' project finance results in a time-varying distortion between wages and the marginal product of labour.

When these conditions are present, welfare can be greater under closed common shock insurance markets than when these markets are open. The introduced distortion in consumption insurance can dampen volatility in the factor market distortions resulting from the financial friction.

4.3.2 LABOUR SUPPLY

An important determinant of the welfare effects of the financial friction is the short run interaction between the factor market wedge τ_N and labour supply Equation 4.34 presents the short run relationship between labour suppy, the factor market wedge τ_N and the unanticipated element of the current period marginal value of wealth $R_t\lambda_t^h$.

$$\Gamma^N(N_t) = \Gamma^\tau(\xi_t, \tau_{Nt}) + \Gamma^0_{t-1} \underbrace{+\log\left[\frac{R_t\lambda_t^h}{\mathbb{E}_{t-1}(R_t\lambda_t^h)}\right]}_{\text{= 0 when systemic risk markets open}} \tag{4.34}$$

where

$$\Gamma^N(N_t) = \log N_t + \log U_2^h(C_t, N_t), \qquad \Gamma^{N\prime} > 0,$$

$$\Gamma^\tau(\xi_t, \tau_{Nt}) = \log(1-\tau_{Nt}) + \log L_t(\xi_t, \tau_{Nt}), \qquad \Gamma_2^\tau(\xi_t, \tau_{Nt}) \text{ is typically negative.}$$

$$\Gamma^0_{t-1} = \log\left(\frac{1-\alpha}{1-\beta^e}\cdot\frac{\beta^e\lambda^h_{t-1}}{\beta^h\lambda^e_{t-1}}\right)$$

A derivation is contained in Appendix 4.D.3. When systemic risk markets are closed, worker household marginal values will tend to rise during downturns. Interest rates and marginal utilities (R_t and λ_t^h) tend to move in opposite directions at busines cycle frequencies, as decreases in productivity or increases in the factor market wedges pull down both interest rates and household consumption.

ative risk aversion. The assumption that worker households are more risk averse than entrepreneurs is internally consistent with our microfoundation. Within the model, entrepreneurs accept a great amount of idiosyncratic consumption risk. All else equal, agents with a greater tolerance for consumption risk would be more suited to entrepreneurship in our model.

4.4 Dynamics

Impulse responses to total factor productivity (TFP) and risk shocks are presented for the log-linearised version of the model in Appendix 4.E.

Following the positive TFP shock, the factor wedge initially rises as leverage increases. This initial increase is dampened when systemic risk markets are open, as insurance payments flow toward entrepreneurs during booms. These insurance transfers reduce leverage relative to the closed markets counterfactual, increasing factor payments to labour and capital, increasing hours worked and amplifying the output response to the positive shock. The consumption responses of both worker and entrepreneur groups of agents are greater when systemic risk markets are open, and so is the investment response.

Immediately following the positive shock, worker households bring forward consumption from future high productivity periods, drawing down their capital stocks before increasing these stocks over the medium run. Entrepreneurs' capital holdings increase sharply immediately following the positive shock, as the return to entrepreneurs' equity increases sharply. In sum, the distribution of capital holdings between households and entrepreneurs shifts sharply in favour to the entrepreneurs following the positive shock. These dynamics in the inequality of wealth holdings are persistent and amplified by the systemic risk markets.

Following the positive shock, credit spreads increase both when systemic risk markets are open and closed. Further, this increase is greater when systemic risk markets are open. However, the magnitudes of the increases in credit spreads are small.

Following the risk shock (an increase in project risk), the model dynamics are quite similar under the two assumptions of closed or open systemic risk markets. There is a transfer of wealth from entrepreneurs to worker households at the onset of the shock, and this transfer slightly increases the factor wedge, reducing wages. The transfer to households also dampens the wealth effect on labour supply, resulting in a slightly larger employment response than under the closed markets counterfactual. These small transfers have a larger effect on the accumulation of capital, which falls significantly further when systemic risk insurance markets are open than in the counterfactual, and in household consumption. It is perhaps surprising that a transfer of wealth from entrepreneurs to households at the onset of the recession *reduces* household consumption over the path of the recession, but the reason is that while this transfer increases the amount of capital held by worker households, it decreases the value of their labour endowment as wages fall further from labour's marginal product than in the counterfactual.

The next section considers the nature of systemic risk insurance flows in more detail.

4.5 WHAT DO EQUILIBRIUM INSURANCE PAYMENTS LOOK LIKE?

Worker households and entrepreneurs in our model hold a set of sophisticated financial assets, which might initially appear wildly removed from the simple real-world portfolios of households. First, the deposits held by our worker households fluctuate in value in response to innovations in asset prices q_t. Second, when systemic risk markets are open, we introduce state-contingent payments x_t that at first glance have no obvious real world counterpart. When we observe real-world bank deposits, they do not tend to revalue in response to the innovations in the values of capital assets, and nor do they respond to macroeconomic shocks.

Figures 4.1 and 4.2 present impulse responses for output, physical asset prices, entrepreneurs' net worth and household deposits. Full impulse responses for a broad range of model variables are available in Appendix 4.A, which also includes the parameterisation used to produce these figures. When systemic risk markets are closed, shocks result in immediate revaluations in household deposits and entrepreneurs' net worth through innovations in the cost of installing new capital. When systemic risk markets are open, the competitive equilibrium allocations result in much more stable paths for worker household' deposits, which are now more 'sticky' in the short run. Entrepreneurs' net worth absorbs most of the fluctuation in physical asset prices, and jump immediately in response to shocks.

In fact, when systemic risk markets are open, households' asset portfolios look much more like real world deposit accounts than when these markets are closed. Conversely, entrepreneurs' net worth looks much more like stock prices, responding dramatically in the short run to macroeconomic shocks.

This prediction of the model suggests that real world deposit accounts, which provide

Figure 4.1: Systemic risk markets and asset values. Log-linearised model. Total factor productivity shock (+1%). Systemic risk markets open (orange, solid), closed (black, dashed).

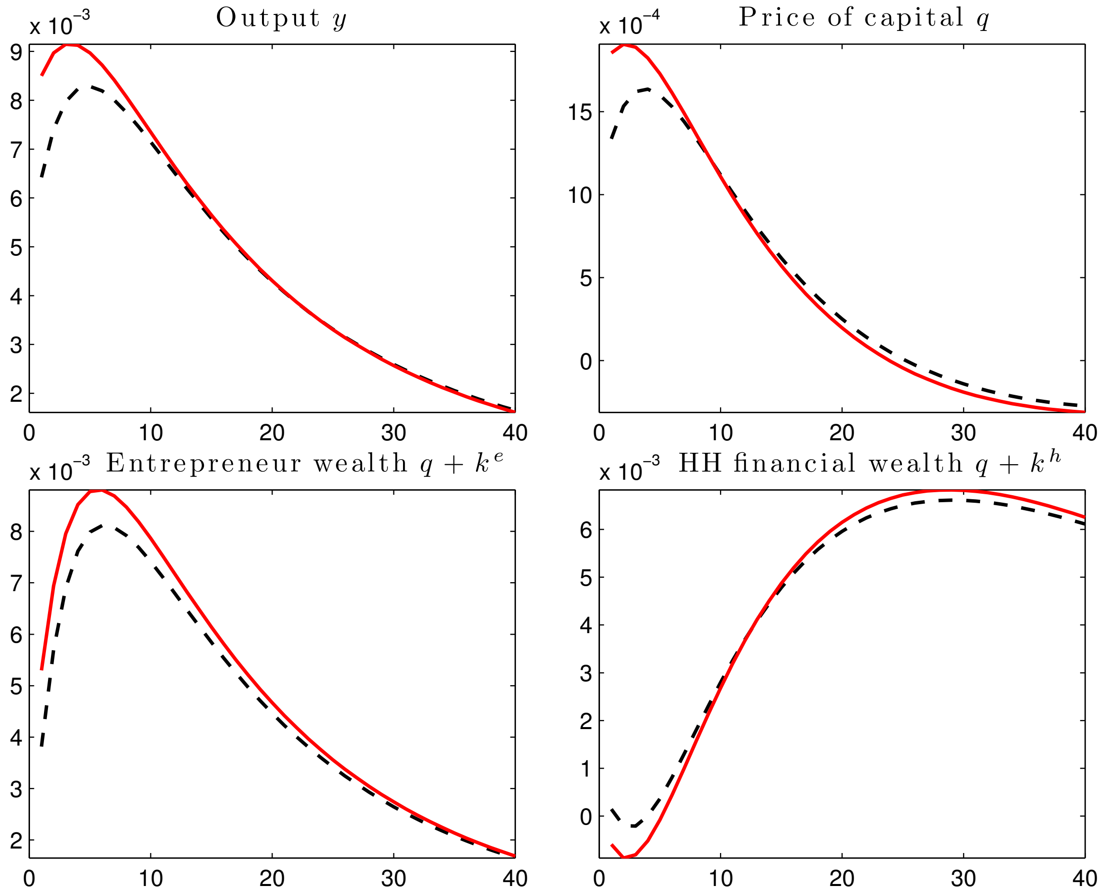
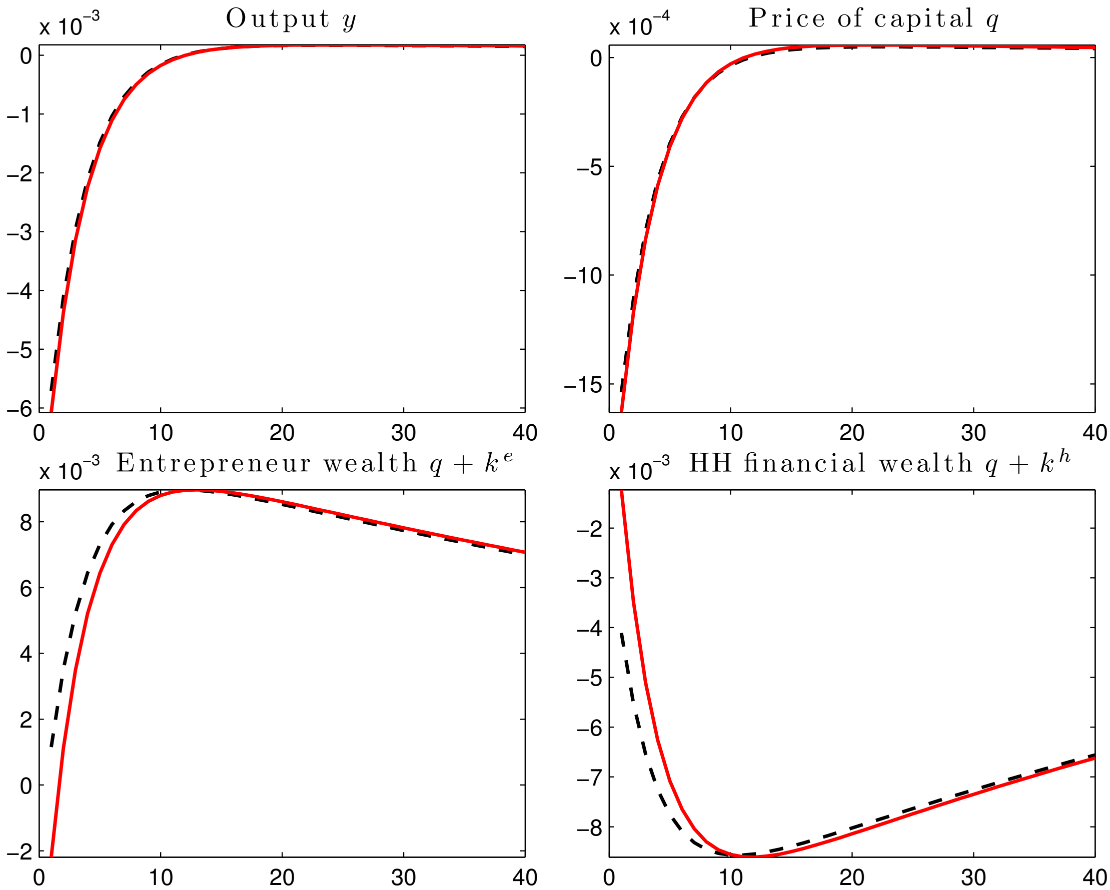


Figure 4.2: Systemic risk markets and asset values. Log-linearised model. Risk shock (+1%). Systemic risk markets open (orange, solid), closed (black, dashed).



a stability in terms of their real purchasing power, may be a good approximation of the (privately) optimal set of state-contingent securities.One interpretation of this result is that the reason why we *don't* see worker households holding portfolios of assets contingent on labour productivity and macroeconomic aggregates is not a result of the associated complexity of these products, but rather a reflection of the demand for and supply of such securities. These securities do not exist because trade in these assets does not offer opportunities for mutual benefit. Deposits work just fine. This arrangement does not however correspond to constrained efficiency.

4.5.1 Could markets implement constrained efficient allocations?

The market equilibria we have considered in this paper restrict trade to one period contracts. We've excluded markets in multiperiod contracts and market exclusion rights. Kilenthong and Townsend (2014) show that permitting trade in such rights will restore constrained efficiency in competitive equilibria in general incomplete markets settings.

