Fondazione Eni Enrico Mattei Working Papers

9-16-2009

Informal Finance: A Theory of Moneylenders

Andreas Madestam

IGIER, Bocconi University, jenny.madestam@statsvet.su.se

Follow this and additional works at: http://services.bepress.com/feem

Recommended Citation

http://services.bepress.com/feem/paper330

This working paper site is hosted by bepress. Copyright © 2009 by the author(s).
Informal Finance: A Theory of Moneylenders

Andreas Madestam

January 2009

Abstract

I study the coexistence of formal and informal finance in underdeveloped credit markets. While weak institutions constrain formal banks, shallow pockets hamper informal lenders. In such economies, informal finance has two effects. By increasing the investment return it decreases borrowers’ relative payoff following default, inducing banks to lend more liberally (disciplinary effect). By channeling bank capital it reduces banks’ agency costs from lending directly to borrowers, limiting banks’ extension of borrower credit (rent-extraction effect). Among other things, the model shows that informal interest rates are higher, borrower welfare lower, and informal finance more prevalent when the rent-extraction effect prevails, consistent with stylized facts in poor societies.

JEL classification: O12; O16; O17; D40.

Keywords: Credit markets; Financial development; Institutions; Market structure.
1 Introduction

Why do some poor borrowers in underdeveloped credit markets take informal loans despite the existence of formal banks, while others obtain funds from both financial sectors simultaneously? Also, why do poor borrowers with equal debt capacity pay very different informal rates of interest? For example, Banerjee and Duflo (2007) document that 95 percent of all borrowers living below $2 a day in Hyderabad, India get credit from informal sources even when banks are present (see Siamwalla et al., 1990 for similar findings from Thailand). Meanwhile, Das-Gupta et al. (1989) provide evidence from Delhi, India where 70 percent of all borrowers obtain capital from both sectors at the same time (see Conning, 2001 and Giné, 2007 for related support from Chile and Thailand). As regards interest rates, Aleem (1990), Banerjee (2003), and others have shown that borrowers with similar characteristics face informal interest rates ranging from 0 to 200 percent annually in India, Pakistan, and Thailand. Despite its empirical importance, the coexistence of formal and informal finance has not received as much attention as recent theoretical work on microfinance (see Ghatak, 1999; Ghatak and Guinnane, 1999; Ghatak, 2000).

In this paper, I provide a theory of informal finance that rationalizes credit market segmentation as well as multiple lending from banks and informal lenders. It suggests that market segmentation leads to higher informal interest rates with adverse welfare effects on borrowers, while multiple lending from both financial sectors induces lower informal interest and improves welfare. My theory is also consistent with the general observations that (i) moneylenders, traders, and landlords who offer informal credit frequently acquire bank funds to service borrowers’ financing needs (Hoff and Stiglitz, 1990, Ghate et al., 1992, and Irfan et al., 1999 remark that formal credit totals three quarters of the informal sector’s liabilities in many Asian countries); (ii) legal protection of creditors is essential to ensure availability of external capital (e.g., La Porta et al., 1997, 1998; Djankov et al., 2007); and (iii) banks in less developed credit markets often have extensive market power (Barth et al., 2004; Beck et al., 2004).

To address the above questions, I construct a simple model in which credit rationing is a result of creditor vulnerability in the bank sector. Specifically, moral hazard at the investment stage prevents banks from extending sufficient funds. By contrast, the informal sector is able to monitor borrowers and induce investment by offering credit to a group of known clients where social ties and social sanctions prevent borrowers from deliberately misusing their loan. The driving factor of the model is the interplay between the different constraints that formal and informal sector impose on credit.

---

1 See also De Soto (2000) for a complementary view stressing the role of property rights protection in promoting financial development.

2 For evidence of the highly personal character of informal lending see, for example, Udry (1990, 1994), Steel et al. (1997), and La Ferrara (2003) for the case of Africa and Aleem (1990), Bell (1990), and Ghate et al. (1992) for the case of Asia. See also Besley et al. (1993) and Banerjee et al. (1994) for theoretical work on rotating savings and credit associations stressing the importance of social sanctions. Anderson et al. (forthcoming) and Karlan (2005, 2007) provide related empirical evidence. Note that my aim is not to explain informal lenders ability to prevent opportunistic behavior, but to understand its implications as in Besley and Coate (1995).
informal lenders face. Banks have unlimited funds but are unable to prevent opportunistic behavior. Informal lenders can control the use of credit but may instead be capital constrained.

I find that informal finance raises investment, disciplines borrowers, and facilitates banks’ rent extraction. By ensuring prudent behavior, informal lenders are able to extend funds when banks cannot. This advantage cuts three ways. Additional informal credit increases the investment of bank-rationed borrowers. Access to (agency-free) informal capital also improves borrowers’ return to investment. Therefore, informal funds discipline borrowers by lowering the relative payoff following default, making it incentive compatible to increase bank lending (the disciplinary effect). Finally, by channeling bank funds informal lenders allow banks to reduce agency costs arising from lending directly to borrowers. Specifically, when extending money to poor borrowers, banks share rent to avoid credit misuse. Lending through informal lenders that are sufficiently rich not to be tempted by diversion means that banks need not share any rent (the rent-extraction effect). In contrast to the previous argument, informal finance thus limits borrowers’ access to bank capital.

The extent to which the disciplinary or the rent-extraction effect dominates depends on the allocation of bargaining power in the bank market. If banks are competitive, informal finance acts as a disciplinary device and expands overall credit provision. Borrowers obtain capital from both financial sectors, with poor informal lenders accessing banks for additional funds. Intuitively, when the surplus of the bank transaction accrues entirely to the banks’ clients, the residual return to an investment increases if banks extend credit to both the informal lender and the borrower. By contrast, informal finance serves as an instrument of rent extraction if the bank is a monopolist. Sufficiently wealthy informal lenders become borrowers’ only source of credit, credit that the informal sector acquires from the bank. As high interest rates increase the monopoly bank’s payoff and the borrowers’ incentive to default, poor bank customers earn a floor surplus above their outside option to limit diversion of bank funds. When the informal lender’s outside option exceeds the offered floor utility, the bank avoids sharing rent if it contracts exclusively with the informal lender. These findings may explain evidence from China indicating that informal finance is more important as the bank market becomes less competitive (Cull and Xu, 2005; Ayyagari et al., 2008; Cheng and Degryse, 2008). The theory’s predictions are also in accordance with Giné’s (2007) observation that poor borrowers in rural Thailand are more likely to access the informal sector alone when bank market power increases.

While financial sector coexistence increases efficiency, welfare is unequally distributed if informal lenders accumulate wealth. Informal finance lifts borrowers out of poverty if the disciplinary effect prevails, as richer informal lenders extend more funds and thus improve borrower incentives. Borrowers are worse off if the rent-extraction effect dominates, however, as they are left completely in the hands of informal lenders, rather than obtaining the bank’s contractual rent. Meanwhile, poor lenders are better off disciplining borrowers, as they receive more bank funds and higher incentive rent, whereas richer lenders prefer the segmented outcome as it preserves their market power. If wealthy informal lenders and bankers have more
say over bank market structure than poor borrowers, these results are consistent with Rajan and Ramcharan’s (2008) finding that banking in the early twentieth century United States was more concentrated in counties with rich landowners who often engaged in lending to farmers. These landlords frequently had ties with the local bank and were, as the model predicts, against bank deregulation.

Variation in informal finance can also be understood within the theory’s framework. As noted, informal finance should be more important in concentrated bank markets where the rent-extraction effect is in force. Likewise, a bad legal environment increases the presence of informal finance if the disciplinary effect dominates, as wealthy informal lenders act as substitutes for institutional quality. Meanwhile, informal finance is less prevalent following institutional decline if informal lenders alone provide all external capital. Worse creditor protection tightens bank credit to informal lenders with a low debt capacity; with no effect on lending if informal lenders are rich enough not to be tempted by diversion. Thus, while bank credit contracts as the legal environment deteriorates, the importance of informal finance can go either way. This helps explain Dabla-Norris and Koeda’s (2008) finding that the empirical relationship between institutions and informal credit is indeterminate, while bank lending narrows as legal protection worsens.

The results also offer insight to the evidence documented by Banerjee (2003) that informal interest rates can be usurious and highly variable. In the model, poor informal lenders charge positive rates of interest even if the adjacent bank market is perfectly competitive. This is because the price of credit reflects the incentive rent informal lenders receive to ensure prudent behavior when forwarding bank funds. I further find that informal interest rates increase if the rent-extraction effect dominates. The reason is that the segmented outcome preserves informal lenders’ market power, while an improved debt capacity forces informal lenders to offer loans at the opportunity cost of funds if they discipline borrowers under bank competition.

The paper relates to three strands of the literature. First, it connects to theoretical work stressing the notion that informal lenders hold a monitoring (or screening) advantage over formal banks. These papers either view informal lenders as bank competitors or as a channel of bank funds. In the former case, by rationalizing multiple lending from banks and informal lenders as an outcome either of exogenous bank credit limits set by the government (Bell et al., 1997) or because banks co-finance projects to draw on the informal sector’s edge in screening out bad loans (Jain, 1999) or in recovering repayments (Varghese, 2005). The joint theme of the channeling theories is their focus on analyzing the negative effects of a bank credit expansion on informal interest rates (Floro and Ray, 1997; Bose, 1998; Hoff and Stiglitz, 1998).\footnote{Bose and Hoff and Stiglitz show that subsidized bank credit induces informal lenders to enter the market leading to higher informal enforcement costs, while Floro and Ray consider how a formal credit expansion increases informal lenders’ ability to collude. The end effect in all three cases is that informal rates increase.}

Although this paper shares some of these ideas there are a number of important differences. First, in earlier work it is not clear whether informal lenders compete with banks or primarily engage in channeling funds. Second, multiple lending theories cannot account for bank lending
to the informal sector or for variation in informal interest rates. Third, channeling theories fail to address the potential agency problem between the formal and the informal lender and do not clarify why channeling of funds occurs in the first place. My model explains why informal lenders take bank credit in both these instances, making competition and channeling a choice variable in a framework where monitoring problems exist between banks and informal lenders, as well as between banks and borrowers. The theory thus extends and reconciles existing approaches by allowing for both competition and channeling of funds while deriving endogenous constraints on informal lending.

