

DISCUSSION PAPER SERIES

No. 7636

VERTICAL RELATIONS UNDER CREDIT CONSTRAINTS

Volker Nocke and John Thanassoulis

INDUSTRIAL ORGANIZATION



Centre for **E**conomic **P**olicy **R**esearch

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP7636.asp

VERTICAL RELATIONS UNDER CREDIT CONSTRAINTS

**Volker Nocke, University of Mannheim, Oxford University, CESifo and CEPR
John Thanassoulis, Oxford University**

Discussion Paper No. 7636
January 2010

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Volker Nocke and John Thanassoulis

ABSTRACT

Vertical Relations under Credit Constraints

We model the impact credit constraints and market risk have on the vertical relationships between firms in the supply chain. Firms which might face credit constraints in future investments become endogenously risk averse when accumulating pledgable income. In the short run, the optimal supply contract therefore involves risk sharing, thereby inducing double marginalization. Credit constraints thus result in higher retail prices. The model offers a concise explanation for several empirical regularities of firm behavior. We demonstrate an intrinsic complementarity between supply and lending providing a theory of finance arms of major suppliers; a monetary transmission mechanism linking the cost of borrowing with short-run retail prices that can help explain the price puzzle in macroeconomics; a theory of countervailing power based on credit constraints; and a motive for outsourcing supply (or distribution) in the face of market risk.

JEL Classification: G32, L14 and L16

Keywords: countervailing power, double marginalization, finance arms, financial companies, market risk, monetary transmission mechanism, outsourcing, price puzzle, risk aversion, risk sharing and vertical contracting

Volker Nocke
Department of Economics
University of Mannheim
L7, 3-5
D-68131 Mannheim
GERMANY

John Thanassoulis
Department of Economics
Oxford University
Manor Road
Oxford OX1 3UL
UK

Email: nocke@uni-mannheim.de

Email:
john.thanassoulis@economics.ox.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=141921

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=161742

Submitted 07 January 2010

We would like to thank Tracy Lewis, Meg Meyer, Jean Tirole, Mike Whinston, and seminar participants at Columbia, Leuven, Northwestern, Oxford, Strathclyde, Texas-Austin, UCLA, UPF, Warwick, the 2009 CEPR Applied IO Conference in Mannheim, the 2009 EARIE Conference in Ljubljana, and the 2009 Duke-Northwestern-Texas IO Theory Conference in Evanston. We are also grateful to the ESRC for financial support (grant RES-000-22-3468).

1 Introduction

Credit constraints have been known to be a part of corporate reality for decades (Hubbard, 1998, and references therein). Massively reduced access to credit has been a feature of the major financial crisis of recent years. It is also well known that firms are subject to substantial market risk – whether on the demand side or supply side. Incorporating insights from the corporate finance literature into an industrial organization model of the vertical supply chain, we study the interaction between credit constraints and market risk, and their effects on short-run retail pricing, long-run investment, and welfare. We show that credit constraints and market risk impact optimal vertical contracting, creating scope for double marginalization, slotting fees, finance arms, and outsourcing. Further, we identify a new monetary transmission mechanism from interest rates to the real economy which acts via firms which are at risk of becoming credit constrained. Finally, the model gives rise to a novel theory of countervailing power based on credit constraints.

Consider a vertical supply chain consisting of a single upstream firm (“he”) supplying a single downstream firm (“she”), and exposed to demand-side risk. The joint-profit maximizing supply contract would involve per unit input prices at the upstream firm’s marginal cost, irrespective of any demand-side risk. But now suppose the downstream firm has some future investment opportunities. The size of the loan she is able to raise to fund the investment, and therefore the actual investment level, depend on the size of the pledgable assets the firm owns. Under the standard assumption that investment is subject to diminishing marginal returns, we show that the profit-maximizing firm becomes endogenously risk averse when accumulating pledgable assets. When pledgable assets are low, the induced investment level is low as well. This implies that the return on the marginal dollar of investment is high, and so the marginal dollar of pledgable assets can be greatly levered through the banking sector.

As a result, the optimal contract between the downstream firm and its upstream supplier involves risk sharing and, hence, double marginalization. The endogenously risk-averse downstream firm wants to insure her level of pledgable income. So she demands a risk-sharing contract in which the supplier bears some loss for poor demand realizations. But for the supplier to recoup these potential losses, he requires payments in high demand states to grow at a rate faster than cost. That is, double marginalization is introduced, causing the retail price of the downstream firm to rise. The cost of the insurance made necessary by the credit constraints is in this sense partly paid for by final consumers.

The optimal supply contract can be thought of as involving a fixed payment from the upstream to the downstream firm and demand-dependent repayments. This may

help explain the increasingly common use of “slotting fees” in the grocery market as well as in other industries such as software and publishing. These fees are fixed payments many retailers require of manufacturers in return for stocking their products.¹ Empirical evidence suggests that an important part of the story is the sharing of risk (Sudhir and Rao, 2006; White et al. 2000), which accords with our model.

It is standard to see the input suppliers and the banking sector as two completely separate industries. However, if the input supplier also provides pledgable income insurance, as in our model, it is no longer clear whether such a separation is indeed optimal. In fact, we demonstrate that there exists an intrinsic complementarity between the provision of insurance and lending. An input supplier with access to funds at the same rate as the banking sector could actually lend on rates that the independent banking sector would find unprofitable. This result may offer an original insight into the existence and profitability of finance arms of major companies such as GE and Cisco. As financial companies lend almost \$1 for every \$2 lent by a mainstream bank, gaining an insight into what makes financial companies effective competitors to banks therefore seems a first-order issue.

The complementarity we find between supply insurance and lending arises because of the countervailing incentives the downstream firm faces when dealing with the insurer and the lender. By pooling insurance and lending, the downstream firm can effectively reduce her temptation to under-report the demand state, which allows for less double marginalization and therefore higher profits.

As the pre-investment degree of risk aversion of the downstream firm is endogenous, it is a function of market-level and firm-level parameters such as the interest rate, the quality of corporate governance, the firm’s asset endowment and her bargaining power in the vertical chain. We demonstrate that if parameters change so as to increase (decrease) the coefficient of absolute risk aversion, then retail prices will rise (fall) in the short run.

We show that an increase in the interest rate that the downstream firm has to pay to finance her investment makes the firm more risk averse when accumulating pledgable assets. Hence, a higher interest rate leads to higher retail prices in the short run and lower investment levels in the long run. This result is a new insight into the *price puzzle*: the macroeconomic link that has been noted between increases in the interest rate and increases in retail prices (Christiano et al., 1999).

Relaxing our assumption that the downstream firm has all of the bargaining power, we show that an increase in the downstream firm’s bargaining power *vis-à-vis* her supplier makes the firm less risk averse when accumulating pledgable assets. Hence, a more

¹Theoretical explanations for this practice have portrayed the slotting fee as a signalling device (Klein and Wright, 2007, and references therein).

powerful downstream firm charges lower retail prices in the short run and invests more in the long run. The model therefore gives rise to a new theory of *countervailing power* (Galbraith, 1952) that is based on credit constraints.

We finally demonstrate a link between market risk and outsourcing. A credit-constrained downstream firm cannot insure herself. By outsourcing input supply, however, the downstream firm can purchase insurance as the upstream supplier is in a unique position to monitor the volumes supplied to the downstream firm. Our result is supported by empirical evidence (Harrigan, 1985; Sutcliffe and Zaheer, 1988) which points in this direction.²

Related Literature. Our paper builds on some existing insights from the industrial organization and corporate finance literatures. On the corporate finance side, we build on Holmstrom and Tirole (1997) in modelling credit constraints as an endogenous outcome, caused by a moral hazard problem associated with the firm's investment project. In contrast to Holmstrom and Tirole, however, we assume that the firm's investment project has decreasing returns. It is this decreasing-returns assumption that gives rise to the firm becoming endogenously risk averse as it means that the extent at which the marginal dollar can be leveraged is decreasing. A related point, but in a model with exogenous credit constraints, is made by Froot et al. (1993). Froot et al. do not study the implications this insight has for vertical contracting and, hence, for pricing and the real economy.

Firms might borrow from their suppliers either for investment purposes (usually through finance arms) or by implication of making late repayments for inputs received (known as "trade credit"). While finance arms are arguably not well understood, there is a literature studying trade credit. Burkart and Ellingsen (2004) note that goods are less divertable to private benefits than money and so they argue that trade credit and bank lending are complementary. Cuñat (2007) suggests that a supplier can enforce repayment as he has a long-term relationship with a downstream buyer over inputs which cannot be supplied by another. Cuñat therefore argues that this long-term relationship can make a supplier able to provide liquidity insurance to a firm which a bank would not. We show that even without any exogenous advantage over banks in terms of the longevity of the relationship, or the divertability of the loan, there is a complementarity between supplier insurance and lending for investment purposes.

On the industrial organization side it is probably fair to say that the literature has been skeptical about the assumption that firms are risk averse. It is perhaps more accepted that small owner-managed firms might inherit the risk aversion of the owner; but, in general,

²The main theoretical arguments in the extant literature have had difficulty with this empirical evidence as they work in the opposite direction. These theories commonly cite problems of incomplete contracting, which mandate integration in the face of risk to save on contracting costs (Mahoney, 1992).

it has proved harder to justify why risk aversion should apply to firms with dispersed ownership. Such explanations would typically require that managers' interests cannot be fully aligned with those of the owners. Here, we demonstrate that risk aversion can result even without such a separation of goals between owners and managers.

Assuming exogenously risk-averse downstream firms, Rey and Tirole (1986) show that the best two-part tariff contract involves double marginalization. Our work demonstrates that risk aversion would be expected if the firms are credit constrained and that double marginalization results not only from two-part tariff contracts but even from the fully optimal contract. More importantly, however, we demonstrate that credit constraints (1) lead to an intrinsic complementarity between supplier insurance and lending (which may explain the existence of finance arms); (2) create a new transmission mechanism linking interest rates with short-run pricing (which can help explain the *price puzzle* in macroeconomics); and (3) provide a theory of countervailing power which predicts that more powerful retailers will charge lower prices.

Plan of the Paper. The model is introduced in Section 2. We consider a credit-constrained downstream firm facing demand-side risk. The paper could easily be rewritten to consider a credit-constrained upstream firm facing supply-side risk. The model is solved and the optimal supply contract characterized in Section 3. The complementarity between lending and insurance is analyzed in Section 4. The link between interest rates and retail prices is studied in Section 5. The theory of countervailing power induced by credit constraints is presented in Section 6. The incentive to outsource production due to market risk is demonstrated in Section 7. Finally, Section 8 concludes, with all omitted proofs contained in the Appendix.