The relevance of our policy analysis depends on the ability for these markets to open and these contracts to be enforceable. If it is likely that these multiperiod contracts and exclusion rights can be traded, then any policy intervention is redundant. Any optimal mechanism would need to implement the deviation between worker households' consumption-leisure marginal rate of substitution and entrepreneurs' risk-consumption marginal rate of substitution. This means that the Kilenthong and Townsend (2014) contracts must either commit workers to future employment at wages that differ from their own consumption-leisure marginal rate of substitution, or must commit entrepreneurs to offer future employment at wages that differ from their own risk-consumption marginal rate of substitution.

In the model, the commitments implied by these contracts would need to be made multilaterally between all workers and all firms. Pairwise agreements between individual worker-firm pairs might help but would not suffice for constrained efficiency, as the labour demanded by individual firms is highly volatile, being sensitive to individual firms' project outcomes. The Bureau of Labor Statistics Job Openings and Labor Turnover survey figures suggest that in a typical month, between 3% and 4% of employed workers are separated from their current employer either through layoffs or quits. This high turnover rate in the data suggests that the implementation of multiperiod labor commitments across households and firms would be costly to manage and enforce. In sum, it is our position that it is unlikely that Kilenthong and Townsend (2014) contracts can restore constrained efficiency in practise, and that there is likely to be a role for Greenwald and Stiglitz (1986) interventions in promoting constrained efficient responses to business cycle shocks in this setting.

4.A Parameterisation

The parameterisation employed in this chapter is identical to that employed in the flexible price model of Chapter 3. The details of this parameterisation are repeated here for convenience.

The worker households' preferences over consumption and labour are described by $u(C,N) = C^{1-\sigma}/(1-\sigma) - N^{1+\psi}/(1+\psi)$, with $\sigma = 2$ and $\psi = 1.5$. The difference between the worker households' and entrepreneurs' intertemporal elasticities of substitution $(1/\sigma)$ are important for our results. Recall that the entrepreneurs' intertemporal elasticity of substitution is equal to one. Typically, when worker households' intertemporal elasticities of substitution are low, the equilibrium insurance flows toward worker households will be greater in recessions. Consequently, the amplification of shocks will be larger.

The worker households' quarterly discount factor is $\beta^h = 0.995$, and constant steady state wealth shares require that the entrepreneurs' quarterly discount factor is $\beta^e = 0.975$. In the steady state, the risk free interest rate ($R = 1/\beta^h$), and the average return to entrepreneurs' equity is ($R^e = 1/\beta^e$).[11] The probability of a bad project outcome is $\pi_1 = 0.00415$, which corresponds to an annual probability of default of 1.66%, the average historical annual default probability of credit rated US firms (Schuermann and Hanson, 2004). and the probability of a type-I error resulting from an audit is $\eta = 0.1$. Small changes in this value have no effect on our main results, so long as η remains strictly positive. The resource cost of auditing (bankruptcy costs), expressed as a share of capital invested in the project is $\kappa = 0.3$, which is high in comparison with microeconomic studies that have found direct bankruptcy costs of between 1% and 6% of firms' assets (see for example Warner, 1977, Weiss, 1990 and Altman, 1984). This is quite a common weakness of costly state verification models. It is difficult to obtain the high interest rate spreads observed in the data with both realistically low default rates and bankruptcy costs. Across parameterisations matching any two of these three variables with available data has little effect on the dynamics of the model. Given that credit spreads resemble an insurance premium in our model, it is possibly the case that greater entrepreneur risk aversion would

[11] It would be possible to write up the model with a common discount factor across worker households and entrepreneurs, but, given entrepreneurs' relatively high expected return to capital, there would need to be some process governing the exit or death of entrepreneurs in the steady state in order to ensure a stable steady state distribution of capital wealth between worker households and entrepreneurs.

help in obtaining the high and volatile credit spreads seen in the data while still retaining realistically low default probabilities and bankruptcy costs. That being said, we would expect entrepreneurs to self select as being relatively risk tolerant, so it is unlikely that greater risk aversion is the best way to bring the model closer to the financial data.

The idiosyncratic risk coefficient is 5, which is sufficiently large that in low states θ_1, individual output is negative in the sense that more capital is destroyed than output produced. The steady state factor wedge is quite sensitive to the entrepreneurs' discount factor (the lower the discount factor, the less savings the entrepreneurs will accrue) and idiosyncratic risk. Under our parameterisation, the steady state factor wedge is τ_N is equal to 10.25%.

The Cobb-Douglas weight on capital is $\alpha = 0.35$. The depreciation rate of capital is $\delta = 0.02$ per quarter. The investment adjustment cost parameter is $\phi = 4$. The persistence of both risk and total factor productivity shocks are set at $\rho_A = 0.93, \rho_\xi = 0.99$, the standard deviation and magnitude of impulse response for both shocks are set at $\sigma_A = 0.01, \sigma_\xi = 0.01$.

4.A.1 DATA

All macroeconomic variables are taken from the St Louis Federal Reserve FRED database, except where otherwise stated. FRED unique identifies are provided in brackets.

The risk free real interest rate is calculated from the Effective Federal Funds rate (FEDFUNDS).[12] The average return on equity is the annualised percentage change in the Russell 3000 Total Market Return Index (RU3000TR_PC1).[13] Both interest rates are converted to real returns by subtracting CPI inflation (CPIAUCSL_PC1). The average annualised credit spread is taken from Gilchrist and Zakrajsek (2012). The average capital leverage ratio is taken from Kalemli-Ozcan et al. (2012) and is set to the upper end of the range of financial leverage (Assets / Equity) values they find for US listed non-financial firms of 2.4.[14]

[12]Using the Effective Federal Funds rate (FEDFUNDS) results in an average real interest rate of 1.35%, which compares with 1.00% derived from the 90 day treasury bill rate (TB3MS).

[13]For comparison, the average real return on equity for United States Banks (USROE) over the time period is equal to 8.65%, which compares with the average of 9.43% calculated from the Russell 3000 Index.

[14]Estimates on this measure vary dramatically, with McGrattan and Prescott (2005) finding an average Assets / Equity ratio of 1.2. This makes quite a big difference to the dynamics of the model, but the main

Table 4.2: Properties of the deterministic steady state

	STEADY STATE VALUE	
	Data	Model
GREAT RATIOS		
Capital-output[†], K/Y	2.74	2.99
Consumption-output, C/Y	0.65	0.75
Hh. consumption share, C^h/C	NA	0.83
Investment-output, I/Y	0.16	0.24
Ent. capital share, K^e/K	0.42	0.42
Labour share, NW/Y	0.56	0.58
FINANCIAL VARIABLES		
Real interest rate[†], $R^f - 1$	1.35%	2.01%
Return on equity[†], $R^e - 1$	9.43%	10.61%
Credit spread[†], $r(\theta_2, \emptyset) - r^b$	2.04%	1.91%

[†] These figures have been converted to annualised values.

Note that in our model, the consumption and investment shares of output add to one. This is not true in the data, which include Government Spending.

4.B Rest of the model

4.B.1 Households

There exists a representative household enjoying consumption C and supplying labour N in each period. Households maximise

$$U_t = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j U^h(C_{t+j}, N_{t+j}), \tag{4.35}$$

where $u_1, -u_{11}, u_2, u_{22} > 0$.

At the start of the period, the representative household holds deposits D_{t-1} which are claims issued by the financial intermediaries. At the end of the period, the financial intermediaries repay the household with interest on their deposits $R_t D_{t-1}$, and wage income $W_t N_t$. The representative household uses these deposit holdings to purchase consumption goods C_t, and the remaining deposits are carried over to the following period D_t,

$$D_t + C_t = R_t D_{t-1} + W_t N_t. \tag{4.36}$$

points made within this paper still hold under a wide range of steady state leverage ratios.

The first order conditions for the representative household can be described as follows:

$$-\frac{U_2^h(C_t, N_t)}{W_t} = U_1^h(C_t, N_t) \tag{4.37}$$

$$U_1^h(C_t, N_t) = \beta \mathbb{E}_t R_{t+1} U_1^h(C_{t+1}, N_{t+1}). \tag{4.38}$$

4.B.2 FINANCIAL INTERMEDIARIES

Financial intermediaries are perfectly competitive and risk neutral. Financial intermediaries provide a deposit asset to households, and intermediate payments from entrepreneurs to households. Between periods, financial intermediaries hold the durable capital good. Within periods, these capital holdings are lent to entrepreneurs.

Loan contracts specify interest rates following successful projects, and recovery rates following unsuccessful projects (bankruptcies). As the probability of project success is variable, intermediaries' diversified portfolios are subject to aggregate risks. This risk is passed on to the household sector depositors.

At the end of each period, all deposits are backed by capital holdings

$$D_t = Q_t K_t^b \tag{4.39}$$

where K_t^f is the amount of capital held by financial intermediaries at the end of period t, and Q_t is the cost of capital in period t. All loan earnings are passed on directly to the representative household through interest on deposits.

$$D_{t-1} R_t = Q_t(1-\delta) K_{t-1}^b + r_t^b K_{t-1}^b \tag{4.40}$$

The first term on the right hand side of equation 4.40 is the value of bank capital holdings in the current period after depreciation δ. The second term is the capital rental income earned through loans to entrepreneurs, net of monitoring costs.

4.B.3 CAPITAL PRODUCERS

Competitive capital producers (indexed by j) combine the combine the consumption good with existing capital to produce new capital goods. Firm j can produce I_{jt} units of the investment good for total cost

$$I_{tj} + \Phi\left(\frac{I_{tj}}{K_{t-1j}}\right) K_{t-1j}, \tag{4.41}$$

where

$$\Phi\left(\frac{I}{K}\right) = \frac{\phi}{2}\left(\frac{I}{K} - \delta\right)^2.$$

In competitive equilibrium, the cost of capital can be described as follows,

$$Q_t = 1 + \Phi'\left(\frac{I_t}{K_{t-1}}\right) = 1 + \phi\left(\frac{I_t}{K_{t-1}} - \delta\right). \tag{4.42}$$

The final condition we require ensures market clearing in the goods market:

$$Y_t = C_t + C_t^e + Q_t I_t + \pi_1 \kappa K_{t-1} \tag{4.43}$$

4.C Derivation of equation 4.32

At the end of period t, we allow agents to trade in claims derivative on the following period's commonly observed exogenous state, $z_{t+1} = (A_{t+1}, \xi_{t+1})$. Let $P_t(z)$ be the price at the end of period t of a security returning one unit of capital at the beginning of period $t+1$, conditional upon the realisation of $z_{t+1} = z$. Let $X_t^h(z)$ denote the amount of securities of type z purchased by the household in period t, and $X_t^e(z)$ the purchases of the entrepreneurs. Market clearing requires that

$$X_t^h(z) + X_t^e(z) = 0 \qquad \forall t, z \tag{4.44}$$

These transactions are settled at the beginning of each period, with units of the capital stock. Consequently, the amount of capital held at the end of period $t-1$ is no longer equal to the amount of capital which can be employed in projects in period t, which follows the settlement of these insurance contracts. We'll denote the capital brought forward by each agent i of type j from period $t-1$ to period t by K_{t-1i}^{pj}. Following the settlement of insurance contracts at the beginning of period t, the capital available to agent i of type j to lend or allocate toward projects becomes K_{ti}^{j}. Note the distinction in time subscript from earlier sections, which accounts for the fact that capital controlled by each agent in the current period t depends on the realisations of aggregate states in period t.

The financial intermediary acts as agent of the household sector in this market. At the end of the period, deposit wealth is used to purchase capital goods and state contingent securities,

$$D_t = Q_t K_t^{pb} + \int_z P_t(z) X_t^b(z) dz_{t+1}. \tag{4.45}$$

At the start of the period, bank capital holdings are augmented following insurance transfers,

$$K_t^b = K_{t-1}^{pb} + X_{t-1}^b(z_t). \tag{4.46}$$

The updated capital stock is then lent to entrepreneurs, and the total proceeds are distributed to the representative household as interest on their deposits.

$$D_{t-1} R_t = Q_t(1-\delta) K_{t-1}^b + r_t^b K_{t-1}^b \tag{4.47}$$

The representative household's capital accumulation constraint remains unchanged.