The second line of literature related to my model explores the interaction between modern and traditional sectors to rationalize persistence of traditional or personal exchange. This includes the studies of Kranton (1996) and Banerjee and Newman (1998). Whereas these papers focus on how market imperfections give rise to institutions that (possibly) impede the development of markets, the present theory focuses on how a given organizational form (informal finance) is affected by changes in wealth, institutional quality, and market power.

Finally, my approach also links to the research of Petersen and Rajan (1995) and Mookherjee and Ray (2002) by emphasizing contractual or market structure as an important determinant of credit availability in less developed financial markets. Similar to Mookherjee and Ray (and in contrast with Petersen and Rajan), I find that bank market power reduces efficiency. Like the present paper, Mookherjee and Ray stress the interplay between debt capacity and market concentration. However, while their analysis concerns the effect of contractual structure on borrowers’ intertemporal savings decisions, I investigate the consequences of market power on the pattern of intermediation between different types of lenders and borrowers.

The model builds on Burkart and Ellingsen’s (2004) analysis of trade credit in a competitive banking and input supplier market. The bank and the borrower in their model are analogous to the competitive formal lender and the borrower in my setting. However, their input supplier and my informal lender differ substantially. Moreover, I explore bank sector market power.

Also, Kochar (1997) empirically invalidates the existence of exogenous constraints as proposed by Bell et al. Another point of difference is that formal-informal coexistence arises as an equilibrium outcome in my setting, while Jain and Varghese derive it by allowing banks to contract on the informal lenders’ presence.

My findings also differ from other theories of intermediation, such as Holmström and Tirole (1997), who cannot explain why borrowers and informal lenders simultaneously take bank credit. Moreover, while financial arrangements in my model have distinct efficiency and distributional features, certification (investors and banks lend to borrowers) and intermediation (investors deposit their funds in banks) are outcome equivalent in Holmström and Tirole.

Varghese (2005) is an exception. However, he does not consider informal lenders’ intermediary function.

See also Arnot and Stiglitz (1991) and Besley and Ghatak (2008).


Burkart and Ellingsen assume that it is less profitable for the borrower to divert inputs than to divert cash. Thus, input suppliers may lend when banks are limited due to potential agency problems.

While the input supplier and the (competitive) bank offer a simple debt contract, the informal lender offers a
In the next section I introduce the model then in Section 3 present equilibrium outcomes. Section 4 deals with welfare. Section 5 examines determinants of informal finance. Section 6 studies informal interest rates. I conclude by discussing robustness issues, consider possible extensions, and point to some policy implications. Formal proofs are relegated to the Appendix.

2 Model

Consider a credit market consisting of risk-neutral entrepreneurs (the borrowers), banks (who provide formal finance), and moneylenders (who provide informal finance). The entrepreneur is endowed with observable wealth \( \omega_E \geq 0 \). She has access to a deterministic production function, \( Q(I) \), where \( I \) is the volume of investment. The production function is assumed to be concave and twice continuously differentiable. To ensure the existence of an interior solution, it is assumed that \( Q(0) = 0 \) and \( Q'(0) = \infty \). In a perfect credit market with interest rate \( r \), the entrepreneur would like to invest enough to attain the first-best level of investment given by \( Q'(I^*) = 1 + r \). However, she lacks sufficient capital to realize this level, \( \omega_E < I^*(r) \), and is thus forced to resort to the bank and/or the moneylender for the remaining funds.\(^{11}\)

Although banks have an excess supply of funds, credit will be limited as the entrepreneur is unable to commit to invest all available resources into her project. Specifically, I assume that the entrepreneur may use (part of) the assets to generate nonverifiable private benefits. Following standard practice (see Hart, 1995), opportunistic behavior resulting in diversion of funds denotes any activity that is less productive than investment, for example, using available resources for consumption or financial saving.\(^{12}\) The diversion activity yields benefit \( \phi < 1 \) for every unit diverted. While investment is unverifiable, the outcome of the entrepreneur’s project in terms of output and/or sales revenue may be verified. The entrepreneur thus faces the following trade-off: either she invests, in which case she realizes the net benefit of production after repaying the bank (and possibly the moneylender), or she profits directly from diverting the bank funds (the entrepreneur still pays the moneylender if she has taken an informal loan). In the case of partial diversion, any remaining returns must be repaid to the bank in full. The bank is assumed not to derive any benefit from resources that are diverted.

As noted in the introduction, moneylenders have a monitoring advantage over banks such that credit granted is fully invested. For simplicity, I assume monitoring cost to be zero.\(^{13}\) The moneylender’s superior knowledge of local borrowers grants him exclusivity and some

---

\(^{11}\) I assume that the entrepreneur accepts the first available contract if indifferent between the contracts offered.

\(^{12}\) Although diversion in its most severe form can be interpreted as “taking the money and run”, milder forms of opportunistic behavior are probably more important. Resources could be used to cover a range of expenditures, for example, school fees or health care costs.

\(^{13}\) It turns out that monitoring cost is irrelevant unless sufficiently prohibitive to prevent banks or entrepreneurs from dealing with the informal sector altogether.
market power, defined in more detail below. In the absence of any contracting problem between the moneylender and the entrepreneur, the moneylender maximizes the joint surplus derived from the investment project and divides the proceeds using Nash Bargaining. A contract is given by a pair \((B, R) \in \mathbb{R}_+^2\), where \(B\) is the amount borrowed by the entrepreneur and \(R\) the repayment obligation. Finally, if the moneylender requires additional funding he turns to the bank.

Following the same logic as above, I assume that the moneylender cannot commit to lend his bank loan and that diversion yields private benefits equivalent of \(\phi < 1\) for every unit diverted. While lending is unverifiable, the outcome of the moneylender’s operation may be verified. The moneylender thus faces the following trade-off: either he lends the bank credit to the entrepreneur, realizing the net-lending profit after compensating the bank, or he benefits directly from diverting the bank loan. In the case of partial diversion, the moneylender repays the remaining amount to the bank in full. Banks do not benefit from assets that are diverted.

Finally, banks have access to unlimited funds at a constant unit cost of zero and offer a contract \((L_i, D_i)\), where \(L_i\) is the loan and \(D_i\) the amount to be repaid, \(i \in \{E, M\}\). When \(\phi\) is equal to zero, legal protection of banks is perfect and there is no agency problem in the sense that a penniless entrepreneur and/or moneylender could raise the amount needed to attain first-best investment. To make the problem interesting, I assume that

\[
\phi > \phi \equiv \frac{Q(I^*(0)) - I^*(0)}{I^*(0)}.
\]

In words, the marginal benefit of diversion yields higher utility than the average rate of return to first-best investment at zero rate of interest [henceforth \(I^*(0) = I^*\)].

The timing is as follows:

1. Banks offer a contract, \((L_i, D_i)\), to the entrepreneur and the moneylender respectively.
2. The moneylender offers a contract, \((B, R)\), to the entrepreneur.
3. The moneylender makes his lending/diversion decision.
4. The entrepreneur makes her investment/diversion decision.
5. Repayments are made.

To distinguish formal from informal finance, I assume that banks are unable to condition their contracts on the moneylender’s contract offer, an assumption empirically supported by Giné

---

14 The assumption that borrowers obtain funds from at most one informal source has empirical support see, for example, Aleem (1990) and Sianwala et al. (1990). The evidence of the extent of informal lenders’ market power is less clear. Informal finance has been documented as competitive (Adams et al., 1984), monopolistically competitive (Aleem, 1990), and as a monopoly (Bhaduri, 1977).

15 Additional sources of income does not alter the main insights. See the final section for a discussion.
If not, the entrepreneur could obtain an informal loan and then approach the bank. Bank credit would then depend on the informal loan and the subsequent certain investment.

3 Equilibrium

I first establish some benchmark results by analyzing each financial sector in isolation. This helps understand how the trade-off between extraction of rents and provision of incentives differ depending on bank market structure and type of lender.

3.1 Benchmark Cases

I begin with the competitive bank market and solve for the subgame-perfect equilibrium outcome. Without loss of generality, I follow Burkart and Ellingsen (2004) and focus on contracts of the form \( \{(L_E, (1 + r) L_E)\}_{L_E \leq \bar{L}_E} \), where \( L_E \) is the loan, \( (1 + r) L_E \) the repayment, and \( \bar{L}_E \) the credit limit.\(^{17}\) The contract implies that a borrower may withdraw any amount of funds until the credit limit binds. For simplicity, entrepreneurs only borrow from one bank at a time. Following a Bertrand argument, competition drives equilibrium bank profit to zero. Nonetheless, credit is limited as investment of bank funds cannot be ensured. Specifically, the entrepreneur chooses the amount of bank funds to invest, \( I \), and the amount of credit, \( L_E \), by maximizing

\[
U_E = \max \{ 0, Q(I) - (1 + r) L_E \} + \phi(\omega_E + L_E - I)
\]

subject to

\[
\begin{align*}
\omega_E + L_E & \geq I, \\
\bar{L}_E & \geq L_E.
\end{align*}
\]

The first part of the expression is the profit from investing, taking the entrepreneur's limited liability into account. The second part denotes the gain from diversion. The full expression is maximized subject to available funds and the credit limit. It follows that neither partial investment nor diversion is optimal. Investing yields the entrepreneur at least \( 1 + r \) on every dollar invested, while diversion leaves her with only \( \phi \). If the entrepreneur plans to divert resources, there is no reason to invest either borrowed or internal funds as the bank would claim all of the returns. Hence, the choice is essentially binary; either the entrepreneur chooses to invest all the money or she diverts the maximum possible. The entrepreneur will not behave opportunistically if the contract satisfies the incentive constraint

\[
Q(\omega_E + \bar{L}_E^u) - (1 + r) L_E^u \geq \phi(\omega_E + \bar{L}_E),
\]

\(^{16}\) See also Bell et al. (1997) for evidence in support of the assumed sequence of events.