2 The Model

We consider a model of a vertically related industry with two firms, a downstream firm D and an upstream firm U . There are two periods: period 0 and period 1.

Period 0. In period 0, U can produce an intermediate input at marginal cost $c \geq 0$. U supplies the input to D which D transforms into a final good on a one-to-one basis at zero cost, and then sells on. When choosing output Q and facing market size z , D faces inverse demand $p(Q/z)$.³ We assume that D is exposed to market risk in that market size z is a random variable with finite support $\{z_1, \dots, z_n\}$. A larger value of z implies that the volume supplied is a smaller proportion of the total market, and so a higher unit price results. We label states in increasing order so that $0 < z_1 < z_2 < \dots < z_n$.

³ D can equivalently be thought of as setting price p and facing demand $zQ(p)$.

Assumption 1 *We make the following standard assumptions on downstream demand:*

- (i) *Marginal revenue $d[Qp(Q/z)]/dQ$ is declining in quantity Q .*
- (ii) *The reservation price exceeds marginal cost at $Q = 0$, $p(0) > c$, and falls below marginal cost, $P(Q) < c$, for Q sufficiently large.*

Assumption 1 implies that, in any demand state z , industry profit $Q[p(Q/z) - c]$ is strictly concave in quantity Q . Moreover, it implies that, in demand state z , industry profit is maximized at quantity $Q = zq(c)$, where $q(c)$ is the unique solution in q to $p(q) + qp'(q) = c$. The downstream price that maximizes industry profit is $p(q(c))$ in every demand state z .

Before the demand state is realized, D offers U a contract of the form $\{Q(z_i), W(z_i)\}$, where $Q(z_i)$ is the input (and output) volume in state z_i , and $W(z_i)$ the associated transfer payment from D to U ; if U rejects D 's offer, both firms make zero profit. (That is, we assume for now that D has all of the bargaining power.) Then, D privately learns the realization of the demand state z and reports state \hat{z} to U . U , for his part, cannot verify the state of demand. D then receives $\hat{Q} = Q(\hat{z})$ units of input from U , transforms the input into a final good, and fetches a retail price of $p(\hat{Q}/z)$ per unit. Finally, D pays $W(\hat{z})$ to U . We assume for notational simplicity that D has no initial assets.⁴ D 's asset level by the end of period 0, a , is therefore given by D 's net profit in that period: $a = \hat{Q}p(\hat{Q}/z) - W(\hat{z})$.

Period 1. In period 1, D has to decide how much to invest in a project. Based on the moral hazard formulation offered by Holmstrom and Tirole (1997), we assume that D is endogenously credit constrained. Specifically, after choosing the investment level I , D can choose whether or not to shirk at the investment stage. If she does not shirk, D makes a gross profit of $\pi(I)$. If she does shirk, instead, the investment project fails and yields a payoff of zero while D receives a benefit proportional to the size of the investment, $B \cdot I$, where $B \leq 1$. (For simplicity, we do not model U 's post-investment role explicitly. Implicitly, we assume here that post-investment either the then risk-neutral D demands her input at marginal cost c or that she does not require U for the investment returns.)

If D wishes to invest more than her pledgable assets, $I > a$, she can choose to verifiably show her asset level a to an external banking sector so as to attempt to secure a loan of $I - a$.⁵ For now, we set the market interest rate to zero so that D has to pay back only

⁴An increase in D 's initial asset level is akin to an increase in D 's bargaining power; see Section 6.

⁵ D can always choose to hide some or all of her assets. As a result, D can only prove that she has *at least* the asset level that she reveals.

the amount of the loan, $I - a$. Any loan has to satisfy the no-shirking condition

$$BI \leq \pi(I) - (I - a) \quad (1)$$

since, otherwise, D would decide to shirk and so would be unable to pay back her loan.

Assumption 2 *We make the following assumptions on the gross return function $\pi(\cdot)$:*

- (i) *The marginal gross return of investment is positive but diminishing: $\pi(I)$ is strictly increasing and strictly concave in I . Further, $\pi'(0) > 1$, and $\pi'(I) < 1$ for I sufficiently large, so that the first-best level of investment, $\hat{I} \equiv \arg \max_I \pi(I) - I$, is strictly positive.*
- (ii) *In equilibrium, any realized value of D 's asset level a is smaller than the level necessary to finance the first-best investment level, $a < (B + 1)\hat{I} - \pi(\hat{I})$. That is, the no-shirking constraint (1) is always binding in equilibrium.⁶*

3 Equilibrium Analysis

We solve the model by backward induction. Suppose D 's asset level at the beginning of period 1 is given by a . By Assumption 2(ii), D chooses an investment level $I(a)$ and an associated loan $I(a) - a$ so that the no-shirking constraint is just binding: while D would like to invest more, the banking sector would be unwilling to provide a larger loan. That is, $I(a)$ is the unique solution in I to

$$BI = \pi(I) - (I - a). \quad (2)$$

Note that Assumption 2(ii) also ensures that at $I(a)$ the marginal gross return satisfies

$$1 < \pi'(I(a)) < 1 + B. \quad (3)$$

The first inequality follows as the investment level is below the first-best level. The second inequality is an implication of credit being constrained at $I(a)$. Since the no-shirking constraint is binding, D 's net payoff at the end of the second period is $\pi(I(a)) - [I(a) - a] \equiv BI(a)$. The following lemma holds:

⁶The assumption that the no-shirking constraint is always binding is for convenience. What is really needed for our main results is that the constraint is binding in the worst demand state(s).

Lemma 1 *D's net payoff, $BI(a) \equiv \pi(I(a)) - [I(a) - a]$, is (i) increasing at a rate greater than B and (ii) strictly concave in the pledgable asset level a .*

Proof. Implicitly differentiating $I(a)$ in equation (2) yields

$$\frac{dI}{da} = \frac{1}{1 + B - \pi'(I)} > 1 \quad \text{and} \quad \frac{d^2I(a)}{da^2} = \frac{\pi''(I) \left[\frac{dI}{da}\right]^2}{1 + B - \pi'(I)} < 0,$$

where the inequalities follow from equation (3) and Assumption 2(i). ■

This is a key preliminary result. It shows that the interaction of credit constraints and diminishing marginal returns to investment make firm D endogenously risk averse with respect to changes in her pledgable asset level a . To get some intuition, suppose that D were not credit constrained. In period 1, it could therefore borrow \hat{I} so as to capture the first-best profit of $\pi(\hat{I}) - \hat{I}$, independently of the realization of a . D would therefore be risk-neutral with respect to end of period 0 assets a . However in fact D is (endogenously) credit constrained. If D can get a loan from the banking sector, then $I(a) - a$ is positive. The positive marginal returns to investment implies that each extra dollar in pledgable income can be leveraged so that $I(a) - a$ is increasing in a . Since marginal returns are diminishing, the rate at which the marginal dollar can be leveraged is decreasing, implying that $d^2[I(a) - a]/da^2 < 0$.

It is probably fair to say that the IO literature has been skeptical about the assumption that firms are risk averse. It is perhaps more accepted that small owner-managed firms might inherit the risk aversion of the owner; but in general it has proved harder to justify why risk aversion should apply to firms with dispersed ownership. Any such explanation would require that managers' interests cannot be fully aligned with those of the owners. The mechanism generating risk aversion in this paper does not require such a separation of goals between owners and managers and may therefore be more generally applicable. (Recall from Footnote 6 that for the mechanism to work all that is required is that D would be credit constrained in the worst period-0 demand state(s).)

The risk aversion will affect the agreement D requires from her supplier U . This will in turn affect the retail prices in period 0 (the "short run") and the expected level of investment in period 1 (the "long run"). Thus credit constraints will – via the supply-chain relationship – affect consumer welfare both in the short and long run. We now determine how.

We now analyze period-0 contracting. If the state is z_i and D truthfully reports it, then she would receive a payoff of $BI(Q_i p(Q_i/z_i) - W_i)$. Suppose instead D were to lie and claim that the state is z_j , thereby requesting volume Q_j in exchange for payment W_j .

This would mean that the retail price received by D would be $p(Q_j/z_i)$. This yields D pledgable income of $a = Q_j p(Q_j/z_i) - W_j$ at the end of period 0. Invoking the Revelation Principle, we restrict attention to contracts that maximize D 's pledgable income when the truth is being told:

Program Bank The optimization program when D uses an independent banking sector is given by

$$\max_{\{Q_i, W_i\}} \sum_{i=1}^n g_i B \cdot I \left(Q_i p \left(\frac{Q_i}{z_i} \right) - W_i \right)$$

subject to the individual rationality constraint for U ,

$$\sum_{i=1}^n g_i \{W_i - Q_i c\} \geq 0, \quad (4)$$

and the incentive constraint at the quantity setting stage for D ,

$$Q_i p \left(\frac{Q_i}{z_i} \right) - W_i \geq Q_j p \left(\frac{Q_j}{z_i} \right) - W_j \text{ for all } j \neq i. \quad (5)$$

This problem is isomorphic to one explored by Hart (1983) in the context of optimal labor contracts. U here maps to workers (the marginal cost c corresponding to workers' reservation wage) in Hart's analysis and D maps to a firm demanding labor specifically. The following proposition then follows:

Proposition 1 (Hart, 1983, Proposition 2) *The solution to Program Bank, $\{Q_i^*, W_i^*\}_{i=1}^n$, has the following properties:*

Property 1 *There is no distortion at the top: $\frac{\partial}{\partial Q} \left[Q_n^* p \left(\frac{Q_n^*}{z_n} \right) \right] = c$.*

Property 2 *There is inefficiently low quantity demanded in all other states:*

$$\frac{\partial}{\partial Q} \left[Q_i^* p \left(\frac{Q_i^*}{z_i} \right) \right] > c \text{ for all } i < n. \quad (6)$$

Property 3 *D 's pledgable income increases in the state:*

$$Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - W_i \geq Q_{i-1}^* p \left(\frac{Q_{i-1}^*}{z_{i-1}} \right) - W_{i-1} \text{ for all } i > 1.$$

Property 4 *U 's payoff increases in the state:*

$$W_i - Q_i^* c \geq W_{i-1} - Q_{i-1}^* c \text{ for all } i > 1. \quad (7)$$

Proof. Hart (1983) yields all four conditions.⁷ We have a strict inequality in his second condition as U is risk neutral here. ■

Note that the optimal contract involves risk-sharing: both U and D are better off in better demand states (Properties 3 and 4). By exploring a general input into a downstream firm D , we obtain important corollaries of the above proposition:

Corollary 1 *The optimal contract with a supplier U when D is subject to credit constraints and market risk results in:*

1. *Retail prices are too high relative to the level that would maximize joint period-0 profit in all except the best demand state. That is, the optimal contract induces double marginalization.*
2. *The optimal contract has the supplier making payments to D which are not recouped in low demand states. Hence, if marginal cost c is sufficiently small, $W(z_i)$ is negative for small realized demand states z_i and positive for large z_i .*

Proof. For part 1, note that equation (6) guarantees that the marginal revenue is above marginal cost at all demand states except for the highest. Hence, as marginal revenue is declining, we must have quantities being below (and, thus, retail prices being above) the industry-profit maximizing levels.