Entrepreneur i's accumulation constraint can be re-written as follows

$$Q_t K_{ti}^{pe} = Y_{ti}(\theta_{ti}) + Q_t(1-\delta) K_{ti}^e - K_{ti}^b \hat{r}_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti} - C_{ti}^e - \int_z P_t(z) X_{ti}^e(z) dz_{t+1} \tag{4.48}$$

$$K_{ti}^e = K_{t-1i}^{pe} + X_{t-1i}^e(z_t). \tag{4.49}$$

Note that the financial intermediaries maximise returns from insurance purchases subject to the household's objective function and stochastic discount factor. The first order necessary conditions for $X_{ti}^j(z_{t+1})$ are

$$P_t(z_{t+1}) \lambda_t^h = \beta \mathbb{E}_t R_{t+1} \lambda_{t+1}^h(z_{t+1}) \qquad \forall z, \tag{4.50}$$

$$P_t(z_{t+1}) \lambda_{ti}^e = \beta^e \mathbb{E}_t R_{t+1i}^e(z_{t+1}, \theta_{t+1i}, \sigma_{t+1i}) \lambda_{t+1i}^e(z_{t+1}, \theta_{t+1i}, \sigma_{t+1i}) \qquad \forall z, \theta, \sigma, i, \tag{4.51}$$

where

$$R_{ti}^e(\theta_{ti}, \sigma_{ti}) = \frac{Y_{ti}(\theta_{ti}) + Q_t(1-\delta) K_{ti}^e - K_{ti}^b \hat{r}_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti}}{Q_{t-1} K_{ti}^e}$$

Combining (4.50,4.51) yields the following optimality condition, which current transfers X_{ti}^{j} must satisfy in competitive equilibrium:

$$\frac{\beta R_t \lambda_t^h}{\lambda_{t-1}^h} = \frac{\beta^e \mathbb{E}_t R_{ti}^e(\theta_{ti}, \sigma_{ti}) \lambda_{ti}^e(\theta_{ti}, \sigma_{ti})}{\lambda_{t-1i}^e} = P_{t-1}(z_t) \qquad \forall i. \tag{4.52}$$

Note that even after the realisation of z_t, it is still necessary to take expectations over entrepreneurs' marginal utility in the current period, which depends on their individual project outcomes observed later within the period.

We can re-write our insurance optimality condition expressed by equation 4.52 in terms of the deviation from perfect aggregate risk insurance:

$$\frac{\beta \lambda_t^h}{\lambda_{t-1}^h} = \frac{\beta^e \mathbb{E}_t \lambda_{ti}^e(\theta_{ti}, \sigma_{ti})}{\lambda_{t-1i}^e} \cdot \frac{\mathbb{E}_t R_{ti}^e(\theta_{ti}, \sigma_{ti}) \lambda_{ti}^e(\theta_{ti}, \sigma_{ti})}{R_t \mathbb{E}_t \lambda_{ti}^e(\theta_{ti}, \sigma_{ti})}. \tag{4.53}$$

The second term on the right hand side is equal to one,[15] which leaves us with full risk sharing of common shocks:

$$\frac{\beta \lambda_t^h}{\lambda_{t-1}^h} = \frac{\beta^e \mathbb{E}_t \lambda_{ti}^e(\theta_{ti}, \sigma_{ti})}{\lambda_{t-1i}^e}. \tag{4.32}$$

[15]This can be shown in two different ways. First, we could substitute in the entrepreneurs' savings condition 4.25. Alternatively, we could note that the entrepreneurs have access to risk free loans and deposits at the gross interest rate R_t. It follows that we could re-write their Euler condition with risk free interest rate R_t.

4.D USEFUL DERIVATIONS

4.D.1 DERIVATION OF EQUATIONS (4.22 , 4.23, 4.24)

Expanding the entrepreneurs' first order conditions for labour demand (4.12) yields

$$\begin{aligned}
0 &= P(\theta_1,\sigma_1)\lambda^e_{ti}(\theta_1,\sigma_1)Y_{Nti}(\theta_1) + P(\theta_1,\sigma_1)\lambda^e_{ti}(\theta_1,\sigma_2)Y_{Nti}(\theta_1)\\
&\quad + P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Nti}(\theta_2) - \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})W_t\\
&= P(\theta_1,\sigma_1)\lambda^e_{ti}(\theta_1,\sigma_1)Y_{Nti}(\theta_1) + P(\theta_1,\sigma_1)\lambda^e_{ti}(\theta_1,\sigma_2)Y_{Nti}(\theta_1)\\
&\quad + P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Nti}(\theta_2) + P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Nti}(\theta_1)\\
&\quad - P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Nti}(\theta_1) - \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})W_t\\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})\theta_1 Y_{Nti} + P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Nti}(\theta_2-\theta_1) - \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})W_t\\
&= \lambda^e_{ti}(\theta_1,\sigma_1)(\theta_1 Y_{Nti} - W_t) + P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Nti}\xi_t\\
&= -(W_t - \theta_1 Y_{Nti}) + P(\theta_2,\emptyset)\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)}Y_{Nti}\xi_t
\end{aligned}$$

$$\begin{aligned}
\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)} &= \frac{1}{P(\theta_2,\emptyset)\xi_t}\frac{W_t-\theta_1 Y_{Nti}}{Y_{Nti}}\\
&= \frac{1}{P(\theta_2,\emptyset)\xi_t}\frac{W_t-[\theta_1+P(\theta_2,\emptyset)\xi_t-P(\theta_2,\emptyset)\xi_t]Y_{Nti}}{Y_{Nti}}\\
&= \frac{1}{P(\theta_2,\emptyset)\xi_t}\frac{W_t-[1-P(\theta_2,\emptyset)\xi_t]Y_{Nti}}{Y_{Nti}}\\
\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)} &= 1-\frac{1}{P(\theta_2,\emptyset)\xi_t}\frac{Y_{Nti}-W_t}{Y_{Nti}} = (4.23\ \text{RHS}).
\end{aligned}$$

Now, re-write (4.21) as follows:

$$P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset) = [P(\theta_2,\emptyset)+P(\theta_1,\theta_2)]\lambda^e_{ti}(\theta_1,\sigma_1) - P(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2),$$

and divide through by $\lambda^e_{ti}(\theta_1,\sigma_1)$:

$$P(\theta_2,\emptyset)\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)} = [P(\theta_2,\emptyset)+P(\theta_1,\theta_2)] - P(\theta_1,\theta_2)\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\sigma_1)},$$

Rearranging and substituting in (4.23) yields

$$
\begin{aligned}
\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\sigma_1)} &= \frac{P(\theta_2,\emptyset)+P(\theta_1,\theta_2)}{P(\theta_1,\theta_2)} - \frac{P(\theta_2,\emptyset)}{P(\theta_1,\theta_2)}\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)} \\
&= \frac{P(\theta_2,\emptyset)+P(\theta_1,\theta_2)}{P(\theta_1,\theta_2)} - \frac{P(\theta_2,\emptyset)}{P(\theta_1,\theta_2)}\left[1-\frac{\tau_{Nti}}{P(\theta_2,\emptyset)\xi_t}\right] \\
&= \frac{P(\theta_2,\emptyset)+P(\theta_1,\theta_2)}{P(\theta_1,\theta_2)} - \frac{P(\theta_2,\emptyset)}{P(\theta_1,\theta_2)} + \frac{P(\theta_2,\emptyset)}{P(\theta_1,\theta_2)}\frac{\tau_{Nti}}{P(\theta_2,\emptyset)\xi_t} \\
\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\sigma_1)} &= 1+\frac{\tau_{Nti}}{P(\theta_1,\theta_2)\xi_t} = (4.22\ \text{RHS}).
\end{aligned}
$$

Now we consider equation (4.24). Substituting (4.19) into (4.13) yields

$$
\begin{aligned}
0 &= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})[Y_{Kti}(\theta_{ti})-\hat{r}_{ti}(\theta_{ti},\sigma_{ti})] + \lambda^e_{ti}(\theta_1,\sigma_1)\left[\mathbb{E}_t\hat{r}_t(\theta_{ti},\sigma_{ti})-r^d_t-\pi_1\kappa\right] \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] + \lambda^e_{ti}(\theta_1,\sigma_1)\left[\mathbb{E}_t\hat{r}_t(\theta_{ti},\sigma_{ti})\right] \\
&\quad - \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})\hat{r}_{ti}(\theta_{ti},\sigma_{ti}) \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\quad + \lambda^e_{ti}(\theta_1,\sigma_1)[P(\theta_1,\theta_2)\hat{r}_{ti}(\theta_1,\theta_2)+P(\theta_2,\emptyset)\hat{r}_{ti}(\theta_2,\emptyset)] \\
&\quad - [P(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2)\hat{r}_{ti}(\theta_1,\theta_2)+P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)\hat{r}_{ti}(\theta_2,\emptyset)] \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\quad + \lambda^e_{ti}(\theta_1,\sigma_1)[P(\theta_1,\theta_2)+P(\theta_2,\emptyset)]\hat{r}_{ti}(\theta_2,\emptyset) \\
&\quad - [P(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2)+P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)]\hat{r}_{ti}(\theta_1,\theta_2) \\
&= \mathbb{E}_t\lambda^e_{ti}(\theta_{ti},\sigma_{ti})Y_{Kti}(\theta_{ti}) - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\quad + \{[P(\theta_1,\theta_2)+P(\theta_2,\emptyset)]\lambda^e_{ti}(\theta_1,\sigma_1)-P(\theta_1,\theta_2)\lambda^e_{ti}(\theta_1,\theta_2)-P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)\}\hat{r}_{ti}(\theta_1,\theta_2) \\
&= \lambda^e_{ti}(\theta_1,\sigma_1)\theta_1 Y_{Kti} + P(\theta_2,\emptyset)\lambda^e_{ti}(\theta_2,\emptyset)Y_{Kti}\xi_t - \lambda^e_{ti}(\theta_1,\sigma_1)\left[r^d_t+\pi_1\kappa\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by (4.21))} \\
&= \theta_1 Y_{Kti} + P(\theta_2,\emptyset)\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\sigma_1)}Y_{Kti}\xi_t - \left[r^d_t+\pi_1\kappa\right] \\
&= \theta_1 Y_{Kti} + \left[1-\frac{\tau_{Nti}}{P(\theta_2,\emptyset)\xi_t}\right]P(\theta_2,\emptyset)Y_{Kti}\xi_t - \left[r^d_t+\pi_1\kappa\right] \qquad \text{(by (4.23))}
\end{aligned}
$$

$$\tau_{Nti} = \frac{Y_{Kti}\left[\theta_1 + P(\theta_2, \emptyset)\xi_t\right] - r_t^d - \pi_1\kappa}{Y_{Kti}}$$
$$\tau_{Nti} = \frac{Y_{Kti} - r_t^d - \pi_1\kappa}{Y_{Kti}} = (4.24 \text{ RHS})$$

4.D.2 DERIVATION OF EQUATION 4.25

First, note that the entrepreneurs' preferences exhibit intertemporal homotheticity, and that projects are constant-returns-to-scale. The combination of these facts results in the consequence that both the choices of each entrepreneur and their returns and risks are scalable in K_{ti}^e. For simplicity, we'll re-write the accumulation constraint as

$$Q_t K_{t+1i}^e = R_{ti}^e K_{ti}^e - C_{ti}^e,$$

where

$$R_{ti}^e = \frac{Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K_{ti}^e - K_{ti}^l \hat{r}_{ti}(\theta_{ti}, \sigma_{ti}) - W_t N_{ti}}{K_{ti}^e}$$

in order to capture the scalability of the entrepreneur's problem. R_{ti}^e is an idiosyncratic shock realised before the consumption decision is taken. In order to describe the entrepreneur's consumption decision, we'll re-write their problem as a Bellman equation formulated after the realisation of R_{ti}^e. We'll denote $W_{ti} = R_{ti}^e K_{ti}^e$, which is the resources the entrepreneur has available when they make their consumption decision in period t.

$$V(W_{ti}) = \max_{C_{ti}^e} \log C_{ti}^e + \beta^e \mathbb{E}_t V(W_{t+1i}),$$

subject to

$$W_{t+1i} = \frac{R_{t+1i}^e}{Q_t}(W_{ti} - C_{ti}^e).$$

Which we can re-write as

$$V(W_{ti}) = \max_{C_{ti}^e} \log C_{ti}^e + \beta^e \mathbb{E}_t V\left(\frac{R_{t+1i}^e}{Q_t}(W_{ti} - C_{ti}^e)\right).$$

Dropping the subscript i, the first order condition for C^e is

$$\frac{1}{C_t^e} = \beta^e \mathbb{E}\frac{R_{t+1}^e}{Q_t} V'\left(\frac{R_{t+1i}^e}{Q_t}(W_{ti} - C_{ti}^e)\right)$$

We proceed by guessing a particular functional form for V, and verifying that this functional form satisfies the conditions above.

Let

$$\hat{V}(W) = \frac{1}{1-\beta^e}\log W + k,$$

where k is some constant.

Assuming, that $\hat{V}$ is the correct value function, we substitute it into the first order condition to solve for consumption:

$$\frac{1}{C_t^e} = \frac{\beta^e}{1-\beta^e}\frac{1}{(W_t - C_t^e)}$$

which rearranges to yield

$$C_t^e = (1-\beta^e)W_{ti}$$

We now verify our guess value function by substituting it into the entrepreneur's Bellman equation,

$$V(W_{ti}) = \max_{C_{ti}^e} \log C_{ti}^e + \beta^e \mathbb{E}_t \frac{1}{1-\beta^e}\log\left(\frac{R_{t+1i}^e}{Q_t}(W_{ti} - C_{ti}^e)\right) + \beta^e k.$$

And substitute in our optimal consumption decision,

$$\begin{aligned}
V(W_{ti}) &= \log[(1-\beta^e)W_{ti}] + \beta^e \mathbb{E}_t \frac{1}{1-\beta^e}\log\left(\frac{R_{t+1i}^e}{Q_t}(W_{ti} - (1-\beta^e)W_{ti})\right) + \beta^e k \\
&= \log W_{ti} + \log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}(\beta^e W_{ti})\right) + \beta^e k \\
&= \log W_{ti} + \log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\log W_{ti} + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}\beta^e\right) + \beta^e k \\
&= \frac{1}{1-\beta^e}\log W_{ti} + \log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}\beta^e\right) + \beta^e k \\
V(W_{ti}) &= \frac{1}{1-\beta^e}\log W_{ti} + k,
\end{aligned}$$

where

$$k = \frac{1}{1-\beta^e}\left[\log(1-\beta^e) + \frac{\beta^e}{1-\beta^e}\mathbb{E}_t \log\left(\frac{R_{t+1i}^e}{Q_t}\beta^e\right)\right].$$

This confirms that our guess value function $\hat{V}$ satisfies the entrepreneur's problem, and that $C_t^e = (1-\beta^e)W_{ti}$ is the optimal consumption decision for our entrepreneur.