\(^{17}\) In a framework similar to the competitive bank market in the present paper, Burkart and Ellingsen (2002) show that \( \{(L_E, (1 + r) L_E)\}_{L_E \leq \bar{L}_E} \) constitutes an optimal contract.
where $L_E^u = \min \{ I^* (r) - \omega_E, \bar{L}_E \}$. Either the entrepreneur borrows and invests efficiently, or she exhausts the credit line extended by the bank. As there is no default in equilibrium, the only equilibrium interest rate consistent with zero profit is $r = 0$.

At low debt capacities, $\omega_E < \omega_E^c$, the temptation to divert resources becomes too large to allow a loan in support of the first-best outcome. In this case, the credit limit is given by the binding incentive constraint

$$Q(\omega_E + \bar{L}_E) - \bar{L}_E = \phi(\omega_E + \bar{L}_E). \quad (2)$$

As an increase in wealth improves the return to investment for a given loan size, the credit line and the investment rise with wealth, see Figure 1. When the entrepreneur is sufficiently wealthy the constraint no longer binds and first-best investment is obtained.

The other benchmark case is one with a monopoly bank in the market. The monopolist sets $L_E$ and $D_E$ by maximizing

$$U_B = D_E - L_E$$

subject to the participation constraint

$$Q(\omega_E + L_E) - D_E \geq Q(\omega_E)$$

and the incentive constraint above. The inequality ensures at least the utility associated with self financing the project. $D_E$ replaces $(1 + r) L_E$ with the borrower choosing whether or not to accept the bank’s take-it-or-leave-it offer and consequently the amount to invest. It follows that the relevant incentive and/or participation constraint must bind, otherwise the bank could increase $D_E$ and earn a strictly higher profit.

For low levels of wealth, $\omega_E < \omega_E^m$, the incentive constraint binds and the bank’s profit may be written as $Q(\omega_E + L_E) - \phi(\omega_E + L_E) - L_E$. The first-order condition of the bank’s profit expression determines the optimal loan size, whereas $D_E$ is defined as the solution to the incentive constraint. Hence, $L_E$ is the unique loan size that solves

$$Q'(\omega_E + L_E) - (1 + \phi) = 0, \quad (3)$$

while $D_E$ is determined by

$$Q(\omega_E + L_E) - D_E = \phi(\omega_E + L_E). \quad (4)$$

A salient feature of this outcome is that entrepreneurs are provided a constant floor rent above their outside option to satisfy the investment level, $I = \omega_E + L_E$, given by equation (3). Since higher wealth is met by a parallel decrease in bank credit to maintain the sub-optimal investment, any wealth improvement is pocketed by the bank. Poor entrepreneurs are thus prevented from accumulating assets. (See Figure 1.)

As wealth increases, $\omega_E \in [\omega_E^m, \bar{\omega}_E^m)$, the participation and the incentive constraint hold simultaneously. A higher debt capacity permits the bank to increase the repayment obligation
such that the entrepreneur is indifferent between taking credit and self financing the project. Since first best is unattainable, the loan size continues to satisfy the incentive constraint. Hence, the repayment obligation is determined by the binding participation constraint while the equilibrium loan size solves

\[ Q(\omega_E) = \phi (\omega_E + L_E). \]  

(5)

Rising wealth induces the bank to increase lending up to the point at which the incentive and the participation constraint hold with equality. Finally, for rich entrepreneurs only the participation constraint binds and the efficient outcome is obtained. Proposition 1 summarizes the findings.

**Proposition 1:** For all \( \phi > \phi^\ell \), there are thresholds \( \omega^c_E > 0 \) and \( \omega^m_E > \omega^m_E > 0 \) such that:

(i) Entrepreneurs with wealth below \( \omega^c_E \) invest \( I < I^\ast \) under competitive banking and bank credit \( (L_E) \) and investment \( (I) \) increase in \( \omega_E \); if \( \omega_E \geq \omega^c_E \) then \( I^\ast \) is invested.

(ii) Entrepreneurs with wealth below \( \omega^m_E \) invest \( I = I' \) as given by equation (3) with a monopoly bank, \( L_E \) decreases in \( \omega_E \), and \( I' \) is independent of \( \omega_E \); if \( \omega_E \in [\omega^m_E, \omega^m_E] \) then \( I \in [I', I^\ast) \) is invested and \( L_E \) and \( I \) increase in \( \omega_E \); if \( \omega_E \geq \omega^m_E \) then \( I^\ast \) is invested.

(iii) Competition increases efficiency, that is, \( \omega^m_E > \omega^c_E \).

My theory thus predicts that monopoly banking reduces lending and investment. Intuitively, whereas the competitive outcome minimizes banks’ aggregate payoff, the monopoly outcome maximizes this return by allowing a monopolist to charge the highest interest rate possible. When increasing the price, the bank lowers the borrower’s incentive to repay. Hence, high interest rates must be coupled with less lending and as a consequence lower investment.

Finally, absent any contracting frictions a moneylender maximizes the joint surplus of the investment project, \( Q(\omega_E + B) - B \). Let \( B^\ast \) denote the loan size that solves the first-order condition \( Q'(\omega_E + B) - 1 \geq 0 \). The efficient outcome, \( B^\ast = I^\ast - \omega_E \), is obtained if the moneylender is sufficiently wealthy, while the outcome is constrained efficient otherwise, with \( B^\ast = \omega_M < I^\ast - \omega_E \). Excess moneylender funds are deposited in the bank earning a zero
rate of interest. Given \( B^* \), the entrepreneur and the moneylender bargain over how to share the project gains using available resources \( \omega_E + B \), with \( \omega_M \geq B \). If they disagree, investment fails and each party is left with her/his wealth or potential loan. In case of agreement, the moneylender offers a contract where the equilibrium repayment, using the Nash Bargaining solution, is

\[
R(B)^* = \arg \max_t \{ Q(\omega_E + B) - t - \omega_E \}^\alpha \{ t - B \}^{1-\alpha} \\
= (1 - \alpha) [Q(\omega_E + B) - \omega_E] + \alpha B,
\]

where \( \alpha \in (0, 1) \) represents the entrepreneur’s bargaining power or equivalently the informal sector’s degree of competitiveness.

### 3.2 The Disciplinary Effect: A Moneylender and Competitive Banks

Moneylenders help capital-constrained entrepreneurs to increase investment by allowing more external funding to be raised. The coexistence of banks and moneylenders also permits the moneylender to access bank credit, introducing additional trade-offs. Although (agency-free) informal credit improves the incentives of the entrepreneur, banks now have to guard against the possibility of diversion on the part of the entrepreneur and the moneylender. Moreover, the informal credit channel allows banks to sidestep the entrepreneur altogether to avoid entrepreneurial agency costs, keeping in mind that the moneylender may still cheat the banks.

If banks compete, some of the credit is provided by the informal lender and some by the bank. As indicated in Proposition 1, all benefits of incremental wealth gains under competitive banking accrue to the borrower, enabling banks to lend more as a result of improved incentives. When banks and moneylenders both extend money, informal capital increases the residual return to the entrepreneur’s project (accounting for the additional repayment to the moneylender) with the end effect equivalent to an increase in internal funds. Informal finance thus disciplines entrepreneurs as it makes them less prone to divert bank credit. Before turning to the precise characterization, I make the additional assumption that

\[
\phi > \phi^*_i(\omega_i) \equiv \frac{Q(I^*) - (I^* - \omega_i)}{I^*}, \tag{6}
\]

where \( i \in \{E, M\} \). As the moneylender’s wealth facilitates the entrepreneur’s constraint (and vice versa), this needs to be incorporated. The condition ensures that diversion benefits exceed the average return to an investment \( I^* \), accounting for entrepreneurial or informal lender wealth. If not, a penniless entrepreneur and/or moneylender could support first best.

Solving backwards and starting with the entrepreneur’s incentive constraint yields

\[
Q(\omega_E + L_E^* + B) - L_E^* - R(B) \geq \phi(\omega_E + L_E), \tag{7}
\]

\[\text{If there is agreement and } Q(\omega_E + B) - R \leq Q(\omega_E), \text{ the entrepreneur receives } Q(\omega_E) \text{ leaving the residual } Q(\omega_E + B) - Q(\omega_E) \text{ to the moneylender.}\]

---

18 If there is agreement and \( Q(\omega_E + B) - R \leq Q(\omega_E) \), the entrepreneur receives \( Q(\omega_E) \) leaving the residual \( Q(\omega_E + B) - Q(\omega_E) \) to the moneylender.
where $L^u_E = \min \{ I^* - \omega_E - B, \bar{L}_E \}$. The only modification from above is that the amount borrowed from the moneylender, $B$, is prudently invested\textsuperscript{19}

If the moneylender requires additional funding, he also turns to a bank and chooses the amount to lend to the entrepreneur, $B$, and the amount of credit, $L_M$, such that the following incentive constraint is satisfied

$$R(\omega + L^u_M) - L^u_M \geq \phi (\omega + \bar{L}_M),$$

where $R(B)$ is a function of the amount lent to the entrepreneur for any pair $(L^u_M, \omega_M)$, with $L^u_M = \min \{ I^* - \omega_M - \omega_E - L^u_E, \bar{L}_M \}$. The left-hand side of the inequality is the moneylender’s net-lending profit, while the right-hand side is the return from borrowing a maximum amount and then diverting all available assets\textsuperscript{20}

It remains to determine the Nash Bargaining outcome. As before, I have

$$R(B)^* = (1 - \alpha) [Q (\omega_E + L^u_E + B) - L^u_E - \omega_E] + \alpha B,$$

the only difference is that each party is compensated for the cost of bank borrowing. I now describe resulting outcomes. Since my purpose is to illustrate poor entrepreneurs’ access to credit, attention is restricted to the range of wealth levels where entrepreneurs are credit rationed by the bank sector\textsuperscript{21}. Remaining cases are briefly discussed in the paper’s final section.