For part 2, note that U 's individual rationality constraint is binding, $\sum_{i=1}^n g_i \{W_i^* - Q_i^*c\} = 0$, while $\{W_i^* - Q_i^*c\}$ is, by equation (7), increasing in i . Hence we must have some state j such that

$$\begin{cases} W_i^* - Q_i^*c \leq 0 & \text{for } i \leq j \\ W_i^* - Q_i^*c \geq 0 & \text{for } i > j \end{cases}.$$

Since U optimally shares in some of the risk, $W_1^* - Q_1^*c < 0$ and $W_n^* - Q_n^*c > 0$. ■

In the absence of either credit constraints or market risk, or both, the optimal supply contract would stipulate quantity $z_i q(c)$ in state z_i , resulting in the retail price $p(q(c))$ that

⁷For D , explicitly, in Hart's notation, we have the revenue function

$$f(z, Q) = Qp\left(\frac{Q}{z}\right),$$

which satisfies Hart's Assumptions 2 (as marginal revenue is positive and declining) and 6 (as profit grows in high demand states). As to his Assumption 5, we require the marginal revenue to grow in high demand states. This is true as

$$\frac{\partial^2 f}{\partial Q \partial z} = \left\{ \frac{\partial \left(\frac{\partial f}{\partial Q} \right)}{\partial \left(\frac{Q}{z} \right)} \right\} \frac{\partial \left(\frac{Q}{z} \right)}{\partial z} =_{\text{sign}} - \left[-\frac{Q}{z^2} \right] > 0,$$

where we have used the fact that the term in curly brackets is negative (as marginal revenue is declining). The other assumptions follow as U is assumed risk neutral and $I(\cdot)$ has been shown to be concave.

maximizes joint period-0 profit. Proposition 1 and Corollary 1 show that the interaction of credit constraints and market risk imply that this (joint period-0 profit maximizing) contract is not an optimal one for the endogenously risk-averse firm D to demand of her supplier. It can be improved by requiring U to share in the risk faced by the downstream firm D . Intuitively, for U to provide such risk sharing, he must earn more in good states than in bad states (Property 4). Since U earns zero profit on average, he must make a loss in the worst state(s). Hence, we can think of U as providing a fixed payment to D , with D then making demand-dependent repayments (Part 2 of Corollary 1).

In essence, D is using U to lower the variance of her end-of-period pledgable income by increasing the proportion which is fixed in advance. However, for U to be able to make back this ex ante committed payment the variable payments made to U must increase in volumes by more than the marginal cost of supply. Hence, double marginalization is created. This double marginalization is optimally spread across (almost) all demand states to reduce the temptation D has to misreport the state of demand. As a result, the optimal risk-sharing contract induces retail prices that are (in almost all demand states) strictly higher than $p(q(c))$. Hence, some of the burden of credit constraints and market risk is borne by consumers.

As discussed above, the optimal contract can be thought of as involving a fixed payment from U to D (and demand-dependent repayments). In the marketing literature, this fixed payment is known as a ‘slotting fee.’ Slotting fees are commonly used in the grocery industry as well as in software and publishing industries. While it has been noted that slotting fees can be rationalized by suppliers signalling the quality of their products to retailers (Klein and Wright, 2007), recent survey evidence suggests that risk sharing is a part of the rationale for slotting fees (Sudhir and Rao, 2006; Bloom et al., 2000).⁸

Our model suggests an alternative interpretation for the empirical results of Chevalier and Scharfstein (1996). These authors study retail prices in the U.S. supermarket industry. In each city studied they compare the prices charged by local supermarkets against those of national chains over a period when some cities faced a bad demand shock (recession) while others did not. Arguing that local stores are more likely to be credit constrained, they offer the empirical finding that firms facing credit constraints raise prices higher than non-constrained firms during bad demand outcomes. The model they offer is one of constant and fixed input prices with the presence of switching costs giving supermarkets a rationale for altering the price level. However, the mechanism we offer also fits the empirical findings without recourse to switching costs. Chevalier and Scharfstein rule out

⁸If U were himself risk averse (either due to credit constraints or otherwise), then the optimal contract would still involve D uploading some risk onto U . Thus, the slotting fee rationale would remain.

an explanation involving credit constrained supermarkets altering their marginal input cost by arguing that “it is hard to think of any reason why [such an] interpretation ... would be true.” We offer such a reason by demonstrating that the credit constraints themselves create the desire to alter input costs and share risks.

Remark 1 *In our analysis, we have allowed for general contracts between the upstream supplier U and the downstream buyer B . Suppose instead that firms were restricted to two-part tariff contracts of the form*

$$W(Q) = f + wQ,$$

where f is a fixed fee and w the per-unit input price. It can be shown that, in this case, the equilibrium contract in period 0, (f^*, w^*) , involves double marginalization (in all demand states), $w^* > c$, and payment of a slotting fee from the upstream firm to the downstream firm,

$$f^* = -q(w^*)(w^* - c) < 0.$$

4 Complementarities between Supplier Insurance and Banking

In the model as presented so far, the supplier U offers her downstream buyer D some pledgable income insurance. The downstream firm D then goes to the banking sector to borrow to fund the investment. If U could borrow and lend at the same (zero) interest rate as banks can, then U could take the place of the bank, providing the loan for investment as well as any pledgable income insurance. In fact, this section shows that borrowing from U and committing not to use a separate banking sector strictly dominates using a banking sector. The reason is that, by having to return to U for a loan, D can commit to charge a lower price and therefore one which is less double marginalized. This is because if she under-reports the state in period 0 and so makes extra profits, U can commit not to allow them to be leveraged. This permits D to credibly discipline herself. As a result, this section will offer a novel explanation for the existence of supplier finance arms.

To derive this result, suppose that D committed not to use a banking sector and only deal with U . D would now be proposing the contract $\{Q_i, T_0^i, T_1^i\}$, where Q_i is quantity of input delivered in period 0 if the state is z_i , T_0^i is a payment from D to U at the end of period 0 (so that D 's investment in period 1 is her revenue minus T_0^i), and T_1^i is a payment from D to U at the end of period 1, after the investment returns are realized.

The program to solve with no bank is as follows.

Program No Bank The optimal program when U provides the loan is given by:

$$\max_{\{Q_i, T_0^i, T_1^i\}} \sum_{i=1}^n g_i \left\{ \pi \left(Q_i p \left(\frac{Q_i}{z_i} \right) - T_0^i \right) - T_1^i \right\},$$

subject to

$$\sum_{i=1}^n g_i \{T_0^i + T_1^i - Q_i c\} \geq 0, \quad (8)$$

$$\left[Q_i p \left(\frac{Q_i}{z_i} \right) - T_0^i \right] \cdot B \leq \pi \left(Q_i p \left(\frac{Q_i}{z_i} \right) - T_0^i \right) - T_1^i, \quad (9)$$

$$\pi \left(Q_i p \left(\frac{Q_i}{z_i} \right) - T_0^i \right) - T_1^i \geq \pi \left(Q_j p \left(\frac{Q_j}{z_j} \right) - T_0^j \right) - T_1^j \text{ for all } j < i. \quad (10)$$

Here, (8) is the individual rationality constraint for U , (9) is D 's no-shirking constraint at the investment stage in period 1, and (10) is D 's incentive constraint when reporting the state of demand in period 0. Note that the optimal contract can involve a large penalty if D were unable to show all of the assets that she should have earned in period 0 according to her demand report. This implies that D can only under-report but not over-report the demand state in period 0. (This parallels the baseline model with a separate banking sector where D cannot claim pledgeable assets that she does not have.)

Note that if D should lie about the state and claim it is j when in fact it is $i > j$, then her assets will in truth be higher than she would have had under state j . However, the size of her loan (T_0^j) is not altered. These extra assets cannot, therefore, be leveraged.⁹

Proposition 2 *Using U as a bank strictly dominates using a separate banking sector.*

Proof. Consider the optimal tariff solving Program Bank: $\{Q_i^*, W_i^*\}$. This is the program when an independent banking sector is used. In state z_i , under this program, D has pledgeable income of $Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - W_i^*$ and invests an amount $I \left(Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - W_i^* \right)$, borrowing the difference between these two.

We first show that U can replicate the optimal contract D would set if using a banking sector. Suppose

$$\begin{aligned} T_1^i &= I \left(Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - W_i^* \right) - \left[Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - W_i^* \right], \\ T_0^i &= W_i^* - T_1^i, \end{aligned}$$

⁹We assume here that any such extra assets could still be invested, although not leveraged. Assuming otherwise would only strengthen our result.

where volumes $\{Q_i^*\}$ are as in the contract with the separate banks, and T_1^i is the size of the loan provided. Then, equation (8), the individual rationality constraint of U , is satisfied with equality by (4). By construction of T_1^i , the credit constraint is binding in every state so that the no-shirking constraint (9) always holds with equality. Finally, from the definition of the loan,

$$\begin{aligned} \pi \left(Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - T_0^i \right) - T_1^i &= B \cdot I \left(Q_i^* p \left(\frac{Q_i^*}{z_i} \right) - W_i^* \right) \\ &\geq B \cdot I \left(Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - W_j^* \right) \text{ for all } j \neq i, \end{aligned}$$

where the inequality is by the incentive constraint (5). The final term is the return available to D if her pledgable assets are $Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - W_j^*$ and she borrows to the point at which the credit constraint binds. We wish to show that this level of borrowing is greater than T_1^j for $j < i$. This is true if and only if having assets of $Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - W_j^*$ and borrowing T_1^j (resulting in investment equal to the level in the right-hand side of (10)) leaves the no-shirking constraint at the investment stage slack. This is shown by noting that, by definition,

$$\pi \left(Q_j^* p \left(\frac{Q_j^*}{z_j} \right) - T_0^j \right) - T_1^j = B \cdot \left[Q_j^* p \left(\frac{Q_j^*}{z_j} \right) - T_0^j \right].$$

Now consider increasing z_j to z_i . As $\pi' > 1 \geq B$ (see equation (3)), we must have

$$\pi \left(Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - T_0^j \right) - T_1^j > B \cdot \left[Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - T_0^j \right].$$

The left-hand side is the profit available if D borrows T_1^j to invest a total of $Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - T_0^j$. Hence, borrowing T_1^j with pledgable assets of $Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - T_0^j - T_1^j = Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - W_j^*$ leaves the credit constraint slack. We thus obtain

$$B \cdot I \left(Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - W_j^* \right) > \pi \left(Q_j^* p \left(\frac{Q_j^*}{z_i} \right) - T_0^j \right) - T_1^j \text{ for } j < i,$$

as required. Hence, the period-0 incentive constraint (10) is actually slack.