We can now substitute this back into the entrepreneurs' accumulation constraint (4.8) to obtain

$$C^e_{ti} = (1-\beta^e)[Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K^e_{ti} - K^l_{ti}\hat{r}_{ti}(\theta_{ti},\sigma_{ti}) - W_tN_{ti}].$$

It follows that

$$Q_tK^e_{t+1i} = \beta^e[Y_{ti}(\theta_{ti}) + Q_t(1-\delta)K^e_{ti} - K^l_{ti}\hat{r}_{ti}(\theta_{ti},\sigma_{ti}) - W_tN_{ti}],$$

and

$$C^e_{ti} = \frac{1-\beta^e}{\beta^e}Q_tK^e_{t+1i}. \tag{4.25}$$

DERIVATION OF EQUATIONS 4.28, 4.30 AND 4.31

First, substitute the national income equation (**??**) into the marginal rates of substitution conditions (4.26) and (4.26) to eliminate W_tN_t,

$$\frac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)} = \frac{(\theta_1 - 1 + \tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}\hat{r}_{ti}(\theta_1,\sigma_1) + r^dK^l_t + \pi_l\kappa K_t}{(\theta_1 - 1 + \tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}\hat{r}_{ti}(\theta_2,\emptyset) + r^dK^l_t + \pi_l\kappa K_t} \tag{4.54}$$

$$\frac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\theta_1)} = \frac{(\theta_1 - 1 + \tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}\hat{r}_{ti}(\theta_1,\sigma_1) + r^dK^l_t + \pi_l\kappa K_t}{(\theta_2 - 1 + \tau_{Nt})\bar{Y}_t + [Q_t(1-\delta)+r^d_t]K^e_{ti} - K^l_{ti}\hat{r}_{ti}(\theta_2,\emptyset) + r^dK^l_t + \pi_l\kappa K_t} \tag{4.55}$$

Note that $\pi_1\theta_1 + \pi_2\theta_2 = 1$, which implies $\theta_1 = 1 - \pi_2\xi_t$, and $\theta_2 = 1 + \pi_1\xi_t$.

Rather than working with specific interest rates $\hat{r}(\cdot)$ and project returns θ, it will be helpful to re-write these conditions in terms of project risk $\xi_t = \theta_2 - \theta_1$, and risk sharing $[\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1)]$. From the financial intermediaries' participation constraint we have

$$K^l_{ti}[(\pi_2 + \pi_1\eta)\hat{r}_{ti}(\theta_2,\emptyset) + \pi_1(1-\eta)\hat{r}_{ti}(\theta_1,\sigma_1)] = r^dK^l_t + \pi_l\kappa K_t$$

which allows us to re-write $\hat{r}_{ti}(\theta_2,\emptyset), \hat{r}_{ti}(\theta_1,\sigma_1)$ in terms of required returns and risk:

$$\hat{r}_{ti}(\theta_2,\emptyset) = r^d + \pi_l\kappa\frac{K_t}{K^l_{ti}} + \pi_1(1-\eta)(\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1))$$

$$\hat{r}_{ti}(\theta_1,\sigma_1) = r^d + \pi_l\kappa\frac{K_t}{K^l_{ti}} - (\pi_2 + \pi_1\eta)(\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1))$$

which we can substitute back into (4.54,4.55) and rearrange to obtain

$$R_tQ_{t-1}\frac{K^e_{ti}}{K^l_{ti}} = (\pi_2\xi_t - \tau_{Nt})\frac{\bar{Y}_t}{K^l_{ti}} - \left[\frac{1}{1-\lambda^e_{ti}(\theta_1,\theta_2)/\lambda^e_{ti}(\theta_1,\theta_1)} - \pi_1(1-\eta)\right](\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1))$$

$$\begin{aligned} R_tQ_{t-1}\frac{K^e_{ti}}{K^l_{ti}} &= (\pi_2\xi_t - \tau_{Nt})\frac{\bar{Y}_t}{K^l_{ti}} + \frac{\lambda^e_{ti}(\theta_2,\emptyset)/\lambda^e_{ti}(\theta_1,\theta_1)}{1-\lambda^e_{ti}(\theta_2,\emptyset)/\lambda^e_{ti}(\theta_1,\theta_1)}\frac{\bar{Y}_t}{K^l_{ti}}\xi_t \\ &\quad - \left[\frac{1}{1-\lambda^e_{ti}(\theta_2,\emptyset)/\lambda^e_{ti}(\theta_1,\theta_1)} - \pi_1(1-\eta)\right](\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1)) \end{aligned}$$

And we equate the right hand sides to solve

$$\frac{\bar{Y}_t\xi_t}{K^l_{ti}(\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1))} = \left[\frac{\dfrac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)}/\dfrac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\theta_1)} - 1}{\dfrac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)} - 1}\right]$$

Note that $\dfrac{\lambda^e_{ti}(\theta_1,\theta_2)}{\lambda^e_{ti}(\theta_1,\theta_1)} > 1$, and $\dfrac{\lambda^e_{ti}(\theta_2,\emptyset)}{\lambda^e_{ti}(\theta_1,\theta_1)} < 1$, making the right hand side strictly greater than 1. The left hand side is the ratio of productive risk to the possible amount of risk sharing following loan restructuring. The equation shows how an increase in the ratio of productive risk to risk sharing increases the entrepreneurs' marginal rates of substitution for consumption across idiosyncratic states. These marginal rates of substitution in turn determine entrepreneurs' precautionary reductions in wage and capital hiring compared with the first best efficient levels, through the wedges specified in equations (4.22) and (4.23). Substituting these factor wedges in place of the marginal rates of substitution yields

$$K^b_{t-1i}[\hat{r}_{ti}(\theta_2,\emptyset) - \hat{r}_{ti}(\theta_1,\sigma_1)] = \bar{Y}_{ti}\left[\frac{\pi_2\xi_t - \tau_{Nti}}{\pi_2+\pi_1\eta}\right] \tag{4.28}$$

Now we can use this solution to solve for the the leverage ratio

$$\frac{\bar{Y}_{ti}}{R_tQ_{t-1}K^e_{t-1i}} = \frac{(\pi_2+\pi_1\eta)\tau_{Nti}}{[\pi_2\xi_t - \tau_{Nti}][\pi_1\eta\xi_t + \tau_{Nti}]}, \qquad \text{(= (4.30 RHS))}$$

and also the ratio of household's to entrepreneurs' capital in terms of the loan coupon interest rate spread $[\hat{r}_{ti}(\theta_2,\emptyset) - r^b_t]$.

$$\frac{K^b_{t-1i}}{K^e_{t-1i}}([\hat{r}_{ti}(\theta_2,\emptyset) - r^b_t] - \pi_1\kappa) = R_tQ_{t-1}\left[\frac{\pi_1(1-\eta)\tau_{Nti}}{\pi_1\eta\xi_t + \tau_{Nti}}\right] + \pi_1\kappa \tag{4.31}$$

4.D.3 DERIVATION OF EQUATION 4.34

From the national income equation (**??**) and the entrepreneurs' optimality condition for relative demand for labour and capital (**??**), we can derive the labour share of national income as a function of the labour wedge:

$$N_t W_t = (1-\alpha)(1-\tau_{Nt})Y_t$$

We can use the expression for leverage (4.30) to eliminate income Y_t,

$$\begin{aligned} N_t W_t &= (1-\alpha)(1-\tau_{Nt})L(\xi_t,\tau_{Nt})R_t Q_{t-1}K^e_{t-1i} \\ &= (1-\alpha)(1-\tau_{Nt})L(\xi_t,\tau_{Nt})\frac{\beta^e}{1-\beta^e}\frac{1}{\lambda^e_{t-1}}R_t && \text{(using 4.25)} \\ &= (1-\alpha)(1-\tau_{Nt})L(\xi_t,\tau_{Nt})\frac{1}{\beta}\frac{\beta^e}{1-\beta^e}\frac{\lambda^h_{t-1}}{\lambda^e_{t-1}}\frac{R_t}{\mathbb{E}_{t-1}(R_t\lambda^h_t)}. && \text{(using 4.38)} \end{aligned}$$

Eliminating wages with the representative households' first order condition for labour (4.37) yields

$$N_t U^h_2(C_t,N_t) = (1-\tau_{Nt})L_t(\xi_t,\tau_{Nt})\left(\frac{1-\alpha}{1-\beta^e}\right)\frac{\beta^e}{\beta}\frac{\lambda^h_{t-1}}{\lambda^e_{t-1}}\left[\frac{R_t\lambda^h_t}{\mathbb{E}_{t-1}(R_t\lambda^h_t)}\right]. \tag{4.34}$$

4.E MODEL DYNAMICS

Figure 4.3: Log-linearised model dynamics: Total factor productivity shock. Systemic risk markets open (red,—), closed (black,– –).