Poor entrepreneurs and poor moneylenders will be credit rationed by the bank. Here the temptation to divert for each of them is too strong to permit bank lending supporting first-best investment. As the surplus of the bank transaction accrues entirely to the entrepreneur and the moneylender, the residual return to an investment increases if both take bank credit. Specifically, the entrepreneur exhausts her bank credit line in addition to borrowing the maximum amount made available by the moneylender. Similarly, the moneylender utilizes all available bank funds and his own capital to service the entrepreneur. Hence, the credit limits will be given by the following binding constraints of the entrepreneur and the moneylender, depending on the bargaining outcome

$$\alpha [Q (I) - \bar{L}_E - \bar{L}_M - \omega_M] + (1 - \alpha) \omega_E = \phi (\omega_E + \bar{L}_E)$$

and

$$(1 - \alpha) [Q (I) - \bar{L}_E - \bar{L}_M - \omega_E] + \alpha \omega_M = \phi (\omega_M + \bar{L}_M),$$

with $I = \omega_E + \bar{L}_E + \omega_M + \bar{L}_M$. To induce the entrepreneur to take informal credit, I assume that $\alpha > \hat{\alpha}$, where $\hat{\alpha}$ denotes the threshold at which she is indifferent between exclusive bank borrowing and obtaining bank and moneylender funds\textsuperscript{22}.

\textsuperscript{19} Since returns are claimed by the bank even if the bank’s credit has been diverted, it is never optimal for the entrepreneur to borrow from the moneylender while diverting bank funds.

\textsuperscript{20} Similar to the entrepreneur, the moneylender faces a binary choice. If he decides to lend all his bank funds in order to repay in full, he earns at least 1, while diversion grants him only $\phi$. If he lends too little to repay the bank loan in full, he may as well divert all funds, since any additional returns are claimed by the bank.

\textsuperscript{21} Specifically, entrepreneurs’ wealth is confined to the range at which they receive the monopoly bank’s floor utility. This keeps the analysis tractable when comparing across bank market structure.

\textsuperscript{22} Results are qualitatively similar when $\alpha = \hat{\alpha}$.

---

\textsuperscript{12} Fondazione Eni Enrico Mattei Working Papers, Art. 330 [2009]

http://services.bepress.com/feem/paper330
For a sufficiently high level of wealth, the moneylender’s return from extending a loan in support of first best and repaying the bank beats the diversion utility and his incentive constraint no longer binds. Strikingly, the moneylender makes no profit in this instance. To see this, suppose the surplus to be shared exceeds the sum of the entrepreneur’s and the moneylender’s outside option (equivalent of diverting bank funds and gaining the moneylender’s wealth). Then the bank will offer the entrepreneur a better contract that simultaneously increases her value of diversion and reduces the surplus to be shared with the moneylender. As long as the surplus is positive, the entrepreneur refrains from diversion in equilibrium although the moneylender will have to concede by lowering his price of credit. This continues until the entrepreneur and the moneylender obtain their respective outside options. The argument rests on the fact that the entrepreneur never finds it optimal to borrow from the moneylender while diverting bank funds, since any additional return is claimed by the bank. Opportunistic behavior on the part of the entrepreneur is thus equivalent with loss of business for the moneylender. The moneylender therefore earns \( \omega_M \), while the entrepreneur remains on her incentive constraint.

Hence, the entrepreneur’s credit limit solves, independent of the bargaining outcome

\[
Q (\omega_E + \bar{L}_E + \omega_M + L_M) - \bar{L}_E - L_M - \omega_M = \phi (\omega_E + \bar{L}_E), \tag{12}
\]

while the investment is given by \( I = I^* \).

When the moneylender is rich enough to self-finance larger parts (or the entire amount) of first-best investment he no longer acquires bank funds. Here the entrepreneur borrows from both a bank and a self-financed moneylender. The entrepreneur’s incentive constraint is still determined by (12), with \( L_M + \omega_M \) replaced by \( B \leq \omega_M \) and \( I = I^* \).

Equilibrium outcomes are summarized in Proposition 2.

**Proposition 2:** For all \( \phi > \phi_i \) and \( \omega_E < \omega^c_E \), entrepreneurs borrow from a bank and a bank-financed moneylender and invest \( I < I^* \) if \( \omega_M < \omega^c_M \) and \( I^* \) if \( \omega_M \in [\omega^c_M, \bar{\omega}^c_M] \). Entrepreneurs borrow from a bank and a self-financed moneylender and invest \( I^* \) if \( \omega_M \geq \bar{\omega}^c_M \).

Moneylenders complement the bank sector in providing external capital. At levels of wealth below \( \omega^c_M \), informal lenders retain some rent to satisfy their lending incentives, whereas entrepreneurs receive the entire surplus generated by the moneylenders’ intermediation activity for wealth above \( \omega^c_M \). Borrowing from a moneylender is thus beneficial in two ways. It raises the investment of bank-rationed entrepreneurs (\( \omega_E < \omega^c_E \)) by making additional funds available. It also permits banks to lend more directly to entrepreneurs, as informal finance disciplines them. Moreover, informal lenders’ edge in avoiding diversion makes sufficiently wealthy moneylenders hostages of the bank sector by forcing them to lend at zero rate of interest.

---

23 The entrepreneur’s credit limit cannot be lower in equilibrium. Otherwise, there would exist a bank contract with a lower limit and a positive informal interest rate preferred by the bank as well as the moneylender.

24 While the entrepreneur could satisfy her needs by only taking informal credit, she borrows from both sectors as I assume that she accepts the first available contract if indifferent. A similar outcome obtains if both lenders offered their contracts simultaneously and moneylenders’ monitoring cost was positive and constant returns to scale.

25 The threshold \( \omega^c_E \) refers to the debt capacity at which first best is realized without informal funds.
3.3 The Rent-Extraction Effect: A Moneylender and a Monopoly Bank

Moneylenders’ monitoring ability also helps banks reduce agency cost by allowing them to channel credit through the informal sector. The implication is that all external funding can be provided by the moneylender if the bank is a monopolist. To see this, recall that a high interest rate increases the monopoly bank’s payoff and the entrepreneur’s incentive to default and so poor bank customers earn rent to avoid diversion of bank credit. The existence of moneylenders modifies this trade-off. The moneylender’s monitoring advantage implies that channeled bank capital saves the incentive rent the bank otherwise share with the poor entrepreneur. On the other hand, forwarded bank money comes at a cost as the bank forgoes part of its surplus to prevent the informal lender from cheating the bank.

Specifically, when the entrepreneur and the moneylender are poor the monopolist lends to both, with the moneylender forwarding all his funds to the entrepreneur. Similar to Proposition 1, the entrepreneur and the moneylender are awarded floor contracts granting them utility above their outside option of pursuing the entrepreneur’s project on their own. At this level of wealth, informal finance enables the bank to decrease the entrepreneur’s net surplus and to minimize the aggregate bank loan needed to satisfy the sub-optimal investment. The bank refrains from channeling the entire loan through the informal sector, however, as the moneylender’s temptation to divert formal credit is too large. The binding incentive constraints and the first-order condition of the bank’s profit expression determine credit extended, $L_E$ and $L_M$, and the aggregate repayment $D$.

More precisely, accounting for the bargaining outcome

$$\alpha [Q (I) - D - \omega_M] + (1 - \alpha) \omega_E = \phi (\omega_E + L_E), \quad (13)$$

$$\alpha [Q (I) - D - \omega_E] + \alpha \omega_M = \phi (\omega_M + L_M), \quad (14)$$

and

$$Q' (I) - (1 + \phi) = 0, \quad (15)$$

with $I = \omega_E + L_E + \omega_M + L_M$. The bank charges an aggregate price, $D = D_E + D_M$, paid in proportion to the share of the surplus kept by each respective borrower.

As the informal lender’s debt capacity improves, his participation and incentive constraint both bind at some point. The increase in moneylender wealth allows the bank to reduce the poor entrepreneur’s part of the aggregate loan to save on the incentive rent shared with her to prevent diversion. Specifically, for the same level of investment [as given by equation (15)], $L_E$ is decreased in step with a climbing $\omega_M$ until the entire loan is extended to the moneylender, giving rise to credit market segmentation. In this instance, the moneylender’s repayment obligation $D_M$ solves the binding participation constraint, accounting for the bargaining outcome

$$(1 - \alpha) [Q (I) - D_M - \omega_E] + \alpha \omega_M = (1 - \alpha) [Q (\omega_E + \omega_M) - \omega_E] + \alpha \omega_M, \quad (16)$$

while the equilibrium loan size $L_M$ satisfies

$$(1 - \alpha) [Q (\omega_E + \omega_M) - \omega_E] + \alpha \omega_M = \phi (\omega_M + L_M), \quad (17)$$
with \( I = \omega_E + \omega_M + L_M \). The participation constraint ensures at least the utility associated with the moneylender self financing the project.

A rich enough moneylender is able to support first best. Equation (16) still determines \( D_M \) and the investment level is given by \( I = I^* \). Finally, when the moneylender is sufficiently wealthy to self finance the investment, the bank and the moneylender compete in the same fashion as described by equation (12) above. Proposition 3 recapitulates the main findings.

**Proposition 3:** For all \( \phi > \phi(\omega_i) \) and \( \omega_E < \omega_E^m \), entrepreneurs borrow from a bank and a bank-financed moneylender and invest \( I = I' \) as given by equation (15) if \( \omega_M < \omega_M^m \). Entrepreneurs borrow exclusively from a bank-financed moneylender and invest \( I = I' \) if \( \omega_M \in [\omega_M^m, \omega_M^s] \) and \( I^* \) if \( \omega_M \in [\omega_M^s, I^* - \omega_E] \). Entrepreneurs borrow from a bank and a self-financed moneylender and invest \( I^* \) if \( \omega_M \geq I^* - \omega_E \).

Moneylenders become increasingly important as their debt capacity improves. While the informal sector merely supplies additional capital at levels of wealth below \( \omega_M^m \), it is poor entrepreneurs’ only source of credit for wealth in-between \( \omega_M^m \) and \( I^* - \omega_E \). Again, moneylenders are beneficial in two ways. As in the competitive case, they help raise bank-rationed (\( \omega_E < \omega_E^m \)) entrepreneurs’ investment.\(^{26}\) Moreover, the informal credit channel allows banks to boost their profit by reducing the surplus otherwise shared with agency prone entrepreneurs.