But as the incentive constraint on the report of the demand state in period 0 is slack, there is room for the transfer of some more risk upstream. Suppose that the quantities are altered to $Q_i^* + \varepsilon$ for all $i < n$ and the tariff W_i^* is increased by εc . The payments T_1^i and T_0^i retain the form given above. This new tariff satisfies (10) for small $\varepsilon > 0$. U remains indifferent, thus continuing to satisfy (8) with equality. By definition of T_1 , (9) is

satisfied with equality. It therefore remains to note that the objective function has grown. This follows as, by Property 2 of Proposition 1, the marginal revenue at states below n exceeds c . ■

The complementarity between supplier insurance and lending results from countervailing incentives being pooled. When applying for a loan, D would like to over-report the size of her assets so as to secure a larger loan. In contrast, in her supply insurance relationship, D would like to under-report the demand state so as to secure a larger insurance payout. By committing to leverage only those assets that are consistent with D 's demand report, U can effectively reduce D 's temptation to under-report the demand state and thus remove some double marginalization from the supply contract.¹⁰

Proposition 2 provides a rationale for suppliers maintaining finance arms, as indeed many major firms do (e.g., GE, Cisco). The finance arm will be able to offer terms which improve on those from a bank by linking the size of the loan to the quantity of input supplied. That a supplier with the same access to capital markets as an external bank can lend on rates that the independent banking sector would find unprofitable, is a new result. Understanding when such non-bank lenders have a comparative advantage over banks is important. In 2008, U.S. financial companies lent just over 608 billion dollars to business borrowers. This figure does not include financial companies lending to private consumers or for real-estate assets.¹¹ This compares with bank lending to businesses of 1.5 trillion dollars (commercial and industrial assets on U.S. bank balance sheets at end 2008). Thus financial companies lend almost \$1 for every \$2 lent by a mainstream bank. Therefore gaining an insight into what makes financial companies effective competitors for banks is arguably a first-order issue.

Carey et al. (1998) note that finance companies are over-represented in loans to higher risk firms. Such a distribution of loans could be explained if there is a complementarity between supplying input and lending, implying that such lenders can make a profit even with risky borrowers, whereas banks cannot. Note that our mechanism does not require that the upstream firm U provides all of the lending to D . Instead, U may cooperate with banks in a consortium of lenders – with the banks providing “inframarginal” lending (the part of the loan that would be provided even in the worst demand state) and U only providing the “marginal” lending that is sensitive to the reported demand state. To

¹⁰The mechanism offered here is related to the literature on countervailing incentives; see, e.g., Lewis and Sappington (1989). These authors show that, with countervailing incentives, the optimal contract may involve pooling in some states. Instead, our focus is to show that by pooling principals (supplier, bank), the agent (buyer) derives a benefit.

¹¹This is drawn from the Federal Reserve G20 statistical release. Available at <http://www.federalreserve.gov/econresdata/releases/statisticsdata.htm>.

enforce this, the borrowing firm must be limited in its access to further lenders for top-up loans. Covenants could be written to this effect.¹²

5 The Cost of Borrowing and Retail Prices

We have shown that the interaction between credit constraints and market risk causes a risk-neutral firm to become endogenously risk averse with respect to its pledgable income. The endogenous risk aversion causes the firm to seek to push risk on to its vertical partners. How risk averse the firm is will depend upon market-level and firm-level parameters. For example, the anticipated interest rate payable on future investment will alter the relationship between pledgable assets and investment levels and so impact on the extent of endogenous risk aversion. Similarly, changes in the quality of monitoring or of corporate governance (the ability to shirk) will alter the level of the credit constraint and so impact endogenous risk aversion and, hence, period-0 contracts. In this section we study the impact of changes in the interest rate payable by the firm on the real economy through short-run pricing and long-run investment decisions. We show that an increase in the interest rate will, under some conditions, increase D 's endogenous risk aversion and thus lead to higher retail prices in the short run and a lower investment level in the long run. The model can thus help explain *price puzzle* effects in macroeconomics.

Before studying the impact of changes in the interest rate (and, in Section 6, in D 's bargaining power), we pose the more general question of how a change in some parameter θ will alter the optimal contract between the credit-constrained D and U in period 0. Let $I(a; \theta)$ denote the (endogenous) investment level as a function of the realized asset level a and some parameter θ . The following lemma demonstrates that if a change in θ increases (decreases) the Arrow-Pratt measure of absolute risk aversion of the investment function, then the period-0 retail price will rise (fall) in all demand states except the highest and those at which the optimal contract involves pooling. As we are working with the optimal period-0 contract, this effect is not an artefact of a restricted contract class (such as linear or two-part tariff contracts).

Lemma 2 *Suppose a change in model parameter θ causes the coefficient of absolute risk aversion, $-\frac{\partial^2 I}{\partial a^2} / \frac{\partial I}{\partial a}$, to increase (decrease) at all pledgable asset levels. Then, at all states $i < n$ at which the optimal contract does not involve pooling, $Q_{i-1}^* < Q_i^* < Q_{i+1}^*$, the optimal quantity sold in period 0 decreases (increases). Hence, the short-run retail price*

¹²Indeed, there is evidence that, if lending is undertaken by a consortium, then covenants are more likely to be required (Bradley and Roberts, 2004).

in such states increases (decreases). The result holds weakly at state $i < n$ if the optimal contract in that state involves pooling, $Q_i^* \in \{Q_{i-1}^*, Q_{i+1}^*\}$.

We now turn to the analysis of the impact of changes in the interest rate payable on borrowed sums. In the following, we provide sufficient conditions under which an increase in the interest rate faced by D increases the relevant measure of D 's risk aversion during the pledgable income accumulation phase. By Lemma 2, this causes the short-run retail price to rise and the expected long-run investment level to decline.¹³ The model thus gives rise to a new monetary transmission mechanism between interest rates and short-run retail pricing, which is distinct from the seminal balance sheet channel of Bernanke and Gertler (1995).¹⁴

Explicitly, we suppose that money borrowed from the external banking sector between periods 0 and 1 needs to be repaid at an interest rate of r . As D is credit constrained, she will borrow as much as her end-of-period-0 assets allow. As D has to pay back $(I-a)(1+r)$ to the bank, the no-shirking constraint in period 1 is now given by

$$h(I, a, r) \equiv BI + (I - a)(1 + r) - \pi(I) \leq 0.$$

Thus, $h(I, a, r)$ measures the ‘‘incentive to shirk’’ at the investment stage. The maximal investment level $I(a, r)$ is implicitly defined by

$$h(I(a, r), a, r) = 0. \tag{11}$$

Consider the partial derivative of $h(I, a, r)$ with respect to the investment level:

$$\frac{\partial h(I, a, r)}{\partial I} = B + 1 + r - \pi'(I) \equiv \gamma(I, r).$$

We may call $\gamma(I, r)$, which measures how the incentive to shirk changes with the investment level, the ‘‘marginal incentive to shirk’’.

By Lemma 2, an increase in the interest rate r will lead to higher retail prices if it increases the coefficient of (absolute) risk aversion of investment returns with respect to pledgable assets. This happens if and only if the induced change in the curvature of I

¹³Over the course of the current Financial Crisis firms have faced historically higher borrowing rates. For two years from July 2007, the spread of corporate debt as compared to U.S. treasuries climbed to levels far in excess of anything experienced over the previous 3 years. Our model predicts a link between the increased cost of borrowing and higher (than myopically optimal) retail prices.

¹⁴In the balance sheet channel, the existing debt position of firms is worsened as their repayments rise and their net worth falls. This leads to the running down of inventories and a reduction in investment in the medium term. See the references in the Introduction and in Tirole (2006) for a model.

w.r.t. a is larger than that in the slope of I w.r.t. a , i.e., if and only if

$$\frac{d}{dr} \ln \left(-\frac{\partial^2 I(a, r)}{\partial a^2} \right) > \frac{d}{dr} \ln \left(\frac{\partial I(a, r)}{\partial a} \right), \quad (12)$$

where (using (11))

$$\frac{\partial I(a, r)}{\partial a} = \frac{1+r}{\gamma} \quad \text{and} \quad -\frac{\partial^2 I(a, r)}{\partial a^2} = \frac{(1+r)(d\gamma/da)}{\gamma^2} = \frac{\partial I(a, r)}{\partial a} \frac{d\gamma/da}{\gamma}. \quad (13)$$

We thus have

$$\frac{d}{dr} \ln \left(-\frac{\partial^2 I(a, r)}{\partial a^2} \right) = \frac{d}{dr} \ln \left(\frac{\partial I(a, r)}{\partial a} \right) + \frac{d}{dr} \ln \left(\frac{d\gamma/da}{\gamma} \right).$$

The inequality (12) therefore holds if and only if $\frac{d}{dr} \left(\frac{d\gamma}{da} \frac{a}{\gamma} \right) > 0$. The term $(d\gamma/da)(a/\gamma)$ is the elasticity of the marginal incentive to shirk with respect to the pledgable asset level. An increase in the interest rate r will thus raise the coefficient of risk aversion (and, by Lemma 2, retail prices) if it increases the elasticity of the marginal incentive to shirk with respect to the pledgable asset level. The following proposition shows that this is indeed the case, provided the investment return function $\pi(\cdot)$ is sufficiently curved (and the curvature does not increase with I).

Proposition 3 *Suppose the curvature of the technology function, $-\pi''(I)$, is sufficiently large in magnitude and declining at higher investment levels (i.e., $\pi'''(I) \geq 0$). Then, an increase in the interest rate causes:*

1. [cf. the price puzzle] *retail prices to rise in the short run (period 0);*
2. *the expected level of investment to decline in the long run.*

The discussion above confirms that the firm's endogenous risk aversion during the pledgable income accumulation phase increases as interest rates rise if the elasticity of the marginal incentive to shirk with respect to assets itself increases as interest rates rise.

To explain this we first demonstrate that this elasticity is positive. Suppose pledgable assets increase by a small percentage. Then the amount that can be invested rises. At the investment margin the last dollar invested is less productive and so the marginal incentive to shirk rises. Hence, the elasticity $\frac{d\gamma}{da} \frac{a}{\gamma}$ is indeed positive.