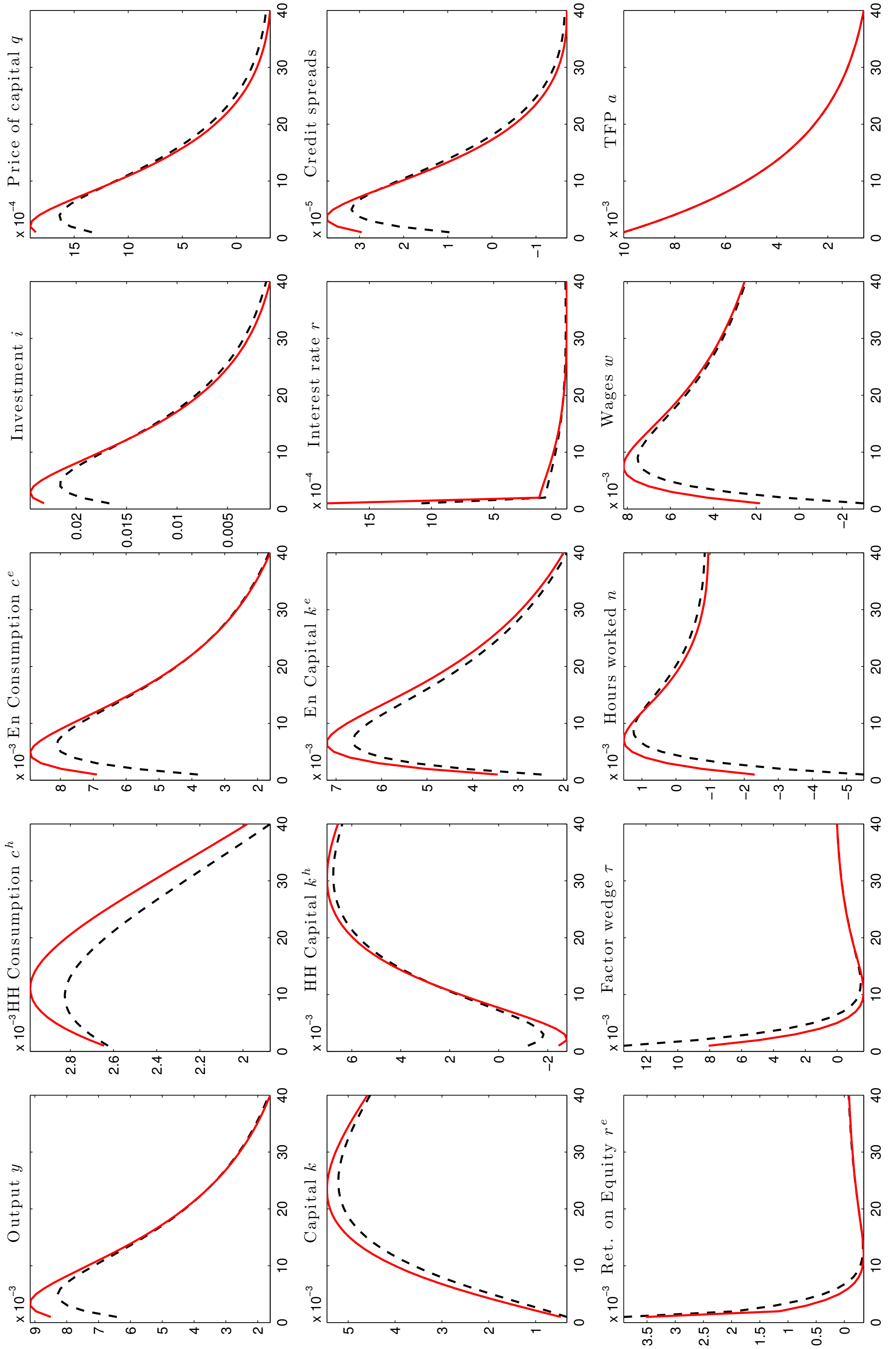
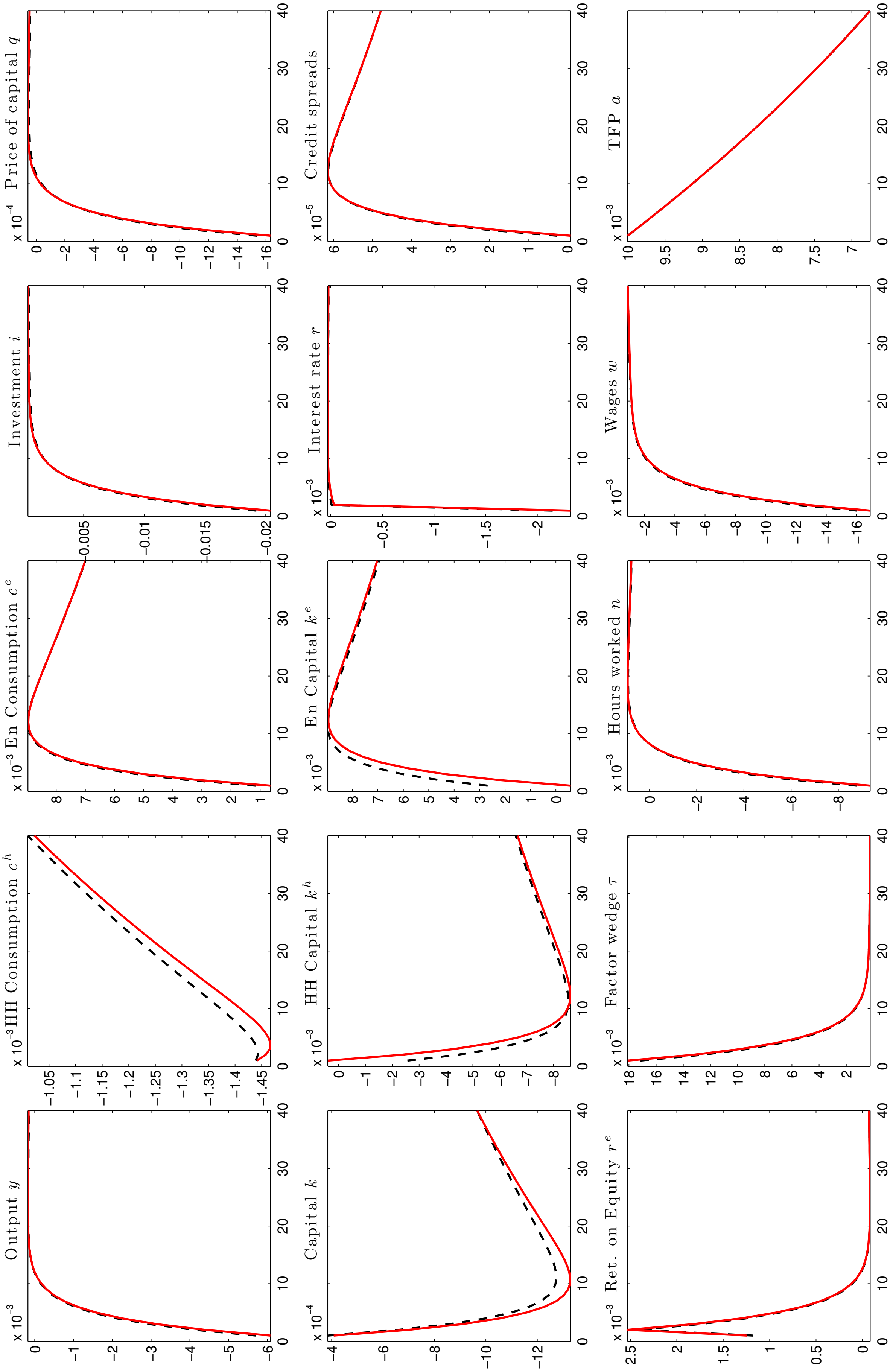


Figure 4.4: Log-linearised model dynamics: Risk shock. Systemic risk markets open (red,—), closed (black,– –).



CHAPTER 5

MONETARY POLICY AND HERDING

When wage, debt or other contracts are not indexed to the price level, fluctuations in the price level resulting from monetary actions redistribute real wealth between agents. An increase in the price level reduces the real value of fixed nominal payments. A nominal output targeting monetary policy regime permits higher price inflation during recessions than an inflation targeting regime. This results in the redistribution of real wealth from creditors to borrowers in downturns and the redistribution of real wealth from borrowers to creditors in booms. Cunning agents might well alter their behaviour in anticipation of these distributional effects of monetary policy. Compared with inflation or price level targeting, nominal output targeting provides agents with an additional incentive to undertake projects which are highly correlated with the state of the economy, increasing business cycle volatility.

CHAPTER 5

INTRODUCTION

This chapter introduces a different type of private information than we have studied in earlier chapters. Here, we assume that firms face a trade-off between idiosyncratic risk and aggregate risk, and that firm insiders' choices about this trade-off are private information, unobservable to outside investors. In addition to this private information friction, we will also consider the restriction of factor payment and managerial compensation contracts to nominal terms. Essentially, restrictions of this sort are required for monetary non-neutrality and the study of monetary policy.

When contracts are not indexed to the price level, fluctuations in the price level redistribute real wealth between agents. An increase in the price level reduces the real value of fixed nominal liabilities, and the value of fixed nominal assets. This means that monetary policy actions will affect the real distribution of wealth. For example, a nominal output targeting monetary policy regime permits higher price inflation during recessions in comparison with an inflation targeting regime. This results in the redistribution of real wealth from creditors to borrowers in downturns (and the redistribution of real wealth from borrowers to creditors in booms). Cunning agents might well alter their behaviour in anticipation of these fluctuations in wealth arising from monetary policy actions. The message of this paper is that compared with inflation or price level targeting, nominal output targeting provides agents with an additional incentive to undertake projects which are highly correlated with the state of the economy, thereby increasing business cycle volatility. This is an example of the phenomenon known as *herding*.

It turns out that herding enables agents to take advantage of the nominal output targeting regime. Under nominal output targeting, the price level is inversely correlated with total output. Agents who choose projects which are correlated with the market can typically benefit from high inflation following project failures and low inflation following project successes. This paper supports this idea with three different examples. In the first example, banks with fixed nominal deposit liabilities outstanding choose to allocate their portfolio of assets across uncorrelated and correlated securities or loan products. The worst possible outcome for these banks is that bad loan performance occurs at the same time as low inflation. This is because low inflation increases the real value of their deposit liabilities. By aligning their loan portfolios with the market, they can ensure that inflation will be high when loan performance is bad—essentially receiving a bailout with the costs passed on to their depositors. In the second example, firm managers are compensated with nominal contracts increasing in individual firms' output. Under nominal output targeting, managers can capitalise on the negative correlation between aggregate output and inflation by herding investment projects with the market, which results in an increase in expected consumption.

These individual examples differ in important ways. In the first example, which considers sticky factor payments to capital, bargains between individual depositors and firms would not necessarily improve outcomes. The real repayment to an individual depositor is not contingent on the portfolio allocation of their bank in equilibrium. Rather, the real value of individual deposits is dependent on the average portfolio allocation decisions across all banks. Social costs resulting from the actions of any bank are shared across all agents. So there is not necessarily an incentive for individual depositor-bank pairs to resolve herding issues. In the second example, this is not the case. There is a greater conflict between the owners of the firm and the managers, and better contract design/monitoring would yield private and social welfare gains. In the first example, bankers have nominal deposit liabilities, and it is risk aversion which discourages banks from specialising in niche products—which would diminish the value of the implicit bailout associated with the nominal output targeting. In the second example, the manager has a nominal compensation asset, and it is risk tolerance which enables them to take advantage of the nominal output targeting regime. Conversely, in the model presented in this paper, there is never a herding equilibrium under an inflation targeting or a price level targeting regime. This is formalised in Proposition 5.1. This is because the source of pecuniary externalities resulting in strategic complementarities is the correlation between the price level and real output. This correlation is eliminated under inflation or price level targeting, which removes any systematic correlation between the price level and output.

The behaviour studied in this paper is an example of *rational herding*, which is a form of strategic complementarity. It is in each agent's best interests to follow a certain action, if and only if they expect their peers to undertake the same action. Even under nominal output targeting, the price level will not be volatile if all agents choose diversified projects—as this would mean that real output volatility would be low, and nominal output can be stabilised with little volatility in the price level. It is the expectation that other agents will choose to herd into the correlated project which implies the expectation that the price level will be volatile and inversely correlated with real output, and it is this inverse correlation between real output and the price level which encourages herding into highly correlated projects under nominal output targeting.

RELATION TO THE LITERATURE

Other examples of rational herding behaviour include bank runs. If as a depositor you suspect that others will withdraw their funds from your bank, it may be rational to attempt to withdraw your funds immediately (Diamond and Dybvig, 1983). This is an example of rational herding, because it is the expectation of others taking a certain action—in this case by withdrawing deposits—which makes it rational for other agents to mimic that same action.

In the context of firms' project choices and banks' lending decisions, it is difficult to identify whether similarity across firms' actions is the result of herding, which could be a sign of constrained inefficiency, or merely that these highly correlated projects or loan products are high yielding. There is definitely a perception that herding is prevalent in financial markets, and there are a number of mechanisms that have been identified to explain herding behaviour. An excellent introduction to the subject is provided by Devenow and Welch (1996).

Perhaps most closely related to the present paper are Farhi and Tirole (2012) and Chari and Phelan (2014). They argue that a failing bank is more likely to receive a direct fiscal bailout if their failure coincides with weakness across peer banks—this is when the systemic costs of bank failure are highest. Foreseeing this, individual banks have an incentive to herd with their peers, ensuring that if they do fail they will likely receive a bailout. Individual banks also know that if they are successful while others are failing, they will be paying for the bailouts of others through higher taxes. Farhi and Tirole (2012) also consider the role of credit easing policies undertaken by central banks in repairing commercial banks' balance sheets in the wake of a crisis and thereby encouraging risk taking ex ante. The present paper shows that these concerns are also at the heart of the design of monetary policy targets, through the correlation between the price level and real output.

The link between monetary policy targets and the efficiency of fixed nominal financial contracts is also the subject of Koenig (2011) and Sheedy (2014). These authors find that nominal output targeting can have a stabilising effect on macroeconomic aggregates. In their models, debtors suffer comparatively large increases in marginal utility in recessions. Debtors would wish to insure this business cycle risk with creditors, but incomplete markets prevent agents from achieving constrained efficient risk sharing with decentralised trade. Nominal output targeting transfers real wealth to debtors in downturns, replicating the missing insurance payments and resulting in a Pareto welfare gain. Those same real wealth transfers are destabilising in the model presented in this paper. This distinction is driven by private information about firm project decisions, which introduces a conflict between incentive to herd and insurance against business cycle risk. Central bank actions which dampen the effects of business cycle volatility on borrowers just encourage these borrowers to increase their exposure to the business cycle, passing on much of the risks to lenders and other agents in the economy.

Traditional studies of monetary policy focus on monetary non-neutralities arising from price and wage setting rigidities, rather than fixed nominal debt contracts. Within the benchmark New Keynesian model, optimal monetary policy stabilises inflation and consequently the welfare relevant output gap (see Goodfriend and King, 1997 and Woodford, 2003, Ch. 6). There is much debate about the generalisability of this result. Generally, when wages are sticky in nominal terms, nominal output targeting can have desirable properties (Garín, Lester, and Sims, 2015). This is because nominal output targeting allows a rise in goods price inflation during downturns, which corresponds to a decrease in real wages. During recessions triggered by technology shocks, labour productivity is low, and the fall in real wages associated with nominal output targeting can restore a tight correlation between real wages and marginal labour productivity.[1]

Despite the considerable academic interest, no country has explicitly pursued a policy of nominal output targeting. As of 2012, 27 central banks were operating under explicit inflation targets (Hammond, 2012). The lack of historical data and experience with nominal output targeting is the primary reason why this paper sticks to simple models that present the intuition behind our results as clearly as possible, but which are not well suited for an econometric evaluation of the relative merits of inflation targeting and nominal output targeting monetary regimes.

[1] See also Hall and Mankiw (1994) and McCallum and Nelson (1999).

5.1 THE MODEL

The model focuses on entrepreneurs' and managers investment decisions, after they have arranged contracts for compensation or factor inputs. After outlining the key features of the framework, we consider two examples. In the first example, a bank with fixed nominal deposits outstanding determines their portfolio of assets. This example could easily be adapted to an entrepreneur operating in the non-financial sector with fixed nominal debts or wage bill. The second example considers a manager of a non-financial firm, whose nominal compensation contract is indexed to the individual output of the firm, but not indexed to prices or aggregate productivity.