### 3.4 Empirical Support and Some Equilibrium Characteristics

The preceding analysis demonstrates that moneylenders perform multiple roles—by raising investment, disciplining borrowers, and facilitating banks’ rent extraction. If we take bank market competition as an indicator of the type of activity informal lenders primarily engage in, the model’s predictions shed light on a series of empirical studies on formal-informal sector interactions. Evidence from China shows that informal finance is more prevalent in the central and the northwest regions where banking competition is scant, while less important in the coastal region where banks are more competitive (Cull and Xu, 2005; Ayyagari et al., 2008; Cheng and Degryse, 2008). Similarly, in Giné’s (2007) study of 2880 households and 606 small businesses in rural Thailand, borrowers are more likely to access the informal sector exclusively when bank competition decreases. These findings are in accordance with the theory’s hypothesis that informal finance is more important in concentrated bank markets where the rent-extraction effect prevails. Bank market power in the adjacent bank market and subsequent credit market segmentation also offers one potential explanation for the evidence reviewed by Banerjee and Duflo (2007), where 95 percent of all borrowers in a survey of 2000 households living below $2 a day in urban Hyderabad, India get loans from informal lending sources despite the existence of a (possibly monopolistic) formal bank sector.

A final observation concerns how banks respond to increases in wealth.

\(^{26}\) The threshold \( \omega_E^m \) refers to the debt capacity at which the entrepreneurs’ incentive and participation constraint both bind. The corresponding investment level in turn varies depending on the wealth of the moneylender.
Proposition 4: (i) For a moneylender with wealth below $\omega_M^{c}$, competitive bank credit ($L_E$ and $L_M$) increase in the entrepreneur’s and the moneylender’s wealth ($\omega_E$ and $\omega_M$). (ii) For $\omega_M < \omega_M^{m}$, aggregate monopoly bank credit ($L_E + L_M$) decreases in $\omega_E$ and $\omega_M$, while $\omega_E$ has an indeterminate effect on $L_E$ and $\omega_M$ has an indeterminate effect on $L_M$. (iii) For $\omega_M \in (\omega_M^{m}, \bar{\omega}_M)$, $L_M$ increases in $\omega_E$ and $\omega_M$.

If moneylenders discipline borrowers, the previous section predicted that banks would lend more liberally to entrepreneurs. A similar finding concerns the effect of an increase in informal lenders’ debt capacity. Climbing moneylender and/or entrepreneurial wealth raises the profitability of lending and investment of competitive bank funds relative to diversion, resulting in more bank credit extended to moneylenders and entrepreneurs. In a sense, richer moneylenders provide entrepreneurs with a stronger commitment device vis-à-vis banks, as increasing informal sector assets make banks more willing to lend directly to poor entrepreneurs.

By contrast, when poor moneylenders assist monopoly banks in extracting rent, the bank reaps the entire benefit of marginal wealth improvements. Although the precise effect depends on each borrower’s share of the surplus, wealth increases are met by a parallel decrease in the aggregate loan, $L_E + L_M$, to satisfy the sub-optimal investment level [given by equation (15)]

When the moneylender’s participation and incentive constraint hold simultaneously, banks channel all their capital through the informal lender. Here, increases in moneylender and/or entrepreneurial wealth improve the moneylender’s outside option of self financing the project, as well as that of diverting bank funds (in case of an increase in moneylender wealth). To satisfy both constraints at equality, bank credit climbs with wealth.

4 Welfare

As anticipated, informal finance improves efficiency in two ways. By lending when banks are restrained by agency issues, moneylenders expand the supply of funds per se. Moreover, if the availability of informal credit disciplines borrowers, this eases poor entrepreneurs’ incentive problem versus the bank sector. If moneylenders facilitate banks’ rent extraction, however, the second effect is absent as increases in borrower wealth are expropriated by the monopoly bank.

Proposition 5: (i) The coexistence of formal and informal finance increases efficiency. (ii) Efficiency is higher if moneylenders discipline borrowers (competitive banking) as opposed to facilitate banks’ rent extraction (monopoly banking), that is, $\bar{\omega}_M > \bar{\omega}_M^{c}$.

While part two of Proposition 5 is consistent with the intuition that a monopoly bank, in order to extract more rent lends less to maintain proper incentives, moneylenders feature distinctly

For example, a boost in moneylender wealth causes a small increase in his bank loan if most of the surplus accrues to the entrepreneur in order to satisfy the moneylender’s incentive constraint. Meanwhile, the entrepreneur’s loan is reduced allowing the bank to seize the entire gain of the improved debt capacity.
in deciding the order of magnitude (see Figure 2). First, if informal lenders discipline entrepren
erreurs, banks boost the wealth accumulation of their borrowers implying that less informal
wealth attains the efficient outcome. Second, since sufficiently rich moneylenders are forced to
lend at the opportunity cost of funds under competitive banking, all incremental wealth gains
contribute to improved investment incentives. Finally, credit market segmentation requires a
higher debt capacity of the moneylender as his outside option alone—not the combined surplus
of the entrepreneur and the informal lender—determines when first best is reached.

To understand how welfare is distributed in the economy, I need to establish the exact
supply of bank funds.

**Lemma 1:** (i) Entrepreneurs obtain more funds from competitive banks for \( \omega_M < I^* - \omega_E \)
and the same amount regardless of bank market structure otherwise. (ii) There exists a thresh-
old \( \hat{\omega}_M(\phi) \in (\omega^c_M, \omega^c_M) \) such that moneylenders with wealth in-between \( \hat{\omega}_M \) and \( I^* - \omega_E \)
obtain more funds from a monopoly bank and moneylenders with wealth below \( \hat{\omega}_M \) obtain
more funds from competitive banks.

Competitive banks supply more credit to entrepreneurs unless moneylenders self finance, in
which case, bank market structure is irrelevant. Moneylenders also take more competitive
bank credit up to first best (\( \omega^c_M \)), then reduce their loan in step with rising wealth. At \( \omega^c_M \), a
monopoly bank continues to extend additional funds, however, as the efficient outcome remains
to be attained (\( \omega_M < \omega^c_M \)). I now determine how bank-financed moneylenders affect welfare.

**Proposition 6:** (i) Entrepreneurs and poor moneylenders, \( \omega_M < \omega^c_M \), are better off when the
disciplinary effect dominates, whereas banks and sufficiently wealthy moneylenders, \( \omega_M > \omega^c_M \), are better off when the rent-extraction effect dominates. (ii) Entrepreneurs prefer a
monopoly bank in isolation over the coexistence of a moneylender and a monopoly bank.

Informal finance supports asset growth of entrepreneurs when the disciplinary effect prevails.
The reason is twofold. Bank competition transfers the entire surplus to the bank borrowers,
allowing more credit to be extended. Moneylenders reinforce this effect by further expanding credit provision and alleviating poor entrepreneurs’ incentive problem. Competition also adds value to poor moneylenders as they receive more bank funds. By contrast, banks and wealthier moneylenders are better off when the rent-extraction effect dominates and credit markets become segmented. This is because the segmented outcome preserves the market power that moneylenders’ enforcement advantage grants them (\(\alpha\) remains unchanged), whereas they are forced to give up any potential rent under competitive banking (\(\alpha\) goes to 1).

The proposition’s second part captures a darker side of informal finance. Poor entrepreneurs receive less funding and consequently lower floor utility from the monopoly bank (for a given investment) if it also extends credit to the moneylender. This loss is sustained when moneylenders provide all external capital, as the bank’s incentive rent yields a value above entrepreneurs’ outside option of doing the project alone with the informal lender. Effectively: the only thing worse than having to borrow from a monopoly bank is to be left in the hands of a moneylender.

If rich moneylenders and bankers have more say over bank market structure than poor entrepreneurs, Proposition 6 provides a political-economy explanation as to why monopoly banking is a pervasive feature of less developed credit markets. In line with the theory, Rajan and Ramcharan (2008) find that bank markets in the early twentieth century United States were more concentrated in counties with wealthy landowners who often engaged in local lending to farmers. These landlords frequently had ties with the local bank and the local store (that offered credit) and, as the model predicts, were against bank deregulation. Note that pro-competitive measures are difficult to implement in such circumstances, not only for the reason that powerful interests stand to lose rent, but because the gainers of such a reform (the poor entrepreneurs), cannot actually compensate banks and moneylenders due to their wealth constraints.

5 The Prevalence of Informal Finance

According to Germidis et al. (1991) and Nissanke and Aryeetey (1998) informal transactions, such as loans made by moneylenders, traders, landlords, and family, account for between one third and three quarters of total credit in Asia and sub-Saharan Africa. The discussion so far indicates that bank-financed informal lenders are more important providers of external capital when they facilitate banks’ rent extraction. It remains to establish this variation formally.

**Proposition 7:** The ratio of informal credit to investment is higher when the rent-extraction effect dominates and \(\omega_M \in (\hat{\omega}_M, I^* - \omega_E)\) and indeterminate with respect to either effect for wealth below \(\hat{\omega}_M\).

Although entrepreneurs obtain more funds from poor moneylenders if the disciplinary effect is in force and banks compete, they also take additional bank credit (Lemma 1) making the exact prediction imprecise. However, as moneylenders become wealthier the theory yields an unambiguous answer: informal finance should be more important if it facilitates banks’
rent extraction. Proposition 7 is a novel prediction of the theory that combines two empirical facts—less developed credit markets are often characterized by a significant degree of bank market power (Barth et al., 2004; Beck et al., 2004) and a large informal financial sector.

A key premise of the model is that informal finance emerges in response to banks’ inability to perfectly enforce their legal claims. I now examine this issue in some detail.

**Proposition 8:** The ratio of informal credit to investment is increasing in creditor vulnerability \(\phi\) if the disciplinary effect dominates and nonincreasing in \(\phi\) if the rent-extraction effect dominates for \(\omega_M \in [\omega_M^c, I^* - \omega_E]\).