Now suppose we move to higher interest rates and again consider a small percentage increase in pledgable assets. At high interest rates less can be invested for given pledgable assets ($\partial I/\partial r < 0$), and the returns to investment at the margin are greater at lower

investment levels due to the curvature of $\pi(\cdot)$. The incentive to shirk at the margin balances this increasing marginal return against the greater interest payment and so changes only modestly. If pledgable income rises by a small percentage then there is now a substantial increase in investment levels due to the greater possible investment returns (due to the curvature of $\pi(\cdot)$).¹⁵ This substantial investment increases the incentive to shirk at the margin by a large amount as the extra investment moves the firm back up $\pi(\cdot)$.¹⁶ Hence, the percentage increase in the marginal incentive to shirk is large. That is, the elasticity rises, $\frac{d}{dr} \left[\frac{d\gamma}{da} \frac{a}{\gamma} \right] > 0$. This delivers Proposition 3. An increase in the firm's cost of borrowing for investment results in a greater coefficient of risk aversion and thus in higher retail prices in the pledgable income accumulation phase.

The fact that investment is lower in the long run now follows as a corollary. For any given realization of assets, we have $\partial I / \partial r < 0$ as the increased payback required lowers the level of borrowing that can be sustained. Further, realized assets are lower for any realization of period-0 market demand due to the increased double marginalization established in the first part of the proposition. Hence, the expected investment level must decline.

5.1 A Numerical Example

One might be concerned as to how large the curvature of π'' must be for our result to hold. To exhibit that infeasibly large curvatures are not required we here consider a very simple family of diminishing marginal returns technology functions: $\pi(I) = I^\lambda / \lambda$ with $\lambda < 1$. This family is chosen as it allows curvature to be arbitrarily small as λ approaches 1. Note that the first best investment level remains at $\hat{I} = 1$ for all λ .

Claim 1 *Consider investment technology $\pi(I) = I^\lambda / \lambda$. For tractability, let us rescale so that assets are $a = 0$ at end of period 0. Restrict attention to $\lambda \in (1/(1+B), 1)$ so as to ensure that credit constraints are binding. Then, the coefficient of absolute risk aversion in period 0 increases with the interest rate for all (allowable) parameter values of λ .*

¹⁵Differentiating (13) yields

$$\frac{\partial^2 I}{\partial a \partial r} = \frac{1}{\gamma} - \frac{1}{\gamma} \frac{\partial I}{\partial a} \frac{d\gamma}{dr} = \frac{1}{\gamma} \left[1 - \frac{\partial I}{\partial a} + \underbrace{\pi''(I) \frac{\partial I}{\partial r} \frac{\partial I}{\partial a}}_{>0} \right],$$

which is positive if $-\pi''$ is not too small.

¹⁶Algebraically, $\frac{d^2 \gamma}{da dr} = -\pi'''(I) \frac{\partial I}{\partial r} \frac{\partial I}{\partial a} - \pi''(I) \frac{\partial^2 I}{\partial a \partial r} > 0$. The inequality follows from the fact that the first term is positive as $\pi''' \geq 0$. And we have just argued that the second term is positive if π'' is large enough.

Proof. From Assumption 2(ii) the firm is credit constrained in period 0 if $a < (B + 1) - 1/\lambda$. As we consider $a = 0$ the restriction on λ indeed ensures that the firm is credit constrained at all considered λ .

From (11), given that $a = 0$ at end of period 0, we have

$$IB = \frac{I^\lambda}{\lambda} - I(1+r) \Rightarrow I(0, r) = [\lambda(B+1+r)]^{-1/(1-\lambda)}. \quad (14)$$

The marginal incentive to shirk is

$$\begin{aligned} \gamma|_{a=0} &= B + 1 + r - \pi'(I) \\ &= \text{from (14)} B + 1 + r - \lambda(B + 1 + r) = (1 - \lambda)(B + 1 + r). \end{aligned}$$

To determine the risk aversion we also require $\frac{\partial \gamma}{\partial a}$. Differentiating $\gamma \equiv B + 1 + r - \pi'(I)$, we obtain

$$\left[\frac{\partial \gamma}{\partial a} \right]_{a=0} = -\pi''(I) \frac{\partial I}{\partial a} = -\pi''(I) \frac{1+r}{\gamma} = (1-\lambda) I^{\lambda-2} \frac{1+r}{\gamma},$$

where we have used (13) giving $\frac{\partial I}{\partial a} = (1+r)/\gamma$.

Now observe from (13) that endogenous risk aversion for D is given by $\frac{\partial \gamma / \partial a}{\gamma}$. Combining the above yields

$$\begin{aligned} \left[\frac{\partial \gamma / \partial a}{\gamma} \right]_{a=0} &= [\lambda(B+1+r)]^{(2-\lambda)/(1-\lambda)} \frac{1+r}{(1-\lambda)(B+1+r)^2} \\ &= \frac{1+r}{1-\lambda} \lambda^2 [\lambda(B+1+r)]^{\lambda/(1-\lambda)}, \end{aligned}$$

which is clearly increasing in r . ■

Hence, risk aversion grows in r for all λ . That is, Proposition 3 applies to this family of investment return functions for all curvature levels.

5.2 Empirical Evidence: The ‘Price Puzzle’

The first part of Proposition 3 is closely related to the ‘price puzzle.’ The price puzzle refers to a long-standing observation in macroeconomics that retail prices appear to rise in the short run when interest rates are raised by the central bank. This is contrary to macroeconomic textbook discussions of the Phillips curve. Textbook macroeconomics would suggest that higher interest rates payable by business to fund investment should lead to a reduction in investment in the economy. This, it is argued, would shrink output

below the equilibrium level. In the simplest rendition of the theory this output gap puts downward pressure on wages and, hence, on prices. Thus higher interest rates would be expected to lead to price falls. However, before this macroeconomic effect occurs, prices (aggregated into an economy-wide price level) seem to first rise for a number of months to a year by a statistically significant amount (Christiano et al., 1999). The exact size of the price puzzle is in dispute as it varies depending on the extent to which the empirical estimation seeks to control for other variables such as inflation expectations (Balke and Emery, 1994; Bernanke et al., 2005). And stronger evidence exists that a price puzzle effect operates at more disaggregated sector levels (Gaiotti and Secchi, 2006). Proposition 3 provides a novel explanation for why retail prices might rise when the cost of borrowing a firm faces rises.

It might be argued that simpler mechanisms are likely to link increasing interest rates with higher retail prices. One such argument might be that if firms rent capital each period an increase in the interest rate payable directly raises firms' marginal cost of production, resulting in higher retail prices. This argument is sensitive to whether the capital stock is fixed or variable in the short run. If the capital stock is fixed in the short run, then interest rate changes would only affect the fixed costs of operation and not the retail prices. And further on its face the Phillips curve approach would seem to rely on some unresponsiveness of the capital stock in the short run. We remain agnostic on the degree of flexibility of capital. We merely note that the link between interest rates on borrowing and retail pricing we have determined in our model operates regardless of the flexibility of the capital stock.

6 Countervailing Power and Credit Constraints

In this section we relax our assumption that D has all of the bargaining power. This allows us to study how changes in D 's bargaining power as compared to her supplier alters the retail prices faced by consumers. We show that an increase in D 's bargaining power corresponds to an increase in the expected level of future investment and this lowers the degree of risk aversion D faces in the pledgable income accumulation phase. This therefore means that a more powerful buyer agrees a bargained supply contract which results in lower retail prices and higher investment levels. The model thus provides a novel *theory of countervailing power* based on credit constraints.

To model a more equal distribution of bargaining power between U and D , we let U receive a payoff from bargaining of β . This is a measure of U 's bargaining strength. We will first confirm that a bargaining power change does not alter the generic structure of

the optimal contract – it remains the case that double marginalization is unavoidable. However, we can demonstrate that as D 's bargaining power increases (U 's bargaining power falls) the extent of double marginalization falls. Thus downstream firms who have greater bargaining power *vis-à-vis* their suppliers sell at lower retail prices and invest more on average.

In this section, therefore, the bargaining analogue of Program Bank is modified by replacing the individual rationality constraint of U (equation (4)) by

$$\sum_{i=1}^n g_i \{W_i - Q_i c\} \geq \beta.$$

Proposition 4 [*Countervailing Power*] *Suppose U has bargaining power and requires an expected profit level of β from the relationship with D . Then:*

1. *The bargained contract is qualitatively identical to the benchmark model. Thus Propositions 1 (contract structure), 2 (finance arms), 3 (price puzzle) apply.¹⁷*
2. *If D 's bargaining power rises (i.e., β falls), then retail prices fall in the short run (period 0) and expected investment rises in the long run (period 1).*

Why do changes in bargaining power between D and U alter the endogenous degree of D 's risk aversion in the short run, thereby leading to retail price and investment level effects? If D 's bargaining power rises then U secures a lower return. This is equivalent to D gaining extra assets in addition to the income she makes through her normal business dealings in period 0. This increase in assets allows the amount borrowed, and thus the investment level, to grow. But at higher investment levels, the marginal return is lower, and thus the marginal incentive to shirk larger. If assets were to increase even further then the amount of extra borrowing would be modest. Hence, the responsiveness of the incentive to shirk with respect to assets is muted. This exactly says that the elasticity of the marginal incentive to shirk with respect to period-0 income declines as D 's bargaining power rises. Lemma 2 guarantees that the degree of risk aversion felt in period 0 declines also. Hence, less insurance is required, so less double marginalization is induced, so retail prices fall and expected investment levels rise.

Proposition 4 provides a novel theory of countervailing power based on credit constraints: consumer prices are lower the larger is the credit-constrained downstream firm's bargaining power *vis-à-vis* her upstream supplier. The term "countervailing power" was coined by Galbraith (1952) but Snyder (2008) notes that formalizing the concept has

¹⁷The same holds for Proposition 5 (outsourcing), which we will prove in Section 7.

proved difficult. Several theories of countervailing power (or buyer power) have recently been proposed in which upstream and downstream firms bargain. One strand of the literature builds on Katz (1987) and models bargaining as a supplier matching the price of some outside option. As such the question of bargaining power does not arise. A second influential strand has considered bilateral bargaining (Chipty and Snyder, 1999; Inderst and Wey, 2007). In this setting, the bargained transfer depends upon the expected incremental cost of supply – and this can differ between buyers of different size. However, without mandating inefficient bargaining, there is typically no retail effect from changes in bargaining power. Here, we are able to offer, to our knowledge, the first model of countervailing power based on credit constraints. Furthermore, in this case, a change in countervailing power has retail price effects.