Within this Chapter, we only consider the entrepreneur's or manager's decision after compensation and factor payments have been agreed. We do not solve the extensive form of the game between factor owners and entrepreneurs. The reason for this is as follows. In our framework, there is no herding externality under price level or inflation targeting, regardless of the particular form of the factor contracts in place ex ante (so long as they are constrained in their indexation to prices). It follows that in the extensive form game, and furthermore in general equilibrium, herding equilibria will not exist under price level or inflation targeting, but could exist under nominal output targeting.

5.1.1 ENTREPRENEURS AND PROJECTS

There exists a unit measure of firm insiders who can make decisions concerning firms' project investments. We'll refer to this group as entrepreneurs, although in one example we consider it will make more sense to think of this group as managers, rather than firm owners. Each entrepreneur has access to two projects, which we'll name *red* and *white*. The return of each white project is perfectly correlated with all other white projects, while the return of each agent's red project is independent of the white project and of all other agents' red projects. Paraphrasing Tolstoy, *All white projects are alike, all red projects succeed or fail in their own way.* Red projects refer to a special individual skill of each agent, while the white project is accessible to all agents. This is a slight departure from typical herding models, where each agent is typically assumed to have access to all projects, and individual projects are equally likely to emerge as the herding project. This distinction is not important for our results. The red project can be thought of as a departure from the industry standard into new technologies, products or manufacturing processes, whereas the white project can be thought of as following the industry standards closely.

Red and white projects produce consumption goods, yielding gross return θ units of the consumption good with a production technology that is increasing and concave in the labour input, $F', -F'' > 0$. Individual entrepreneurs' are price takers in labour markets. The productivity multiplier θ is equal to $\bar{\theta}$ with probability π in the case of red projects and $\pi' < \pi$ in the case of white projects. Otherwise $\theta = \underline{\theta} < \bar{\theta}$. White projects are not only correlated, but also earn a lower expected real return relative to red projects.[2] Project returns are public information ex post. Each entrepreneur's choice of project is private information both ex ante and ex post. Entrepreneur j's nominal compensation is contingent on their individual project output θ_j and the price level P, and is given by $Z(\theta, P)$. Entrepreneurs enjoy real consumption according to strictly increasing utility function $U(Z(\theta, P)/P)$.

Agents choose which project to undertake, red or white. For ease of exposition, we allow agents randomise over projects, choosing the probability of undertaking the white projects q in order to maximise expected utility. The agents' problem in the first period can be written as follows:

$$
\begin{aligned}
\max_{q\in[0,1]} \; \mathbb{E}\, q & \left(\pi' U\left(\frac{Z(\bar{\theta}, P(Q,\bar{\theta}))}{P(Q,\bar{\theta})}\right) + (1-\pi')U\left(\frac{Z(\underline{\theta}, P(Q,\underline{\theta}))}{P(Q,\underline{\theta})}\right)\right) \\
& + (1-q)\pi' \left[\pi U\left(\frac{Z(\bar{\theta}, P(Q,\bar{\theta}))}{P(Q,\bar{\theta})}\right) + (1-\pi)U\left(\frac{Z(\underline{\theta}, P(Q,\bar{\theta}))}{P(Q,\bar{\theta})}\right)\right] \\
& + (1-q)(1-\pi') \left[\pi U\left(\frac{Z(\bar{\theta}, P(Q,\underline{\theta}))}{P(Q,\underline{\theta})}\right) + (1-\pi)U\left(\frac{Z(\underline{\theta}, P(Q,\underline{\theta}))}{P(Q,\underline{\theta})}\right)\right],
\end{aligned}
\tag{5.1}
$$

where Q is the mean choice of q across all agents. The equilibrium outcome Q influences the price level P as we will see in the following section. When entrepreneurs choose the white project, there are two possible outcomes. When their project output is high, the price level will be that associated with high output, and vice-versa. When entrepreneurs

[2] In the examples we present in this paper, the assumption that the correlated white project earns lower returns than the uncorrelated red projects does not conflict with the Consumption Capital Asset Pricing Model, which predicts that the market returns of *marketable financial assets* that are highly correlated with real output should be higher than those of uncorrelated assets.

choose the red project, there are four possible outcomes. High or low individual output can occur at the same time as high or low aggregate output and associated price levels.

5.1.2 MONETARY AUTHORITY

Goods prices are flexible, the monetary authority can implement any price level it chooses. This means that there are no monetary policy errors in equilibrium—the monetary authority always hits their target. This assumption is quite strict, but provides a useful benchmark. As there are no monetary policy errors in this framework, price level targeting is equivalent to inflation targeting. We'll analyse two *monetary policy regimes*: price level targeting and nominal output targeting.

Definition 5.1 *Under* price level targeting*, the monetary authority chooses* $P = 1$.

Under nominal output targeting, we set the expected price level equal to that under price level targeting, and nominal output constant across white project outcomes. Define Q as the mean of the population of individual agents' choices of q, and θ_w is the outcome of the white project. Also, let real output be denoted by

$$Y(Q, \theta_w) = Q\theta_w + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]. \tag{5.2}$$

The first term on the right hand side captures the total output of all agents who have chosen the white project. The second term captures the total output of all agents who have chosen their individual red projects.

Definition 5.2 *Under* nominal output targeting*, the monetary authority chooses* $P(Q, \theta_w)$ *such that*

(1) $\mathbb{E}[P(Q, \theta_w)] = \pi' P(Q, \bar{\theta}) + (1 - \pi')P(Q, \underline{\theta}) = 1$ *and*

(2) $P(Q, \underline{\theta})Y(Q, \underline{\theta}) = P(Q, \bar{\theta})Y(Q, \bar{\theta})$.

The first condition ensures that the expected price level under nominal output targeting is equal to that under price level targeting. The second condition states that nominal output is constant across white project outcomes.

It will be helpful to re-write our definition of nominal output targeting as a function mapping agents' actions to the price level. Lemma 5.1 provides this result.

Lemma 5.1 *Under nominal output targeting, the price level is contingent on the share of agents selecting the high white project (Q) and the outcome of the white project (θ_w). The price level is given by*

$$
\begin{aligned}
P(Q,\bar{\theta}) &= \frac{Q\underline{\theta} + (1-Q)[\pi\bar{\theta} + (1-\pi)\underline{\theta}]}{Q[(1-\pi')\bar{\theta} + \pi'\underline{\theta}] + (1-Q)[\pi\bar{\theta} + (1-\pi)\underline{\theta}]} \\
P(Q,\underline{\theta}) &= \frac{Q\bar{\theta} + (1-Q)[\pi\bar{\theta} + (1-\pi)\underline{\theta}]]}{Q[(1-\pi')\bar{\theta} + \pi'\underline{\theta}] + (1-Q)[\pi\bar{\theta} + (1-\pi)\underline{\theta}]}
\end{aligned}
\tag{5.3}
$$

The proof of lemma 5.1 is contained in Appendix 5.A.

Our definition of monetary policy rules does not permit errors. That is, the monetary authority hits their selected target with certainty. This is a strong assumption, and in practise, it is likely that there are not only policy errors, but that these errors will be correlated with business cycle shocks in a predictable way. Nevertheless, under inflation or price level targeting, we should see less volatility in the price level, and therefore less opportunity to manipulate actions in order to take advantage of fluctuations in the real value of nominal contracts over the business cycle. Under nominal output targeting, which aims to produce a predictable negative correlation between output and price inflation, we would expect there to be greater opportunity to manipulate actions in order to take advantage of this negative correlation.

5.1.3 Competitive equilibrium

Definition 5.3 *A schedule of actions $\{Q, P\}$ is a* competitive equilibrium *under a given monetary policy regime if and only if $q^* = Q$ and $P(Q,\cdot)$ is consistent with the given monetary policy regime.*

a. A symmetric equilibrium is called a herding equilibrium *if and only if all agents choose the common white project with probability 1, ($Q = 1$).*

b. A symmetric equilibrium is called a diversification equilibrium *if and only if all agents choose their individual red projects with probability 1, ($Q = 0$).*

It is important to note that *rational herding* in financial economics is a term reserved for situations where agents pursue a given action because they expect other agents to pursue the same action. The *herding equilibria* that we identify in this paper share this property—it is the expectation that others will choose the white project which encourages agents to choose the white project. In general, the observation that investment decisions are highly correlated is not necessarily evidence of herding, but could merely be an indication that the highly correlated project offers high returns.

The symmetric equilibrium concept we employ in this paper is narrow, particularly so given that the subject of the paper is monetary policy and business cycles—typically analysed in dynamic general equilibrium models. In this paper, agents' compensation schemes are taken as given and do not respond to monetary policy regimes nor to expectations of output and herding. It is likely that in practise, compensation contracts as well as project sizes and factor prices would respond to expectations of agents' actions and monetary policy regimes.

Proposition 5.1 ensures us that this abstraction away from the endogenous nature of contracts and factor prices under subgame perfect or general equilibrium is not important for the main finding of this paper. Proposition 5.1 states that under weak conditions, namely that agent's compensation is increasing in output and that the red project has a greater expected return than the white project, there is no herding equilibrium under the price level targeting regime. It follows that even in more richly specified models, these conditions are sufficient to ensure that any model will not exhibit a herding equilibrium under price level or inflation targeting, but may indeed exhibit a herding equilibrium under nominal output targeting.

Proposition 5.1 *If agents' compensation is strictly increasing in project output ($Z(\bar{\theta}, P) > Z(\underline{\theta}, P)$ for all P) then the only symmetric equilibrium under price level targeting is the diversification equilibrium where all agents choose their individual red projects ($Q = 0$).*

Proof. By Definition 5.1, the price level P is equal to 1 in all states under price level targeting. Substituting this into the agents' objective function (5.1) yields

$$\max_{q\in[0,1]} \mathbb{E}\, q \left[\pi' U\left(Z(\bar{\theta},1)\right) + (1-\pi')U\left(Z(\underline{\theta},1)\right)\right] + (1-q)\left[\pi U\left(Z(\bar{\theta},1)\right) + (1-\pi)U\left(Z(\underline{\theta},1)\right)\right]$$

Recall that the expected return of agents' individual red projects is greater than that of the correlated white project by assumption, $\pi > \pi'$. It follows that all agents will choose their respective red projects, $q^* = 0$. In symmetric equilibrium, $Q = q^* = 0$, consistent with the diversification equilibrium. ■

We'll now consider two quite different examples of our model in which a herding equilibrium exists under nominal output targeting. In both examples, Proposition 5.1 holds, so there is no herding equilibrium under price level targeting. In the first example, banks with nominal deposit liabilities choose portfolio allocations. In the second example, managers with nominal compensation contracts choose investment projects.

5.2 Example 1: Banks

In this example, the agents will be banks. Banks enter the period with a fixed stock of nominal deposit liabilities outstanding D, which are due at the end of the period. Banks hold portfolios of risky loan assets, with nominal repayment rates that are increasing in the price level—it is easier for mortgagors to repay their debts when inflation is high. For simplicity, we'll assume that the nominal revenue of each bank is proportional to the price level. Consequently, bank revenue is equal to $P\theta$. Rather than choosing investment projects as before, here we assume that banks allocate their funds into different loan products. Banks choose the portfolio weight q that is allocated to the *white* product, which is highly correlated with other banks. For example, the white product may be residential mortgage loans, or maturity mismatch as in Farhi and Tirole (2012). The remaining holdings of the bank ($(1 - q)$ share of their portfolio) is allocated to the individual bank's *red* product, which is uncorrelated with the rest of the market. This is considered to be a product in which the bank can develop a specialisation, for example agricultural real estate, trade finance or credit card products.

We assume that the bank's exposure to the red and white projects is not contractable—depositors cannot demand that the bank specialise in their individual red project—but it is important to note that the source of risk to depositors is fluctuations in the price level, which are dependent on the common actions across all banks. Given that deposits are risk free in nominal terms, each depositor is indifferent between the portfolio choices of their own individual bank, which have no discernible effect on the volatility of the price level.

Each individual bank's nominal income is the difference between revenues and deposit liabilities:

$$Z(\theta, P) = \theta P - D. \tag{5.4}$$

Herding in this example will be driven by an aversion to bad outcomes. More generally, any form of weak concavity in the preference function U would be sufficient. For this example we will impose a particular form of risk aversion suitable for considering the incentives facing banks. We will assume that when bank income falls below $\underline{\theta} - D$, a real resource cost of the product of e and the shortfall will be paid. This cost can be considered as the cost of raising additional equity in response to particularly bad outcomes. Bank utility is defined as follows:

$$U\left(\frac{Z(\theta, P(Q, \theta_w))}{P(Q, \theta_w)}\right) = \begin{cases} \dfrac{Z(\theta, P(Q, \theta_w))}{P(Q, \theta_w)} & \text{when} \quad \dfrac{Z(\theta, P(Q, \theta_w))}{P(Q, \theta_w)} \geq \underline{\theta} - D, \\ \dfrac{Z(\theta, P(Q, \theta_w))}{P(Q, \theta_w)} - e\left[\dfrac{(\underline{\theta} - D) - Z(\theta, P(Q, \theta_w))}{P(Q, \theta_w)}\right] & \text{otherwise.} \end{cases} \tag{5.5}$$

Of course, under the price level targeting regime, where $P(Q, \theta_w) = 1$, income will always be sufficiently high to avoid cost e. Under nominal output targeting, the equity cost will be triggered when bank income is low and the price level is also low. The price level will be low when total output is high. It follows that banks can eliminate these equity issuance costs by choosing the white project, which ensures that they receive high real income when the price level is low.