Rich enough moneylenders emerge as substitutes for institutional quality when they discipline entrepreneurs under competitive banking. Deteriorating creditor protection boosts the profitability of diversion relative to investment for poor entrepreneurs, inducing a shift to agency-free informal credit. By contrast, if moneylenders are the only providers of external funds and first best remains out of reach, worse legal protection cuts the funding of the monopoly bank’s sole customer to avoid opportunistic behavior. Conversely, if the efficient outcome is attained, better institutions are irrelevant since diversion no longer tempts the moneylender.

In sum, less efficient creditor protection increases the prevalence of informal finance when entrepreneurs obtain money from both financial sectors, while the opposite holds true if moneylenders alone provide all external capital. Using firm-level data for 26 countries in Eastern Europe and Central Asia, Dabla-Norris and Koeda (2008) broadly confirms Proposition 8 by showing that the empirical relationship between institutions and informal credit is indeterminate, while bank lending contracts as creditor protection worsens.

### 6 Informal Interest Rates

Banerjee (2003) reviews evidence of informal interest rates ranging from 0 to 80 percent annually in India, between 26 and 43 percent in Thailand, and between 18 and 200 percent in Pakistan. Can the model explain such heterogeneity?

**Proposition 9:** (i) Bank-rationed moneylenders always charge a positive rate of interest. (ii) The informal interest rate is higher when the rent-extraction effect dominates and increases as credit markets become segmented.

My theory offers two main explanations for the reported variation. First, poor moneylenders charge positive rates of interest even if the adjacent bank market is perfectly competitive. This

---

29 At levels of wealth below \(\omega_M^c\), the effect of changes in creditor protection is sensitive to how the project gains are shared between the entrepreneur and the moneylender.

30 Dabla-Norris and Koeda do not account for bank market power, barring any conclusion in this respect.

31 For additional evidence, see also Udry (1990, 1994) and La Ferrara (2003) for the case of Africa and Das-Gupta et al. (1989), Aleem (1990), and Siamwalla et al. (1990) for the case of Asia.
is because the price of credit reflects the incentive rent moneylenders receive to ensure prudent behavior when forwarding bank funds. Competition from the bank sector is thus softened as excessive lending to poor entrepreneurs and/or poor moneylenders would result in diversion.

Second, the exact magnitude of informal interest rates depends on bank market power and the subsequent function performed by informal lenders. At low debt capacities, diminishing returns to scale imply that entrepreneurs pay lower informal interest if banks compete due to the larger informal loan size that competition generates. As moneylender wealth climbs, informal lenders are forced to lend at the opportunity cost of funds if they help discipline borrowers. Meanwhile, the segmented outcome that results from facilitating banks’ rent extraction preserves moneylenders’ market power. Hence, moneylenders assisting banks in extracting rent always charge higher interest rates. Moreover, credit market segmentation increases interest rates further as entrepreneurs lose their outside option when dealing exclusively with the informal sector. In effect, without the bank’s incentive rent entering the bargaining, entrepreneurs are charged the same price that they would have paid if no bank were present. Terms offered to the same borrower may thus vary from an effective interest rate of zero to very high rates\(^{32}\).

Another empirical regularity that can be understood within my framework concerns the observed variation between informal and formal lending rates. Define the lending-rate gap as

\[
r^g = \left( \frac{R}{B} - \frac{D_M}{L_M} \right),
\]

\(^{(18)}\)

\(r^g\) is the difference between informal interest charged and formal interest paid by moneylenders. To simplify, I restrict attention to the interval at which credit market segmentation occurs.

**Proposition 10:** The lending-rate gap, \(r^g\), is: (i) Positive when the disciplinary effect dominates for \(\omega_M < \omega_M^*\) and zero for \(\omega_M \geq \omega_M^*\). (ii) Positive when the rent-extraction effect dominates if \(\alpha < \tilde{\alpha}\) and negative if \(\alpha > \tilde{\alpha}\) for \(\tilde{\alpha} \in (\hat{\alpha}, 1)\).

If banks compete and moneylenders discipline borrowers, there is a net increase in the price per dollar of credit extended unless moneylenders are held hostage by the banks and forced to offer zero-interest loans. Intuitively, poor moneylenders enjoy access to inexpensive competitive bank credit and keep some gains of the investment project, while the entire surplus accrues to the entrepreneurs when moneylenders’ debt capacity improves. By contrast, although informal interest rates are higher when moneylenders assist the monopoly bank, informal lenders are charged more on the credit they take. However, since credit market segmentation preserves moneylenders’ market power, a net increase in the price per dollar of credit extended will occur if moneylenders keep a sufficiently large share of the project’s surplus\(^{33}\).

---

\(^{32}\) The order of magnitude depends on the curvature of the production function and the debt capacity of the entrepreneur and the moneylender (with the informal interest rate decreasing in borrower and moneylender wealth).

\(^{33}\) On the other hand, if entrepreneurs retain most of the surplus, \(\alpha > \tilde{\alpha}\), moneylenders charge the same price per dollar as the monopoly bank on the bank’s part of the informal loan and a lower rate on any remaining internal capital, resulting in lower overall interest (than the bank rate) and a negative interest differential.
7 Discussion and Concluding Remarks

A worthwhile question is why the bank does not merge with the moneylender, making him the local branch manager of the bank? The straightforward answer is that "bringing the market inside the firm" at best replicates the market outcome, as the branch manager has to be incentivized to act responsibly with the bank funds. However, the merger also adds a new dimension, the employer-employee relationship, which opens up for opportunistic behavior on the part of the bank as well.\[34\] Moreover, if there is output uncertainty and monitoring costs are positive, bank managers may be tempted to reduce monitoring and subsequent monitoring costs only to claim that entrepreneurial failure was due to exogenous events. (Allowing for output uncertainty would not change the theory’s main predictions.) Hence, the overall effect is likely to be efficiency reducing, confirming why this kind of organizational design is uncommon in developing credit markets.\[35\]

A related concern is whether the key insights would be altered if informal monitoring was less efficient, if other sharing rules governed the moneylender’s and the entrepreneur’s exchange, or if agents engaged in side payments? As regards the first objection, suppose the entrepreneur fails to invest a fraction \(\delta \in (0, 1)\) of the moneylender’s funds.\[36\] It can be shown that for \(\delta\) sufficiently small, equilibrium outcomes remain the same. Pertaining to the choice of sharing rule, the Nash Bargaining solution produces an efficient outcome similar to Coasian bargaining since utility is transferable. Any sharing rule therefore yields quantitatively similar results in terms of the ensuing investment. Finally, as briefly noted, side payments do not change the equilibrium outcomes since poor entrepreneurs or moneylenders are unable to compensate the other party and/or the bank due to their wealth constraints. That is, available funds are always used most efficiently in production.

As the model stands, the informal lender’s occupational choice is restricted to lending money. In a more general setting he may have additional sources of income, such as holding land or trading. This will not weaken the results. Complementary sources of income (and/or collateral) make it less tempting to behave opportunistically, enabling the bank to extend more funds or extract more rent. The model’s predictions thus apply to a broader class of phenomena characterized as informal finance, including credit extended by traders, landlords, and family.\[37\]

If rising entrepreneurial wealth is allowed, rich entrepreneurs only take competitive bank credit. With a monopoly bank, little changes if the entrepreneur’s wealth climbs and wealth disparity is maintained. Here the bank is indifferent between dealing with the (relatively) richer moneylender alone and lending a small amount to the entrepreneur and the remainder to the moneylender. If the entrepreneur is the richer party, the outcome resembles the one analyzed

---

34 Similar in spirit to Williamson’s (1985) arguments of why “selective interventions” are hard to implement.
36 The value \(\delta\) could be a deadweight loss or, alternatively, a benefit accruing directly to the entrepreneur.
37 Additional reasons why a landlord engages in lending include the linking of credit and land transactions to increase tenants’ work effort, as in Braverman and Stiglitz (1982), Mitra (1983), and Basu (1987) amongst others.
in detail above, now with the bank gradually reducing its loan to the poor moneylender. If entrepreneurs and moneylenders are equally affluent though short of first best, both receive credit. Finally, rich entrepreneurs only take monopoly bank credit.\footnote{Conning (2001) and Giné (2007) document that rich borrowers resort exclusively to formal lenders.} Except in this last case, the conclusion’s regarding welfare, prevalence of informal finance, and pricing of informal transactions remain qualitatively unchanged.

The model’s findings yield some useful policy insights. In general, stronger institutions improve efficiency, but not necessarily in favor of poor borrowers. If credit markets are segmented, pro-competitive measures may be more useful than enhanced legal protection of banks if the objective is to ease borrowers’ access to bank finance. Bank competition not only expands financial access, but also reduces the interest charged by informal lenders. This is consistent with the recent conclusions of the Indian Committee on Financial Sector Reforms (CFSR), which argues that increased bank competition is paramount to improving poor people’s financial standing (CFSR, 2008). Programs that strengthen borrowers outside options (similar to the empowerment strategies of poor tenants documented in Banerjee et al., 2002) further diminish the reliance on credit provided by the informal sector and the monopoly banks. This points to the importance of alternative credit schemes as pioneered by the growing microfinance movement.\footnote{See Morduch (1999) for a survey of the issue.} In fact, microfinance programs may present a more viable alternative if powerful vested interests (in the form of wealthy informal lenders and monopoly banks) are opposed to bank market reforms.
Appendix

The following result will be helpful in the subsequent analysis.

**Lemma A2:** \( Q' (\omega_E + \bar{L}_E) - (1 + \phi) < 0 \).

**Proof.** Part (i): When the entrepreneur (henceforth E) borrows exclusively from a competitive bank (henceforth B) and the credit limit binds,

\[
Q (\omega_E + \bar{L}_E) - \bar{L}_E - \phi (\omega_E + \bar{L}_E) = 0. \tag{A1}
\]

This constraint is only binding if \( Q' (\omega_E + \bar{L}_E) - (1 + \phi) < 0 \). Otherwise, \( \bar{L}_E \) could be increased without violating the constraint. ■

**Proof of Proposition 1**

Proposition 1 is proved in the main text, except for the comparative static results, the existence, and the uniqueness of \( \omega^c_E, \omega^m_E, \) and \( \bar{\omega}^m_E \).