7 Outsourcing

In our model, the upstream firm U provides (partial) insurance to his downstream buyer D . An obvious question is whether the insurance can instead be provided by a third party. The answer comes in two parts. First, if the third party can verifiably observe the input supply, then D may decide to source the input from U at marginal cost c and separately secure insurance from the third party. However, the retail price implications are unchanged as the insurance would induce double marginalization for the same reason as before.

Second, if the third party cannot verifiably observe the input supply, then U and D would have an incentive to collude and under-report the supply of input from U to D . (Of course, this is not possible when U provides insurance.) This would prevent a third party from providing insurance to D . In this case, we obtain the following result:

Proposition 5 *The credit-constrained downstream firm D strictly prefers to outsource input production to U rather than produce in-house at the same cost.*

Proof. Suppose D were to produce the input in-house at marginal cost c . In this case, in effect the supply contract would satisfy $W_i = cQ_i$ for all states i . Hence, for any demand state realization, the integrated firm would maximize its payoff by solving

$$\max_{Q_i} \sum_{i=1}^n g_i B \cdot I \left(Q_i p \left(\frac{Q_i}{z_i} \right) - cQ_i \right).$$

This is solved where $\partial [Q_i p(Q_i/z_i)] / \partial Q = c$ for all z_i . That is, the integrated firm would implement the non-double-marginalized retail price. However, by Proposition 1,

Property 2, though implementable, this is not the optimal tariff when D is outsourcing input production to U . Hence, D strictly prefers outsourcing to U . ■

Our model thus provides a new rationale for credit constrained firms exposed to market risk to outsource supply: the suppliers can provide revenue insurance that a third party cannot to the same extent.

There are many reasons why outsourcing might be a good idea. But the relationship between market risk and outsourcing is still a topic of debate. Empirically, there exists evidence supporting our theoretical results. For example, both Harrigan (1985) analyzing executive interviews and Sutcliffe and Zaheer (1988) experimentally find evidence that firms do move more production outside the firm when exposed to demand risk. However the dominant theoretical view is, arguably, that contractual incompleteness combined with demand risk would act to increase vertical integration (see Mahoney, 1992, for a survey and discussion).¹⁸ Our model suggests a force pushing against integration, which is responsive to market risk.

8 Conclusions

In this paper, we analyze a model of vertical relations between a downstream buyer and her upstream supplier. The downstream buyer is endogenously credit constrained which means that the scale of her investment is constrained by the level of her pledgable assets. Assuming that the downstream buyer's investment technology exhibits diminishing marginal returns, we find that the firm becomes endogenously risk averse when accumulating pledgable income.

As a result, the optimal contract between the (endogenously) risk-averse downstream firm and her upstream supplier involves risk sharing. This is true even if the supplier is himself risk averse or credit constrained. However, such a contract comes at a cost to consumers in the form of higher prices. Demand-dependent repayments to the supplier raise the downstream firm's effective marginal cost, inducing an increase in consumer prices. Thus double marginalization is a necessary feature of optimal supply contracts under credit constraints.

As supplier-insurance and lending for investment are subject to countervailing incentives, their pooling within one principal allows the downstream firm to reduce the double marginalization problem. So our model can explain why finance arms of major companies (such as GE) can lend profitably when banks cannot. Why such non-bank lending

¹⁸Carlton (1979) offers the same conclusion but in a model of unadjustable input volumes.

arrangements should exist and be thriving is currently not settled in the literature. Our model offers a contribution to this debate.

As the downstream firm's risk aversion is endogenous, it is affected by changes in market-level and firm-level parameters. This creates a new channel through which interest rates can affect short-run retail prices as well as long-run investment levels. As interest rates rise, the sensitivity of the firm's investment to pledgable income increases and so can be shown to make the firm more risk averse when accumulating pledgable assets. The increase in risk aversion results in greater insurance being demanded from the supplier which increases double marginalization in the supply contract and thus retail prices. This is potentially important in explaining empirically observed price dynamics such as the *price puzzle* in macroeconomics.

Relaxing our assumption that the downstream firm has all of the bargaining power, our model predicts that an increase in the credit-constrained firm's bargaining power with respect to her supplier will reduce the firm's endogenous risk aversion when accumulating pledgable assets. This generates a new theory of countervailing power as a more powerful downstream firm will set lower retail prices and invest more on average.

Finally, our model predicts that if input supply is not verifiable, then risk-averse firms exposed to market risk will gain by outsourcing supply (or sales). Once outsourced, the firm can enact a value-enhancing supply contract with insurance features. The same insurance cannot be provided by a third party if input supply is not verifiable by that third party.

These results have all been demonstrated in a model of downstream credit constraints and demand-side risk. However, the results are more general and would apply analogously to a model of upstream credit constraints and supply-side risk. There is, to our knowledge, little current empirical evidence which directly isolates the impact of credit constraints on pricing levels; though evidence of risk sharing and double marginalization is widespread. We have, in addition, reported much empirical evidence that appears to be in line with our predictions concerning finance arms, interest rates and outsourcing.

While the shape of the optimal contract between the credit-constrained (and thus risk-averse) downstream firm and her upstream supplier does not rely on the cause of the downstream firm's risk aversion, we would have been unable to obtain several of our results without explicitly modelling the interaction between the credit constraints and risk aversion. First, the complementarity between lending and insurance (giving rise to a theory of finance arms) obviously requires a role for credit constraints. Second, the relationship between interest rates and short-run retail prices and long-run investment levels relies on the interest rate altering the endogenous degree of risk aversion via the

credit constraints. Third, and similarly, the theory of countervailing power is based on bargaining power altering the endogenous degree of risk aversion via the credit constraints. In contrast, the results on double marginalization and slotting fees arise from the risk-sharing motive of the risk-averse downstream firm. The same is true for our results on outsourcing. Nevertheless, the IO literature has been skeptical about modelling firms as being risk averse. Our paper shows that risk-averse firms are to be expected as long as there is some chance of the firms being credit-constrained. The mechanism through which the endogenous risk aversion is generated does not rely on any separation of goals between owners and managers.

A Omitted Proofs

Proof of Lemma 2. We aim to show that:

if $Q_{i-1}^* < Q_i^* < Q_{i+1}^*$, then $\frac{\partial}{\partial \theta} \left[\frac{-\frac{\partial^2 I}{\partial a^2}}{\frac{\partial I}{\partial a}} \right] =_{\text{sign}} -\frac{\partial}{\partial \theta} Q_i^*(\theta)$ for all $i < n$;

if $Q_i^* \in \{Q_{i-1}^*, Q_{i+1}^*\}$, then $\frac{\partial}{\partial \theta} Q_i^*(\theta) = 0$ or $\frac{\partial}{\partial \theta} \left[\frac{-\frac{\partial^2 I}{\partial a^2}}{\frac{\partial I}{\partial a}} \right] =_{\text{sign}} -\frac{\partial}{\partial \theta} Q_i^*(\theta)$ for all $i < n$.

We first characterize the optimal period-0 contract in some more detail. Result 1 in Hart (1983) shows that the set of incentive constraints in Program Bank can be replaced with the following set of (local) constraints:

$$Q_i \geq Q_{i-1} \text{ for all } i \in \{2, \dots, n\}, \quad (15)$$

$$Q_i p \left(\frac{Q_i}{z_i} \right) - W_i \geq Q_{i-1} p \left(\frac{Q_{i-1}}{z_i} \right) - W_{i-1} \text{ for all } i \in \{2, \dots, n\}. \quad (16)$$

We now show that (16) must be satisfied with equality. Suppose not at some state i . Consider increasing W_i to $W_i + \varepsilon$ and lowering W_{i-1} to $W_{i-1} - \varepsilon(g_i/g_{i-1})$. (16) remains satisfied if $\varepsilon > 0$ is small. The individual rationality constraint of U , equation (4), is unaffected by construction. D 's objective function changes by

$$\begin{aligned} & B\varepsilon g_i \left\{ -I' \left(Q_i p \left(\frac{Q_i}{z_i} \right) - W_i \right) + I' \left(Q_{i-1} p \left(\frac{Q_{i-1}}{z_{i-1}} \right) - W_{i-1} \right) \right\} \\ & \geq B\varepsilon g_i \left\{ -I' \left(Q_i p \left(\frac{Q_i}{z_i} \right) - W_i \right) + I' \left(Q_{i-1} p \left(\frac{Q_{i-1}}{z_i} \right) - W_{i-1} \right) \right\} \\ & > 0, \end{aligned}$$

where the first inequality follows from $z_i > z_{i-1}$ and the concavity of the investment

function $I(\cdot)$, while the second equality follows by assumption on (16). But this is a contradiction to the optimality of the contract. Hence, constraint (16) must be satisfied with equality.

We next express the optimal period-0 contract purely in terms of quantities $\{Q_i\}$. From (16),

$$W_i - Q_i c = [W_{i-1} - Q_{i-1} c] + \Delta \Pi_i,$$

where

$$\Delta \Pi_i = Q_i \left[p \left(\frac{Q_i}{z_i} \right) - c \right] - Q_{i-1} \left[p \left(\frac{Q_{i-1}}{z_i} \right) - c \right].$$

The term $\Delta \Pi_i$ measures the industry profit gain if D does not lie and claim the state is marginally worse than it is (reporting $i-1$ instead of i). Iterating, we obtain

$$W_i - Q_i c = \sum_{j=2}^i \Delta \Pi_j + [W_1 - Q_1 c].$$

From the individual rationality constraint for U ,

$$\begin{aligned} 0 &= g_1 [W_1 - Q_1 c] + g_2 \{ \Delta \Pi_2 + [W_1 - Q_1 c] \} + g_3 \{ \Delta \Pi_3 + \Delta \Pi_2 + [W_1 - Q_1 c] \} \\ &\quad + \dots + g_n \left\{ \sum_{j=2}^n \Delta \Pi_j + [W_1 - Q_1 c] \right\} \\ \Rightarrow -[W_1 - Q_1 c] &= \sum_{k=2}^n \left(g_k \sum_{j=2}^k \Delta \Pi_j \right) = \sum_{k=2}^n \left(\Delta \Pi_k \sum_{j=k}^n g_j \right), \end{aligned} \quad (17)$$

where we have swapped the order of summation in the second expression. Thus, equation (17) gives W_1 . Furthermore, we have

$$W_i - Q_i c = \sum_{j=2}^i \Delta \Pi_j - \sum_{j=2}^n \left(\Delta \Pi_j \sum_{k=j}^n g_k \right) \text{ for } i \geq 2. \quad (18)$$

Note that the second term on the right-hand side of (18) is independent of i .