This comes at a cost for depositors. Under inflation targeting, the real value of deposits is constant. Under nominal output targeting, when banks choose correlated white projects the real value of deposits will be subject to risk, decreasing when output is low and increasing when output is high. This result is formalised by Proposition 5.2:

Proposition 5.2 *1. Under price level targeting, the only symmetric equilibrium is the diversification equilibrium. 2. Under nominal output targeting: (a) if* $e > \frac{(\pi - \pi')}{\pi'(1-\pi)(1-\pi')}$ *then there are two equilibria, a diversification equilibrium and a herding equilibrium; (b) otherwise, there is only the diversification equilibrium.*

The proof of 5.2 is contained in Appendix 5.B.

5.3 EXAMPLE 2: MANAGERS

In this example, our agents will be risk tolerant (modeled as risk neutral) managers with nominal compensation contracts that are increasing in the real output of the firm. For example, managers' nominal bonuses may be linked to sales targets. Managers can choose to follow industry trends (the white project) with the result that their output will be highly correlated with the market. Alternatively they may attempt to specialise their output into a niche product (their individual red project) in which case their output will be uncorrelated with that of their peers.

For each type of project, managers must exert some utility cost $e > 0$ in order to attain success probabilities π, π', otherwise projects enjoy success with probability zero. Agents' compensation contracts are incentive compatible if the nominal payoff to agents following successful projects $Z(\bar{\theta})$ is greater than the nominal payoff to agents following unsuccessful projects plus the effort cost scaled up by the expected price level,

$$Z(\bar{\theta}) \geq Z(\underline{\theta}) + \mathbb{E}[P(Q, \theta_w)]e. \tag{5.6}$$

By definitions 5.1 and 5.2 which specify our monetary policy regimes, the expectation of the price level is equal to 1 under both regimes, $\mathbb{E}[P(Q, \theta_w)] = 1$. Assuming that the incentive compatibility constraint is binding, equation 5.6 can be re-written as follows,

$$Z(\bar{\theta}) = Z(\underline{\theta}) + e. \tag{5.7}$$

As the managers' project choices are not verifiable, only project returns θ are contractable. A financial contract is specified by state contingent consumption allocations of the agent, written in nominal terms $(Z(\bar{\theta}), Z(\underline{\theta}))$.

We can re-write (5.1) to reveal each individual manager's objective function,

$$\begin{aligned}\max_{q\in[0,1]} \; \mathbb{E}\; q &\left[\pi' \frac{Z(\bar{\theta})}{P(Q,\bar{\theta})} + (1-\pi')\frac{Z(\underline{\theta})}{P(Q,\underline{\theta})}\right] \\ &+ (1-q)\pi' \left[\pi \frac{Z(\bar{\theta})}{P(Q,\bar{\theta})} + (1-\pi)\frac{Z(\underline{\theta})}{P(Q,\bar{\theta})}\right] \\ &+ (1-q)(1-\pi') \left[\pi \frac{Z(\bar{\theta})}{P(Q,\underline{\theta})} + (1-\pi)\frac{Z(\underline{\theta})}{P(Q,\underline{\theta})}\right].\end{aligned} \tag{5.8}$$

We now solve for symmetric equilibria across monetary policy regimes. Under price level targeting, Proposition 5.1 holds. There is only one equilibrium, where managers diversify across projects. Under nominal output targeting the story changes. As long as the expected return of the white project is not too low, there will be two equilibria. The herding equilibrium where all managers choose the white project, and the diversification equilibrium where all managers choosing their individual red projects.

Proposition 5.3 *1. Under price level targeting, the only symmetric equilibrium is the diversification equilibrium. 2. Under nominal output targeting: (a) if* $\frac{\bar{\theta}}{\underline{\theta}} > \frac{\pi}{1-\pi}\frac{1-\pi'}{\pi'}$*, then there are two equilibria, a diversification equilibrium and a herding equilibrium; (b) otherwise, there is only the diversification equilibrium.*

The proof of Proposition 5.3 is contained in Appendix 5.C.

5.A Proof of Lemma 5.1

Proof. Definition 5.2 part (1) can be re-written as

$$1 = \pi' P(Q, \bar{\theta}) + (1 - \pi') P(Q, \underline{\theta})$$

Rearranging part (2) of Definition 5.2 in terms of $P(Q, \bar{\theta})$ allows us to eliminate this term from the above condition:

$$P(Q, \bar{\theta}) = P(Q, \underline{\theta}) \frac{Y(Q, \bar{\theta})}{Y(Q, \underline{\theta})},$$

$$1 = \pi' P(Q, \underline{\theta}) \frac{Y(Q, \bar{\theta})}{Y(Q, \underline{\theta})} + (1 - \pi') P(Q, \underline{\theta}).$$

Recall from Equation 5.2 that real output is given by

$$Y(Q, \theta_w) = Q\theta_w + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}].$$

Substituting this into the previous condition yields

$$1 = \pi' P(Q, \underline{\theta}) \frac{Q\bar{\theta} + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]}{Q\underline{\theta} + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]} + (1 - \pi') P(Q, \underline{\theta}),$$

$$P(Q, \underline{\theta}) = \frac{Q\underline{\theta} + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]}{Q[\pi'\bar{\theta} + (1 - \pi')\underline{\theta}] + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]}.$$

From here, we can substitute this solution into part (2) of Definition 5.2 to complete the proof:

$$\begin{aligned} P(Q, \bar{\theta}) &= P(Q, \underline{\theta}) \frac{Q\bar{\theta} + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]}{Q\underline{\theta} + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]} \\ &= \frac{Q\bar{\theta} + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]}{Q[\pi'\bar{\theta} + (1 - \pi')\underline{\theta}] + (1 - Q)[\pi\bar{\theta} + (1 - \pi)\underline{\theta}]}. \end{aligned}$$

■

5.B Proof of Proposition 5.2

Proof. Part 1 follows directly from Proposition 5.1. To consider Part 2, first we substitute the conditions for bank income (5.4) and bank preferences (5.5) into the general agents' problem specified by equation 5.1:

$$\max_{q\in[0,1]} \mathbb{E}\, q\left(\pi'\left(\frac{\bar{\theta}P(Q,\bar{\theta})-D}{P(Q,\bar{\theta})}\right)+(1-\pi')\left(\frac{\underline{\theta}P(Q,\underline{\theta})-D}{P(Q,\underline{\theta})}\right)\right)$$
$$+(1-q)\pi'\left[\begin{array}{l}\pi\left(\frac{\bar{\theta}P(Q,\bar{\theta})-D}{P(Q,\bar{\theta})}\right)\\ +(1-\pi)\left(\frac{\underline{\theta}P(Q,\bar{\theta})-D}{P(Q,\bar{\theta})}-e\left[\frac{(\underline{\theta}-D)-[\underline{\theta}P(Q,\bar{\theta})-D]}{P(Q,\bar{\theta})}\right]\right)\end{array}\right]$$
$$+(1-q)(1-\pi')\left[\pi\left(\frac{\bar{\theta}P(Q,\underline{\theta})-D}{P(Q,\underline{\theta})}\right)+(1-\pi)\left(\frac{\underline{\theta}P(Q,\underline{\theta})-D}{P(Q,\underline{\theta})}\right)\right]. \tag{5.9}$$

We can simplify the bankers' problem by eliminating D and price level terms $P(Q,\theta_w)$ where possible,

$$\max_{q\in[0,1]} q\left(\pi'\bar{\theta}+(1-\pi')\underline{\theta}\right)+(1-q)\left[\pi\bar{\theta}+(1-\pi)\underline{\theta}\right]$$
$$-(1-q)\pi'(1-\pi)\underline{\theta}e\,\mathbb{E}\left[\frac{1-P(Q,\bar{\theta})}{P(Q,\bar{\theta})}\right]. \tag{5.10}$$

The first term of equation 5.10 is the expected real return of the white product. The second term is the expected real return of the red product. The third term is the expected real resource cost associated with equity issuance under circumstances when both real revenues and the price level are low. Individual bankers will place a high portfolio weight q on the correlated white project if the difference in expected incomes between red and white projects is smaller than their expected cost of equity issuance in states where their individual red project fails and the price level is low. We can rearrange to solve for q,

$$q^*=1 \text{ if } \left(\pi'\bar{\theta}+(1-\pi')\underline{\theta}\right)>\left[\pi\bar{\theta}+(1-\pi)\underline{\theta}\right]-\pi'(1-\pi)\underline{\theta}e\,\mathbb{E}\left[\frac{1-P(Q,\bar{\theta})}{P(Q,\bar{\theta})}\right] \tag{5.11}$$

and $q^*=0$ otherwise.

We can use the definition of nominal output targeting and equation 5.3 replace $P(Q,\bar{\theta})$ and re-write the condition in terms of Q and parameters. Further rearranging yields

$$q^*=1 \text{ if } e\,\mathbb{E}\left[\frac{Q(1-\pi')(\bar{\theta}-\underline{\theta})}{Q\underline{\theta}+(1-Q)[\pi\bar{\theta}+(1-\pi)\underline{\theta}]}\right]>\frac{(\pi-\pi')(\bar{\theta}-\underline{\theta})}{\pi'(1-\pi)\underline{\theta}} \tag{5.12}$$

and $q^*=0$ otherwise.

A diversification equilibrium requires that $q^*=Q=0$. The right hand side of equation 5.12 is strictly positive. When $\mathbb{E}[Q]=0$, the left hand side is equal to zero for all positive values of e. It follows that $q^*=0$ and the model permits a diversification equilibrium.

A herding equilibrium requires $q^* = Q = 1$. Substituting $Q = 1$ into equation 5.12 reveals that a herding equilibrium can occur when

$$e > \frac{(\pi - \pi')}{\pi'(1-\pi)(1-\pi')}. \tag{5.13}$$

When the correlated white portfolio offers returns which are close to the higher returning red portfolio (that is, when $\pi - \pi'$ is small), the herding equilibrium is permitted even when risk aversion or the resource costs of equity issuance e are small. ■

5.C Proof of Proposition 5.3

Proof. Part 1 holds by Proposition 5.1. The proof of part 2 follows.

First, use the incentive compatibility constraint (5.7) to eliminate $Z(\bar{\theta})$ from the managers' objective function (5.8). Then eliminate constants $Z(\underline{\theta}), e$ to obtain:

$$\max_{q\in[0,1]} \mathbb{E} \quad q\frac{\pi'}{P(Q,\bar{\theta})} + (1-q)\pi\left[\frac{\pi'}{P(Q,\bar{\theta})} + \frac{(1-\pi')}{P(Q,\underline{\theta})}\right]. \tag{5.14}$$

It follows that

$$q^* = 1 \quad \text{if} \quad \mathbb{E}\,\frac{\pi'}{P(Q,\bar{\theta})} > \mathbb{E}\,\pi\left[\frac{\pi'}{P(Q,\bar{\theta})} + \frac{(1-\pi')}{P(Q,\underline{\theta})}\right], \tag{5.15}$$

and $q^* = 0$ otherwise. Rearranging yields

$$q^* = 1 \quad \text{if} \quad \mathbb{E}\,\frac{P(Q,\underline{\theta})}{P(Q,\bar{\theta})} > \frac{\pi}{1-\pi}\frac{1-\pi'}{\pi'} \tag{5.16}$$

and $q^* = 0$ otherwise. Substituting in the monetary policy rule specified in equation 5.3 yields

$$q^* = 1 \quad \text{if} \quad \mathbb{E}\,\frac{Q\bar{\theta} + (1-Q)[\pi\bar{\theta} + (1-\pi)\underline{\theta}]}{Q\underline{\theta} + (1-Q)[\pi\bar{\theta} + (1-\pi)\underline{\theta}]} > \frac{\pi}{1-\pi}\frac{1-\pi'}{\pi'} \tag{5.17}$$

and $q^* = 0$ otherwise. After rearranging we obtain

$$q^* = 1 \quad \text{if} \quad \mathbb{E}\,[Q] > 1 + \frac{\underline{\theta}\dfrac{\pi}{1-\pi}\dfrac{1-\pi'}{\pi'} - \bar{\theta}}{(\bar{\theta}-\underline{\theta})\left(1 + \pi\left[\dfrac{\pi}{1-\pi}\dfrac{1-\pi'}{\pi'} - 1\right]\right)} \tag{5.18}$$

and $q^* = 0$ otherwise. Given that $Q \in [0,1]$, $q^* = Q = 1$ can only be a symmetric equilibrium when the second term on the right hand side is negative. This occurs when $\frac{\bar{\theta}}{\underline{\theta}} > \frac{\pi}{1-\pi}\frac{1-\pi'}{\pi'}$. If so, then $q^* = 1$ when $\mathbb{E}[Q] = 1$. That is, a herding equilibrium will exist. Otherwise, the only equilibrium is the diversification equilibrium. ■

BIBLIOGRAPHY

Salem Abo-Zaid. On credit frictions as laborincome taxation. *Economics Letters*, 118(2): 287–292, 2013.