**Lemma A3:** There exist unique thresholds \( \omega^c_E (\phi) > 0, \omega^m_E (\phi), \) and \( \bar{\omega}^m_E (\phi) \) such that:

(i) \( Q (\omega_E + \bar{L}_E) - \bar{L}_E - \phi (\omega_E + \bar{L}_E) = 0 \) for \( \omega_E = \omega^c_E (\phi) \) and \( \omega_E + \bar{L}_E = I^* \);

(ii) \( \phi (\omega_E + L_E) - Q (\omega_E) = 0 \) for \( \omega_E = \omega^m_E (\phi) \) and \( \omega_E + L_E = I \), with the investment level given by equation (3) in the main text;

(iii) \( \phi (\omega_E + L_E) - Q (\omega_E) = 0 \) for \( \omega_E = \bar{\omega}^m_E (\phi) \) and \( \omega_E + L_E = I^* \); and

(iv) \( \bar{\omega}^m_E (\phi) > \omega^m_E (\phi) > 0 \) and \( \bar{\omega}^m_E (\phi) > \omega^c_E (\phi) \).

**Proof.** Part (i): The threshold \( \omega^c_E \) is the smallest wealth level that satisfies \( \omega_E + \bar{L}_E = I^* \). As equation (A1) yields the maximum incentive-compatible investment level, \( \omega^c_E \) satisfies

\[
Q (I^*) - I^* (1 + \phi) + \omega^c_E = 0. \tag{A2}
\]

The threshold is unique if \( \bar{L}_E \) is increasing in \( \omega_E \). Differentiating (A1) with respect to \( \bar{L}_E \) and \( \omega_E \) I obtain

\[
\frac{d\bar{L}_E}{d\omega_E} = \frac{\phi - Q' (\omega_E + \bar{L}_E)}{Q' (\omega_E + \bar{L}_E) - (1 + \phi)} > 0,
\]

where the inequality follows from Lemma A2, \( Q' (I) \geq 1 \), and \( \phi < 1 \). Finally, \( \omega^c_E > 0 \) is a result of the assumption that \( \phi > \phi_L \) [equation (I)].

Part (ii): The threshold \( \omega^m_E \) is the smallest wealth level at which E’s incentive constraint equals her participation constraint allowing E to invest \( \omega_E + L_E = I \), with I given by equation (3) in the main text. Thus, \( \omega^m_E \) satisfies

\[
\phi I - Q (\omega^m_E) = 0. \tag{A3}
\]
The threshold is unique if \( L_E \) is decreasing in \( \omega_E \) when the equilibrium is given by equations (3) and (4) in the main text. Differentiating (3) and (4) with respect to \( L_E \) and \( \omega_E \) using Cramer’s rule I obtain
\[
\frac{dL_E}{d\omega_E} = -1.
\]
Finally, \( \omega^m_E > 0 \) follows from the assumption that \( \phi > \phi^c \).

Part (iii): The proof is analogous to the proof of Part (ii) and omitted.

Part (iv): Solving for \( \omega^m_E \) and \( \bar{\omega}^m_E \) and combining the two expressions, yields
\[
Q \left( \bar{\omega}^m_E \right) \times I' = Q \left( \omega^m_E \right) I^*,
\]
with \( I' \) given by equation (A1) in the main text. By concavity, \( I^* > I' \) and hence \( \omega^m_E > \omega^m_E \). Solving for \( \omega^m_E \) and \( \bar{\omega}^m_E \) and combining the two expressions, yields \( Q \left( I^* \right) - I^* = Q \left( \bar{\omega}^m_E \right) - \omega^c_E \), where \( \omega^m_E > \omega^c_E \) follows from concavity. □

**Lemma A4:** (i) If \( \omega_E \leq \omega^c_E \) under competitive banking then \( L_E \) and \( I \) increase in \( \omega_E \). (ii) If \( \omega_E \leq \omega^m_E \) with a monopoly bank then \( L_E \) decreases in \( \omega_E \) and \( I \) is independent of \( \omega_E \); if \( \omega_E \in (\omega^m_E, \bar{\omega}^m_E) \) then \( L_E \) and \( I \) increase in \( \omega_E \).

**Proof.** Part (i): The proof that \( d\bar{L}_E/d\omega_E > 0 \) is provided in Lemma A3. As equation (A1) also determines the investment level, \( dI/d\omega_E > 0 \) follows.

Part (ii): When \( \omega_E \leq \omega^m_E \), the proof that \( dL_E/d\omega_E < 0 \) is provided in Lemma A3. Differentiating equations (3) and (4) in the main text and the investment condition, \( \omega_E + L_E = I \), with respect to \( I \) and \( \omega_E \) using Cramer’s rule I obtain
\[
\frac{dI}{d\omega_E} = 0.
\]
When \( \omega_E \in (\omega^m_E, \bar{\omega}^m_E) \), the relevant equations are given by (5) in the main text, the binding participation constraint, \( Q(\omega_E + L_E) - D_E = Q(\omega_E) \), and the investment condition, \( \omega_E + L_E = I \). Differentiating (5), the binding participation constraint, and the investment condition with respect to \( L_E, I, \) and \( \omega_E \) using Cramer’s rule I obtain
\[
\frac{dL_E}{d\omega_E} = Q'(\omega_E) - \frac{\phi}{\phi} > 0
\]
and
\[
\frac{dI}{d\omega_E} = Q'(\omega_E) \frac{\phi}{\phi} > 0,
\]
where the first inequality follow from \( Q'(I) \geq 1 \) and \( \phi < 1 \). □

**Proof of Proposition 2**

I show the existence and the uniqueness of \( \omega^c_E, \omega^c_M, \) and \( \bar{\omega}^c_M \) and proceed with the equilibrium outcomes.
Lemma A5: There exist unique thresholds $\omega_E^c(\phi) > 0$, $\omega_M^c(\phi)$, and $\bar{\omega}_M^c(\phi)$ such that:

(i) $Q(\omega_E + \bar{L}_E) - \bar{L}_E - \phi (\omega_E + \bar{L}_E) = 0$ for $\omega_E = \omega_E^c(\phi)$ and $\omega_E + \bar{L}_E = \ell^*$;

(ii) $\alpha [Q(\omega_E + \bar{L}_E + \omega_M + \bar{L}_M) - \bar{L}_E - \bar{L}_M - \omega_M] + (1 - \alpha) \omega_E - \phi (\omega_E + \bar{L}_E) = 0$ and $(1 - \alpha) [Q(\omega_E + \bar{L}_E + \omega_M + \bar{L}_M) - \bar{L}_E - \bar{L}_M - \omega_E] + \alpha \omega_E - \phi (\omega_M + \bar{L}_M) = 0$ for $\omega_M = \omega_M^c(\phi)$ and $\omega_E + \bar{L}_E + \omega_M + \bar{L}_M = \ell^*$;

(iii) $Q(\omega_E + \bar{L}_E + \omega_M) - \bar{L}_E - \omega_M - \phi (\omega_E + \bar{L}_E) = 0$ for $\omega_M = \bar{\omega}_M^c(\phi)$ and $\omega_E + \bar{L}_E + \omega_M = \ell^*$; and

(iv) $\bar{\omega}_M^c(\phi) > \omega_M^c(\phi) > 0$.

Proof. Part (i): The proof is provided in Lemma A3.

Part (ii): The threshold $\omega_M^c$ is the smallest wealth level that satisfies $\omega_E + \bar{L}_E + \omega_M + \bar{L}_M = \ell^*$ when $E$ and the moneylender (henceforth $M$) utilize bank funds as given by equations (10) and (11) in the main text. Using (10) and (11) to solve for the maximum incentive-compatible investment level I have that, for a given level of $E$'s wealth, $\omega_E$, $\omega_M^c$ satisfies

$$Q(\ell^*) - \ell^*(1 + \phi) + \omega_E + \omega_M^c = 0. \tag{A4}$$

The threshold is unique if both $\bar{L}_E$ and $\bar{L}_M$ are increasing in $\omega_M$. Differentiating (10) and (11) with respect to $\bar{L}_E$, $\bar{L}_M$ and $\omega_M$ using Cramer’s rule I obtain

$$\frac{d\bar{L}_E}{d\omega_M} = \frac{\alpha [Q'(\ell) - 1]}{\phi [1 + \phi - \alpha Q'(\ell)]} > 0$$

and

$$\frac{d\bar{L}_M}{d\omega_M} = \frac{\phi [Q'(\ell) - \phi] - \alpha [Q'(\ell) - 1]}{\phi [1 + \phi - \alpha Q'(\ell)]} > 0,$$

where the inequalities follow from Lemma A2, $Q'(\ell) \geq 1$, and $\phi < 1$.

Part (iii): The threshold $\bar{\omega}_M^c$ is the smallest wealth level that satisfies $\omega_E + \bar{L}_E + \omega_M = \ell^*$ at which $M$ is able to self finance $E$. Thus, for a given level of $E$’s wealth, $\omega_E$, $\bar{\omega}_M^c$ satisfies

$$Q(\ell^*) - \ell^*(1 + \phi) + \omega_E + \bar{\omega}_M^c\phi = 0. \tag{A5}$$

The threshold is unique if $\bar{L}_E$ ($\bar{L}_M$) is increasing (decreasing) in $\omega_M$ when the relevant constraints are given by equation (12) in the main text and the first-order condition $Q'(\ell) - 1 = 0$. Differentiating (12) and the first-order condition with respect to $\bar{L}_E$, $\bar{L}_M$, and $\omega_M$ using Cramer’s rule I obtain

$$\frac{d\bar{L}_E}{d\omega_M} = 0$$

and

$$\frac{d\bar{L}_M}{d\omega_M} = -1.$$

Part (iv): Combining (A4) and (A5), yields $\omega_M^c = \phi \bar{\omega}_M^c$, where $\bar{\omega}_M^c > \omega_M^c$ follows from $\phi < 1$. Finally, $\omega_M^c > 0$ is a result of the assumption that $\phi > \frac{\ell}{\omega_E}$ [equation (6)]. ■
Lemma A6: If (i) $\omega_E < \omega_E^c$ and $\omega_M < \omega_M^c$ then the entrepreneur borrows from a bank and a bank-financed moneylender. If (ii) $\omega_E < \omega_E^c$ and $\omega_M \geq \omega_M^c$ then the entrepreneur borrows from a bank and a self-financed moneylender.