We now discuss the pledgable assets which will be available to D at the end of period 0, given any realization of the state. We have

$$\begin{aligned} a_i &= Q_i p \left(\frac{Q_i}{z_i} \right) - W_i \\ &= Q_i \left[p \left(\frac{Q_i}{z_i} \right) - c \right] - \sum_{j=2}^i \Delta \Pi_j + \sum_{j=2}^n \left(\Delta \Pi_j \sum_{k=j}^n g_k \right) \end{aligned}$$

First, we show that assets are increasing in the state:

$$\begin{aligned}
a_{i+1} - a_i &= Q_{i+1} \left[p \left(\frac{Q_{i+1}}{z_{i+1}} \right) - c \right] - Q_i \left[p \left(\frac{Q_i}{z_i} \right) - c \right] - \Delta \Pi_{i+1} \\
&= Q_i \left[p \left(\frac{Q_i}{z_{i+1}} \right) - p \left(\frac{Q_i}{z_i} \right) \right] \geq 0
\end{aligned} \tag{19}$$

The inequality follows as $z_{i+1} > z_i$. Next, we consider $\partial a_i / \partial Q_l$ for some realization of risk i and some contracted quantity at state l . From the last equation:

$$\begin{aligned}
\frac{\partial a_{i+1}}{\partial Q_l} - \frac{\partial a_i}{\partial Q_l} &= \frac{\partial}{\partial Q_l} \left\{ Q_i \left[p \left(\frac{Q_i}{z_{i+1}} \right) - p \left(\frac{Q_i}{z_i} \right) \right] \right\} \\
&= \begin{cases} 0 & \text{if } i \neq l, \\ \text{MR}_{i+1}(Q_i) - \text{MR}_i(Q_i) > 0 & \text{if } i = l, \end{cases}
\end{aligned} \tag{20}$$

where $\text{MR}_{i+1}(Q_i)$ is the marginal revenue in state z_{i+1} , evaluated at output Q_i . The final line follows as marginal revenue grows in higher demand states (see Footnote 7).

Hence, we have demonstrated that D 's problem can be rewritten as: maximize $E[I(a, \theta)]$ over $\{Q_i\}$, subject to (15) only, with the transfers being determined by (17) and (18).

Now, we turn to period 1. Suppose that the model parameter is at the level θ_1 and the optimal contract is $\{Q_i^*(\theta_1)\}$. The pledgable assets conditional on the state are given above. Recall that (15) must hold. Consider some state $l < n$ and suppose that $Q_{l-1}^* < Q_l^* < Q_{l+1}^*$. In this case, $E[I(a, \theta_1)]$ is maximized with respect to Q_l^* . That is $E \left[\frac{\partial I}{\partial a}(a, \theta_1) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} = 0$. Expanding, using (20), this states:

$$\left[\sum_{j=1}^l g_j \frac{\partial I}{\partial a}(a_j, \theta_1) \right] \frac{\partial a_l}{\partial Q_l} + \left[\sum_{j=l+1}^n g_j \frac{\partial I}{\partial a}(a_j, \theta_1) \right] \frac{\partial a_{l+1}}{\partial Q_l} = 0. \tag{21}$$

As the investment returns function $I(\cdot)$ is increasing, and using (20), we must have $\partial a_l / \partial Q_l < 0 < \partial a_{l+1} / \partial Q_l$.

Suppose that the model parameter rises slightly to $\theta_2 > \theta_1$. We now derive the optimal change in Q_l . To this end, we use the Taylor expansion identity that

$$\frac{\partial I}{\partial a}(a, \theta_2) = \frac{\partial I}{\partial a}(a, \theta_1) \left[1 + (\theta_2 - \theta_1) \frac{\frac{\partial^2 I}{\partial a \partial \theta}(a, \theta_1)}{\frac{\partial I}{\partial a}(a, \theta_1)} \right]$$

We aim to sign $E \left[\frac{\partial I}{\partial a} (a, \theta_2) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)}$, which can be rewritten as

$$\begin{aligned} E \left[\frac{\partial I}{\partial a} (a, \theta_2) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} &= E \left[\left\{ \frac{\partial I}{\partial a} (a, \theta_2) - \frac{\partial I}{\partial a} (a, \theta_1) \right\} \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} \\ &= (\theta_2 - \theta_1) E \left[\frac{\frac{\partial^2 I}{\partial a \partial \theta} (a, \theta_1)}{\frac{\partial I}{\partial a} (a, \theta_1)} \frac{\partial I}{\partial a} (a, \theta_1) \frac{\partial a}{\partial Q_l} \right]. \end{aligned}$$

Expanding out and using (20), yields

$$\begin{aligned} E \left[\frac{\partial I}{\partial a} (a, \theta_2) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} &= (\theta_2 - \theta_1) \left[\sum_{j=1}^l g_j \frac{\frac{\partial^2 I}{\partial a \partial \theta} (a_j, \theta_1)}{\frac{\partial I}{\partial a} (a_j, \theta_1)} \frac{\partial I}{\partial a} (a_j, \theta_1) \right] \underbrace{\frac{\partial a_l}{\partial Q_l}}_{<0} \\ &\quad + (\theta_2 - \theta_1) \left[\sum_{j=l+1}^n g_j \frac{\frac{\partial^2 I}{\partial a \partial \theta} (a_j, \theta_1)}{\frac{\partial I}{\partial a} (a_j, \theta_1)} \frac{\partial I}{\partial a} (a_j, \theta_1) \right] \underbrace{\frac{\partial a_{l+1}}{\partial Q_l}}_{>0}. \end{aligned}$$

Suppose that an increase in assets a reduces the Taylor quotient,

$$\frac{\partial}{\partial a} \left[\frac{\frac{\partial^2 I}{\partial a \partial \theta} (a, \theta_1)}{\frac{\partial I}{\partial a} (a, \theta_1)} \right] < 0. \quad (22)$$

As assets increase in the state (from (19)), we have

$$\begin{aligned} E \left[\frac{\partial I}{\partial a} (a, \theta_2) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} &< (\theta_2 - \theta_1) \frac{\frac{\partial^2 I}{\partial a \partial \theta} (a_l, \theta_1)}{\frac{\partial I}{\partial a} (a_l, \theta_1)} \left[\sum_{j=1}^l g_j \frac{\partial I}{\partial a} (a_j, \theta_1) \right] \frac{\partial a_l}{\partial Q_l} \\ &\quad + (\theta_2 - \theta_1) \frac{\frac{\partial^2 I}{\partial a \partial \theta} (a_l, \theta_1)}{\frac{\partial I}{\partial a} (a_l, \theta_1)} \left[\sum_{j=l+1}^n g_j \frac{\partial I}{\partial a} (a_j, \theta_1) \right] \frac{\partial a_{l+1}}{\partial Q_l} \\ &= 0, \end{aligned} \quad (23)$$

where the equality follows from (21). This therefore proves that $Q_l^*(\theta_2) < Q_l^*(\theta_1)$, and so retail prices would be higher under parameter θ_2 than θ_1 if (22) holds. The reverse result follows analogously by reversing the inequality if an increase in assets raises the Taylor quotient (i.e., if the inequality in (22) is reversed).

Hence, we have shown that if at state l with model parameter θ , $Q_{l-1}^* < Q_l^* < Q_{l+1}^*$, then

$$\frac{\partial}{\partial \theta} Q_l^* (\theta) =_{\text{sign}} \frac{\partial}{\partial a} \left[\frac{\frac{\partial^2 I}{\partial a \partial \theta} (a, \theta_1)}{\frac{\partial I}{\partial a} (a, \theta_1)} \right] =_{\text{sign}} - \frac{\partial}{\partial \theta} \left[\frac{-\frac{\partial^2 I}{\partial a^2}}{\frac{\partial I}{\partial a}} \right],$$

where the last equality follows algebraically. The last term is the rate of change of the

coefficient of risk aversion and so proves result 1 of the lemma.

Finally, we consider the case of pooling. We seek to modify the proof above to show that the pooled quantity falls weakly as we move to θ_2 if $\frac{\partial}{\partial a} \left[\frac{\frac{\partial^2 I}{\partial a \partial \theta}(a, \theta_1)}{\frac{\partial I}{\partial a}(a, \theta_1)} \right] < 0$. Consider the largest pooled state z_l , where $Q_{l-1}^* = Q_l^* < Q_{l+1}^*$. Note that $l < n$ as we know that at state n , the efficient quantity $z_n q(c)$ is delivered, while at state $n-1$, there is strictly too little quantity: $Q_{n-1}^* < z_{n-1} q(c) < z_n q(c) = Q_n^*$. As we have $Q_{l-1}^*(\theta_1) = Q_l^*(\theta_1)$, the optimization over state l is constrained, so that $E \left[\frac{\partial I}{\partial a}(a, \theta_1) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} \leq 0$. If $E \left[\frac{\partial I}{\partial a}(a, \theta_1) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} = 0$, then the identical proof to above applies showing that $Q_l^*(\theta_2) \leq Q_l^*(\theta_1)$. The inequality is weak as Q_l^* will only be able to fall if Q_{l-1}^* does. Suppose instead that $E \left[\frac{\partial I}{\partial a}(a, \theta_1) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} < 0$. The Taylor expansion around θ_2 is now given by

$$E \left[\frac{\partial I}{\partial a}(a, \theta_2) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} = E \left[\frac{\partial I}{\partial a}(a, \theta_1) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)} + (\theta_2 - \theta_1) E \left[\frac{\partial^2 I}{\partial a \partial \theta}(a, \theta_1) \frac{\partial a}{\partial Q_l} \right]_{Q_l^*(\theta_1)},$$

which is strictly negative for $\theta_2 - \theta_1$ small. Hence, again we have $Q_l^*(\theta_2) \leq Q_l^*(\theta_1)$. Finally, we obtain

$$Q_{l-1}^*(\theta_2) \leq Q_l^*(\theta_2) \leq Q_l^*(\theta_1) = Q_{l-1}^*(\theta_1),$$

where the first inequality follows by (15), the second inequality has just been shown, and the equality follows by assumption.