Philippe Aghion and Patrick Bolton. An incomplete contracts approach to financial contracting. *Review of Economic Studies*, 59(3):473–94, 1992.

David Alary and Christian Gollier. Debt contract, strategic default, and optimal penalties with judgement errors. *Annals of Economics and Finance*, 5(2):357–372, November 2004.

Edward Altman. A further empirical investigation of the bankruptcy cost question. *Journal of Finance*, 39(4):1067–89, 1984.

Guido Ascari and Lorenza Rossi. Trend Inflation and Firms PriceSetting: Rotemberg Versus Calvo. *Economic Journal*, 122(563):1115–1141, 09 2012.

Guido Ascari, Efrem Castelnuovo, and Lorenza Rossi. Calvo vs. Rotemberg in a trend inflation world: An empirical investigation. *Journal of Economic Dynamics and Control*, 35(11):1852–1867, 2011.

Szilrd Benk, Max Gillman, and Michal Kejak. Credit Shocks in the Financial Deregulatory Era: Not the Usual Suspects. *Review of Economic Dynamics*, 8(3):668–687, July 2005.

Ben S. Bernanke, Mark L. Gertler, and Simon Gilchrist. The financial accelerator in a quantitative business cycle framework. In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1, Part C, chapter 21, pages 1341–1393. Elsevier, 1 edition, 1999.

Kim C Border and Joel Sobel. Samurai accountant: A theory of auditing and plunder. *Review of Economic Studies*, 54(4):525–40, October 1987.

Charles W. Calomiris and Urooj Khan. An assessment of tarp assistance to financial institutions. *Journal of Economic Perspectives*, 29(2):53–80, 2015.

Linda Cameron. Investor Protection and the New Zealand Stock Market. Policy Perspectives Paper 07/02, New Zealand Treasury, October 2007.

Charles T Carlstrom and Timothy S Fuerst. Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis. *American Economic Review*, 87(5):893–910, December 1997.

Charles T. Carlstrom, Timothy S. Fuerst, and Matthius Paustian. Optimal Contracts, Aggregate Risk, and the Financial Accelerator. Working Paper 1420, Federal Reserve Bank of Cleveland, October 2014.

V. V. Chari and Christopher Phelan. Too Correlated to Fail. Economic Policy Paper 14-3, Federal Reserve Bank of Minneapolis, July 2014.

V. V. Chari, Patrick J. Kehoe, and Ellen R. McGrattan. Business Cycle Accounting. *Econometrica*, 75(3):781–836, 05 2007.

Lawrence Christiano and Daisuke Ikeda. Government Policy, Credit Markets and Economic Activity. NBER Working Papers 17142, National Bureau of Economic Research, Inc, June 2011.

Lawrence J. Christiano, Roberto Motto, and Massimo Rostagno. Risk Shocks. *American Economic Review*, 104(1):27–65, January 2014.

Harold L Cole and Narayana R Kocherlakota. Efficient Allocations with Hidden Income and Hidden Storage. *Review of Economic Studies*, 68(3):523–42, July 2001.

Mariacristina De Nardi, Eric French, and David Benson. Consumption and the Great Recession. *Economic Perspectives*, (Q I):1–16, 2012.

Marco Del Negro, Marc P. Giannoni, and Frank Schorfheide. Inflation in the Great Recession and New Keynesian Models. NBER Working Papers 20055, National Bureau of Economic Research, Inc, April 2014.

Andrea Devenow and Ivo Welch. Rational herding in financial economics. *European Economic Review*, 40(3-5):603–615, April 1996.

Douglas W Diamond and Philip H Dybvig. Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy*, 91(3):401–19, June 1983.

Peter Diamond. Aggregate demand management in search equilibrium. *Journal of Political Economy*, 90(5):881–94, 1982.

Avinash Dixit. Whither Greenwald-Stiglitz? Conference presentation, October 2003.

Alfred Duncan and Charles Nolan. Disputes, Debt and Equity. Discussion Paper 2014-21, Adam Smith Business School, University of Glasgow, December 2014.

Tore Ellingsen and Eirik Gaard Kristiansen. Financial contracting under imperfect enforcement. *The Quarterly Journal of Economics*, 126(1):323–371, 2011.

Emmanuel Farhi and Jean Tirole. Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts. *American Economic Review*, 102(1):60–93, February 2012.

Emmanuel Farhi and Ivan Werning. A Theory of Macroprudential Policies in the Presence of Nominal Rigidities. NBER Working Papers 19313, National Bureau of Economic Research, Inc, August 2013.

Fabio Fornari and Livio Stracca. What does a financial shock do? First international evidence. *Economic Policy*, 27(71):407–445, 07 2012.

Douglas Gale and Martin Hellwig. Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies*, 52(4):647–63, October 1985.

Jordi Galí. *Introduction to Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Introductory Chapters. Princeton University Press, 2008.

Julio Garín, Robert Lester, and Eric Sims. On the desirability of nominal gdp targeting. NBER Working Papers 21420, National Bureau of Economic Research, Inc, 2015.

John Geanakoplos and H. M. Polemarchakis. Pareto improving taxes. Cowles Foundation Discussion Papers 1662, Cowles Foundation for Research in Economics, Yale University, 2008.

Simon Gilchrist and Egon Zakrajsek. Credit Spreads and Business Cycle Fluctuations. *American Economic Review*, 102(4):1692–1720, June 2012.

Marvin Goodfriend and Robert King. The New Neoclassical Synthesis and the Role of Monetary Policy. In *NBER Macroeconomics Annual 1997, Volume 12*, NBER Chapters, pages 231–296. National Bureau of Economic Research, Inc, June 1997.

Austan D. Goolsbee and Alan B. Krueger. A retrospective look at rescuing and restructuring general motors and chrysler. *Journal of Economic Perspectives*, 29(2):3–24, 2015.

Edward J. Green and Soo-Nam Oh. Can a "credit crunch" be efficient? *Quarterly Review*, (Fall):3–17, 1991.

Bruce C Greenwald and Joseph E Stiglitz. Externalities in Economies with Imperfect Information and Incomplete Markets. *The Quarterly Journal of Economics*, 101(2): 229–64, May 1986.

Robert E. Hall. The Cyclical Response of Advertising Refutes Counter-Cyclical Profit Margins in Favor of Product-Market Frictions. NBER Working Papers 18370, National Bureau of Economic Research, Inc, September 2012.

Robert Ernest Hall and N. Gregory Mankiw. Nominal income targeting. In *Monetary Policy*, pages 71–94. National Bureau of Economic Research, Inc, 1994.

Gill Hammond. *State of the art of inflation targeting*. Number 29 in Handbooks. Centre for Central Banking Studies, Bank of England, 2012.

Oliver Hart and John Moore. Incomplete contracts and renegotiation. *Econometrica*, 56 (4):755–85, 1988.

Joseph G. Haubrich. Imperfect state verification and financial contracting. Working Paper 9506, Federal Reserve Bank of Cleveland, 1995.

Michael C Jensen. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review*, 76(2):323–29, May 1986.

Urban Jermann and Vincenzo Quadrini. Macroeconomic Effects of Financial Shocks. *American Economic Review*, 102(1):238–71, February 2012.

Sebnem Kalemli-Ozcan, Bent Sorensen, and Sevcan Yesiltas. Leverage across firms, banks, and countries. *Journal of International Economics*, 88(2):284–298, 2012.

Weerachart T. Kilenthong and Robert M. Townsend. A Market Based Solution to Price Externalities: A Generalized Framework. NBER Working Papers 20275, National Bureau of Economic Research, Inc, July 2014.

Nobuhiro Kiyotaki. A mechanism design approach to financial frictions. September 2010.

Nobuhiro Kiyotaki and John Moore. Credit Cycles. *Journal of Political Economy*, 105 (2):211–48, April 1997.

Frank H. Knight. *Risk, Uncertainty, and Profit*. Hart, Schaffner & Marx; Houghton Mifflin Co., Boston, MA, 1921.

Evan F. Koenig. Monetary policy, financial stability, and the distribution of risk. Technical report, 2011.

Stefan Krasa and Anne P. Villamil. Optimal contracts when enforcement is a decision variable. *Econometrica*, 68(1):119–134, January 2000.

Michael U. Krause, David Lopez-Salido, and Thomas A. Lubik. Inflation dynamics with search frictions: A structural econometric analysis. *Journal of Monetary Economics*, 55(5):892–916, July 2008.

Arvind Krishnamurthy. Collateral constraints and the amplification mechanism. *Journal of Economic Theory*, 111(2):277–292, 2003.

Campbell Leith and Ding Liu. The Inflation Bias under Calvo and Rotemberg Pricing. SIRE Discussion Papers 2014-018, Scottish Institute for Research in Economics (SIRE), 2014.

Jonathan Macey and Joshua Mitts. Finding order in the morass: The three real justifications for piercing the corporate veil. *Cornell Law Review*, 100(1):99–156, 2014.

Bennett T. McCallum and Edward Nelson. Nominal income targeting in an open-economy optimizing model. *Journal of Monetary Economics*, 43(3):553–578, June 1999.

Ellen McGrattan and Edward Prescott. Taxes, regulations, and the value of u.s. and u.k. corporations. *Review of Economic Studies*, 72(3):767–796, 2005.

Nahum D. Melumad and Dilip Mookherjee. Delegation as commitment: The case of income tax audits. *RAND Journal of Economics*, 20(2):139–163, 1989.

Bruce D. Meyer and James X. Sullivan. Consumption and Income Inequality and the Great Recession. *American Economic Review*, 103(3):178–83, May 2013.

Atif Mian and Amir Sufi. *House of Debt*. University of Chicago Press, 1 edition, 2015.

Atif R. Mian and Amir Sufi. What explains high unemployment? The aggregate demand channel. NBER Working Papers 17830, National Bureau of Economic Research, Inc, February 2012.

Dilip Mookherjee and Ivan Png. Optimal auditing, insurance, and redistribution. *The Quarterly Journal of Economics*, 104(2):399–415, May 1989.

Dale Mortensen and Christopher Pissarides. Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, 61(3):397–415, 1994.

Christopher J. Nekarda and Valerie A. Ramey. The Cyclical Behavior of the Price-Cost Markup. NBER Working Papers 19099, National Bureau of Economic Research, Inc, June 2013.

Kalin Nikolov. Collateral amplification under complete markets. *Journal of Economic Dynamics and Control*, 45(C):80–93, 2014.

Charles Nolan and Christoph Thoenissen. Financial shocks and the US business cycle. *Journal of Monetary Economics*, 56(4):596–604, May 2009.

NZHR. Inquiry into finance company failures, 2011. Report of the New Zealand House of Representatives Commerce Committee. Forty-ninth Parliament. Lianne Dalziel, Chairperson.

Gabor Pinter, Konstantinos Theodoridis, and Tony Yates. Risk news shocks and the business cycle. Bank of England working papers 483, Bank of England, December 2013.

A. Mitchell Polinsky and Steven Shavell. *The Theory of Public Enforcement of Law*, volume 1 of *Handbook of Law and Economics*, chapter 6, pages 403–454. Elsevier, 2007.

Julio J Rotemberg. Sticky Prices in the United States. *Journal of Political Economy*, 90 (6):1187–1211, December 1982.

Julio J. Rotemberg. Cyclical Wages in a Search-and-Bargaining Model with Large Firms. NBER Working Papers 12415, National Bureau of Economic Research, Inc, August 2006.

Til Schuermann and Samuel Hanson. Estimating probabilities of default. Staff Reports 190, Federal Reserve Bank of New York, 2004.

Moritz Schularick and Alan M. Taylor. Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review*, 102 (2):1029–61, April 2012.

Kevin D. Sheedy. Debt and Incomplete Financial Markets: A Case for Nominal GDP Targeting. *Brookings Papers on Economic Activity*, 48(1 (Spring):301–373, 2014.

Robert J. Shiller. Derivatives Markets for Home Prices. NBER Working Papers 13962, National Bureau of Economic Research, Inc, April 2008.

Adam Smith. *An Inquiry into the Nature and Causes of the Wealth of Nations*. 1776. 5th Edition published 1904.

Robert M. Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265–293, October 1979.

Neil Wallace. The mechanism-design approach to monetary theory. In Benjamin M. Friedman and Michael Woodford, editors, *Handbook of Monetary Economics*, volume 3, chapter 01, pages 3–23. Elsevier, 1 edition, 2010.

Jerold B Warner. Bankruptcy costs: Some evidence. *Journal of Finance*, 32(2):337–47, 1977.

Lawrence A. Weiss. Bankruptcy resolution: Direct costs and violation of priority of claims. *Journal of Financial Economics*, 27(2):285–314, 1990.

Michael Woodford. *Interest and Prices*. Princeton University Press, Princeton, NJ, 2003.

Francesco Zanetti. Financial Shocks and Labor Market Fluctuations. Economics Series Working Papers Number 746, University of Oxford, Department of Economics, May 2015.