If, instead, $\frac{\partial}{\partial a} \left[\frac{\frac{\partial^2 I}{\partial a \partial \theta}(a, \theta_1)}{\frac{\partial I}{\partial a}(a, \theta_1)} \right] > 0$, then consider the smallest pooled state and repeat the argument of the paragraph above. ■

Proof of Propostion 3. From the discussion in the main text, an increase in the interest rate r will raise the coefficient of risk aversion (and, by Lemma 2, retail prices) if

$$\frac{d}{dr} \ln \left(\frac{d\gamma/da}{\gamma} \right) > 0. \quad (24)$$

Since $\partial I/\partial a > 0$, $\gamma > 0$ (Assumption 2) and so $d\gamma/da = -\pi''(I)\partial I/\partial a > 0$. Equation (24) holds if

$$\frac{d^2\gamma}{dadr}\gamma - \frac{d\gamma}{da}\frac{d\gamma}{dr} > 0,$$

where

$$\frac{d\gamma}{dr} = 1 - \pi''(I)\frac{\partial I}{\partial r} \text{ and } \frac{d^2\gamma}{dadr} = -\pi'''(I)\frac{\partial I}{\partial r}\frac{\partial I}{\partial a} - \pi''(I)\frac{\partial^2 I}{\partial a \partial r}$$

Differentiating (13) yields $\frac{\partial^2 I}{\partial a \partial r} = \frac{1}{\gamma} - \frac{1}{\gamma} \frac{\partial I}{\partial a} \frac{d\gamma}{dr}$. So equation (24) holds if

$$\left\{ -\pi'''(I) \frac{\partial I}{\partial r} \frac{\partial I}{\partial a} \gamma \right\} - \pi''(I) + \pi''(I) \frac{d\gamma}{dr} \frac{\partial I}{\partial a} - \frac{d\gamma}{da} \frac{d\gamma}{dr} > 0$$

Note that the brace is positive as $\partial I / \partial r < 0$ and $\pi'''(I) \geq 0$ is assumed in the statement of the proposition. Noting that $\frac{d\gamma}{da} = -\pi''(I) \frac{\partial I}{\partial a}$, a sufficient condition for equation (24) to hold is $\frac{d\gamma}{dr} \frac{\partial I}{\partial a} < \frac{1}{2}$. As $\frac{\partial I}{\partial a} > 0$, this is satisfied if $\frac{d\gamma}{dr}$ is negative, or at least not too large and positive. Given that $\frac{d\gamma}{dr} = 1 - \pi''(I) \frac{\partial I}{\partial r}$, the result follows if $\pi''(I)$ is sufficiently negative. Hence, an increase in the interest rate results in a larger coefficient of (absolute) risk aversion. Part 1 of the proposition follows from Lemma 2.

For part 2, note that the investment levels fall for any realization of assets as $\partial I / \partial r < 0$ and, further, realized assets are lower for any realization of period-0 market demand (except the largest), due to the lower equilibrium volumes. ■

Proof of Proposition 4. We first consider part 1. From Program Bank, with β replacing 0 on the RHS of equation (4), it is straightforward to see that an increase in β (the minimum expected profit level for U) is isomorphic to reducing D 's initial asset endowment by the same amount. Adding β to all the transfers converts one problem into the other, and so the optimal contracts take an identical shape and differ in a fixed payment of β . Next, a reduction in D 's initial asset endowment of β is equivalent to an alteration in the investment technology to $\pi(I) - \beta$. This follows from consideration of equation (2). Hence the bargained contract is of the form we have analyzed and so Propositions 1 (contract structure), 2 (finance arms), 3 (price puzzle) and 5 (outsourcing) to come, continue to apply.

We now turn to part 2. Given a required profit of β for U , part 1 has established that the optimal contract can be found by reducing D 's assets by β and proceeding as above. Hence the investment function the downstream faces $[I(a; \beta)]$ is defined implicitly by:

$$IB = \pi(I) - (I - [a - \beta]) \tag{25}$$

Part 2 then follows from Lemma 2 if $\frac{\partial}{\partial \beta} \left[-\frac{\partial^2 I}{\partial a^2} / \frac{\partial I}{\partial a} \right] > 0$. As before, let γ denote the marginal incentive to shirk, i.e., $\gamma = B + 1 - \pi'(I)$. Implicit differentiation of (25) yields

$$\frac{\partial I}{\partial a} = \frac{1}{\gamma} \text{ and } -\frac{\partial^2 I}{\partial a^2} = \frac{\partial \gamma / \partial a}{\gamma} \frac{\partial I}{\partial a}$$

Hence, following the same steps as in the proof of Proposition 3, raising D 's bargaining

power lowers endogenous risk aversion if and only if $\frac{\partial}{\partial \beta} \ln \left(\frac{\partial \gamma / \partial a}{\gamma} \right) > 0$. Now, note that

$$\frac{\partial}{\partial \beta} \left[\frac{\partial \gamma / \partial a}{\gamma} \right] =_{\text{sign}} \gamma \frac{\partial^2 \gamma}{\partial a \partial \beta} - \frac{\partial \gamma}{\partial \beta} \frac{\partial \gamma}{\partial a} \quad (26)$$

We have $\frac{\partial \gamma}{\partial \beta} = -\pi''(I) \frac{\partial I}{\partial \beta} < 0$ as increasing β increases investment ($\partial I / \partial \beta = -1/\gamma < 0$), and $\partial \gamma / \partial a > 0$ as $-\partial^2 I / \partial a^2$ is positive. Finally, we have $\frac{\partial^2 \gamma}{\partial a \partial \beta} = -\frac{\partial}{\partial a} \left[\frac{-\pi''(I)}{\gamma} \right]$ which is positive as the denominator is increasing with a while the numerator is decreasing in a as

$$\frac{\partial}{\partial a} [-\pi''(I)] = -\underbrace{\pi'''(I)}_{\geq 0} \underbrace{\frac{\partial I}{\partial a}}_{> 0} < 0.$$

Combining these inequalities confirms that (26) is positive, as required.

For the investment level result, note first that $\partial I / \partial \beta < 0$ for any realized period-0 income level a . Next, by the first part of the proof, an increase in β results in more double marginalization and therefore lower period-0 income a for any realization of market demand. Thus we have shown part 2. ■

References

- [1] Balke, N. and K. Emery (1994), “Understanding the Price Puzzle,” *Federal Reserve Bank of Dallas Economic and Financial Policy Review*, 4th Quarter, 15-26.
- [2] Bernanke, B., J. Boivin and P. Eliasch (2005), “Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach,” *Quarterly Journal of Economics*, 120(1), 387-422
- [3] Bernanke, B. and M. Gertler (1995), “Inside the Black Box: The Credit Channel of Monetary Policy Transmission,” *Journal of Economic Perspectives*, 9(4), 27-48.
- [4] Bloom P., G. Gundlach and J. Cannon (2000), “Slotting Allowances and Fees: Schools of Thought and the Views of Practicing Managers,” *Journal of Marketing*, 64(2), 92-108
- [5] Bradley, M. and M. Roberts (2004), “The Structure and Pricing of Corporate Debt Covenants,” SSRN working paper no. 466240.
- [6] Burkart, M. and T. Ellingsen (2004), “In-Kind Finance: A Theory Of Trade Credit,” *American Economic Review*, 94(3), 569-590.

- [7] Carey M., M. Post and S. Sharpe (1998), "Does Corporate Lending by Banks and Finance Companies Differ? Evidence on Specialization in Private Debt Contracting," *Journal of Finance*, 53(3), 845-878.
- [8] Carlton, D. (1979), "Vertical Integration in Competitive Markets Under Uncertainty," *Journal of Industrial Economics*, 27(3), 189-209.
- [9] Chevalier, J. and D. Scharfstein (1996), "Capital-Market Imperfections And Countercyclical Markups: Theory and Evidence," *American Economic Review*, 86(4), 703-725.
- [10] Chipty, T. and C. Snyder (1999), "The Role of Firm Size in Bilateral Bargaining: A Study of the Cable Television Industry," *Review of Economics and Statistics*, 81(2), 326-340.
- [11] Christiano, L., M. Eichenbaum and C. Evans (1999), "Monetary Policy Shocks: What Have We Learnt And To What End?", *Handbook of Macroeconomics*, Volume 1A. Eds. Taylor and Woodford eds. North Holland Press.
- [12] Cuñat, V. (2007), "Trade Credit: Suppliers as Debt Collectors and Insurance Providers," *Review of Financial Studies*, 20(2), 491-527.
- [13] Froot K., D. Scharfstein and J. Stein (1993), "Risk Management: Coordinating Corporate Investment and Financing Policies," *Journal of Finance*, 48(5), 1629-1658.
- [14] Gaiotti, E. and A. Secchi (2006), "Is there a cost channel of monetary policy transmission? An investigation into the pricing behavior of 2,000 firms," *Journal of Money, Credit and Banking*, 38(8), 2013-2037.
- [15] Galbraith, J. (1952), *American Capitalism: The Concept of Countervailing Power*, Boston: Houghton Mufflin.
- [16] Harrigan, K. (1985), "Vertical Integration and Corporate Strategy," *Academy of Management Journal*, 28(2), 397-425.
- [17] Hart, O. (1983), "Labour Contracts under Asymmetric Information: An Introduction," *Review of Economic Studies*, 50(1), 3-35.
- [18] Holmstrom, B. and J. Tirole (1997), "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112(3), 663-691.

- [19] Hubbard, G. (1998), "Capital-Market Imperfections and Investment," *Journal of Economic Literature*, 36, 193-225.
- [20] Inderst, R. and C. Wey (2007), "Buyer power and supplier incentives," *European Economic Review*, 51, 647-667.
- [21] Katz, M. (1987), "The Welfare Effects of Third-Degree Price Discrimination in Intermediate Good Markets," *American Economic Review*, 77(1), 154-167.
- [22] Klein, B. and J. Wright (2007), "The Economics of Slotting Contracts," *Journal of Law and Economics*, 50(3), 421-455.
- [23] Lewis, T. and D. Sappington (1989), "Inflexible Rules in Incentive Problems," *American Economic Review*, 79(1), 69-84.
- [24] Mahoney, J. (1992), "The Choice of Organizational Form: Vertical Financial Ownership Versus Other Methods of Vertical Integration" *Strategic Management Journal*, 13(8), 559-584.
- [25] Rey, P. and J. Tirole (1986), "The Logic Of Vertical restraints," *American Economic Review*, 76(5), 921-939.
- [26] Shin, H. and R. Stulz (1998), "Are Internal Capital Markets Efficient?" *Quarterly Journal of Economics*, 113(2), 531-552.
- [27] Snyder, C. (2008), "Countervailing Power," in: *The New Palgrave Dictionary in Economics*, Second Edition, Eds. Durlauf and Blume. North Holland Press.
- [28] Sudhir, K. and V. Rao (2006), "Do Slotting Allowances Enhance Efficiency or Hinder Competition?" *Journal of Marketing Research*, 43(2), 137-155.
- [29] Sutcliffe, K. and A. Zaheer (1988), "Uncertainty in the Transaction Environment: An Empirical Test," *Strategic Management Journal*, 19(1), 1-23.
- [30] Tirole, J. (2006), *The Theory of Corporate Finance*, Princeton University Press.
- [31] White, C., L. Troy and N. Gerlich (2000), "The Role of Slotting Fees and Introductory Allowances in Retail Buyers' New-Product Acceptance Decisions," *Journal of the Academy of Marketing Science*, 28(2), 291-298.