Proof. In what follows, I consider E’s and M’s incentive constraints given that B breaks even. Five distinct cases need to be analyzed as E may borrow from: (1) B exclusively; (2) B and a bank-financed M; (3) a bank-financed M exclusively; (4) a self-financed M exclusively; (5) B and a self-financed M.

Part (i): First, consider $\omega_M < \omega_M^c$. Recognizing the concavity of $Q(I)$ and $Q'(I) \geq 1$, it follows that E and M prefer Case (2) to Cases (3), (4), and (5) for any $\alpha$. Finally, for $\alpha > \hat{\alpha}$ as defined in the main text, E prefers Case (2) to Case (1) as well. Next, when $\omega_M \in [\omega_M^c, \bar{\omega}_M^c)$, $\omega_E + \omega_M$ accounts for the interval of credit lines such that $\omega_M < I^* - \omega_E - L_E$, for a given $\omega_E$ and $\omega_M$. From the main text we know that Case (2) leaves E with the entire surplus, while M is indifferent between lending or not and so Case (2) remains the equilibrium outcome when $\omega_M \in [\omega_M^c, \bar{\omega}_M^c)$.

Part (ii): Here, $\omega_E + \omega_M$ accounts for the interval of credit lines such that $\omega_M \geq I^* - \omega_E - L_E$, for a given $\omega_E$ and $\omega_M$. The only difference from Part (ii) is that M refrains from bank borrowing when he is able to self finance large parts of the first-best investment, making Case (2) irrelevant. Thus, Case (5) is the only possible outcome since in Cases (3) and (4), E would have to share part of a (possibly smaller) surplus with M.

Proof of Proposition 3

I show the existence and the uniqueness of $\omega_E^m, \omega_M^m,$ and $\bar{\omega}_M^m$ and proceed with the equilibrium outcomes.

Lemma A7: There exist unique thresholds $\omega_E^m(\phi), \omega_M^m(\phi),$ and $\bar{\omega}_M^m(\phi)$ such that:

(i) $\phi (\omega_E + L_E) - \alpha Q (\omega_E + B) - (1 - \alpha) \omega_E + \alpha B = 0$ for $\omega_E = \omega_E^m(\phi)$ and $\omega_E + L_E + B = 1$, with the investment level given by equation (15) in the main text;

(ii) $\phi (\omega_M + L_M) - (1 - \alpha) [Q (\omega_E + \omega_M) - \omega_E] - \alpha \omega_M = 0$ for $\omega_M = \omega_M^m(\phi)$ and $\omega_E + \omega_M + L_M = 1$, with the investment level given by equation (15) in the main text;

(iii) $\phi (\omega_M + L_M) - (1 - \alpha) [Q (\omega_E + \omega_M) - \omega_E] - \alpha \omega_M = 0$ for $\omega_M = \bar{\omega}_M^m(\phi)$ and $\omega_E + \omega_M + L_M = I^*$; and

(iv) $\bar{\omega}_M^m(\phi) > \omega_M^m(\phi) > 0$.

Proof. Part (i): The threshold $\omega_E^m$ is the smallest wealth level at which E’s incentive constraint equals her participation constraint allowing E to invest $\omega_E + L_E + B = 1$, with $I$ given by equation (15) in the main text. Thus, for a given level of M’s wealth, $\omega_M, \omega_E^m$ satisfies

$$\phi (I - B) - \alpha Q (\omega_E^m + \omega_M) - (1 - \alpha) \omega_E^m + \alpha \omega_M = 0.$$  \hfill (A6)
The threshold is unique if \( L_E + L_M \) decrease in \( \omega_E \) when the equilibrium is given by equations (13) to (15) in the main text. [The same reasoning applies when \( \omega_M \in (\omega_M, 1^* - \omega_E) \).]

Differentiating (13) to (15) with respect to \( L \) using Cramer’s rule I obtain

\[
\frac{dL_E}{d\omega_E} = \frac{1 - \alpha - \phi}{\phi}
\]

and

\[
\frac{dL_M}{d\omega_E} = \frac{\alpha - 1}{\phi},
\]

with \( dL_E/d\omega_E + dL_M/d\omega_E = -1 \). To show \( \omega^m_E > 0 \), let \( \alpha \rightarrow \hat{\alpha} \) in (A6). This yields \( \phi (I - B) - Q(\omega^m_E) = 0 \), where \( \omega^m_E > 0 \) follows from the assumption that \( \phi > \hat{\phi} \). Then let \( \alpha \rightarrow 1 \). Here, \( \phi (I - B) - Q(\omega^m_E + \omega_M) + \omega_M = 0 \). Note that \( \omega^m_E \) decreases in \( \omega_M \) for \( \omega_M < I^* - \omega_E \). As \( \omega_M \) approaches \( I^* - \omega_E \), I have that \( \phi (I^* - \omega_M) - Q(I^*) + I^* - \omega^m_E = 0 \), which is identical to (A5). If \( \omega^m_E = 0 \) then \( \omega_M = I^* \), but this contradicts \( \omega^m_E < I^* \). Hence, \( \omega^m_E > 0 \).

Part (ii): The threshold \( \omega^m_M \) is the smallest wealth level at which M’s incentive constraint equals his participation constraint allowing an investment of \( \omega_E + \omega_M + L_M = I^* \), with \( I \) given by equation (15) in the main text. Thus, for a given level of E’s wealth, \( \omega_E, \omega^m_M \) satisfies

\[
\phi (I - \omega_E) - (1 - \alpha) [Q(\omega_E + \omega^m_M) - \omega_E] - \alpha \omega^m_M = 0.
\]  

(A7)

The threshold is unique if \( L_E + L_M \) decrease in \( \omega_M \) when the equilibrium is given by equations (13) to (15) in the main text. Differentiating (13) to (15) with respect to \( L_E, L_M, \) and \( \omega_M \) using Cramer’s rule I obtain

\[
\frac{dL_E}{d\omega_M} = -\frac{\alpha}{\phi}
\]

and

\[
\frac{dL_M}{d\omega_M} = \frac{\alpha - \phi}{\phi},
\]

with \( dL_E/d\omega_M + dL_M/d\omega_M = -1 \).

Part (iii): The threshold \( \omega^m_M \) is the smallest wealth level at which M’s incentive constraint equals his participation constraint allowing an investment of \( \omega_E + \omega_M + L_M = I^* \). Thus, for a given level of E’s wealth, \( \omega_E, \omega^m_M \) satisfies

\[
\phi (I^* - \omega_E) - (1 - \alpha) [Q(\omega_E + \omega^m_M) - \omega_E] - \alpha \omega^m_M = 0.
\]  

(A8)

The threshold is unique if \( L_M \) is increasing in \( \omega_M \) when the equilibrium is given by equations (16) and (17) in the main text. Differentiating (16) and (17) with respect to \( L_M \) and \( \omega_M \) using Cramer’s rule I obtain

\[
\frac{dL_M}{d\omega_M} = \frac{(1 - \alpha) Q'(\omega_E + \omega_M) + \alpha - \phi}{\phi} > 0,
\]
where the inequality follows from \( Q'(I) \geq 1 \) and \( \phi < 1 \).

Part (iv): Combining (A7) and (A8), yields \((I' - \omega_E)\left\{ (1 - \alpha) [Q(\omega_E + \omega_M^m) - \omega_E] + \alpha \omega_M^m \right\} = (I^* - \omega_E)\left\{ (1 - \alpha) [Q(\omega_E + \omega_M^m) - \omega_E] + \alpha \omega_M^m \right\}, \) with \( I' \) given by equation (15) in the main text, and hence \( \omega_M^m > \omega_M^m \). Finally, \( \omega_M^m > 0 \) follows from the assumption that \( \phi > \phi \).

Lemma A8: If (i) \( \omega_E < \omega_E^m \) and \( \omega_M < \omega_M^m \) then the entrepreneur borrows from a bank and a bank-financed moneylender. If (ii) \( \omega_E < \omega_E^m \) and \( \omega_M \in [\omega_M^m, I^* - \omega_E] \) then the entrepreneur borrows exclusively from a bank-financed moneylender. If (iii) \( \omega_E < \omega_E^m \) and \( \omega_M \geq I^* - \omega_E \) then the entrepreneur borrows from a bank and a self-financed moneylender.

Proof. I proceed by considering B’s utility given that the relevant (incentive or participation) constraint of E and M is satisfied.

Part (i): There are two distinct cases to consider when \( \omega_E < \omega_E^m \) and \( \omega_M < \omega_M^m \). First, when the incentive constraints of E and M bind, B prefers lending to both as opposed to only one of them as this minimizes the aggregate loan size needed to satisfy \( I' \) [given by equation (15) in the main text]. When M’s participation and incentive constraint hold simultaneously, B can either: (1) scale up the loan to E and M, allowing the investment to rise above \( I' \); or (2) maintain \( I = I' \) by reallocating the loan from E to M in response to an increase in M’s wealth. Suppose that Case (1) is a candidate equilibrium, as defined by equations (13) to (15) in the main text. An increase in \( \omega_M \) allows B to increase \( L_M \) up to the point at which M’s incentive constraint equals his participation constraint. M’s additional loan raises E’s return to investment and permits a larger loan to E as well. Hence, an increase in M’s debt capacity increases B’s utility by (differentiating \( U_B = Q(I) - (1 - \alpha) [Q(\omega_E + \omega_M) - \omega_E] - \alpha \omega_M - \phi (\omega_E + L_E) - L_E - L_M \) with respect to \( \omega_M \))

\[
\frac{dU_B}{d\omega_M} = \frac{Q'(\omega_E + \omega_M) [Q'(I) - (1 + \phi)] + \phi}{\phi},
\]

where \( Q'(I) < 1 + \phi \) as \( I > I' \). Meanwhile, Case (2) implies that an increase in \( \omega_M \) is met by an increase in \( L_M \) and a subsequent decrease in \( L_E \) satisfying \( dL_M/d\omega_M + d\omega_M/d\omega_M = -dL_E/d\omega_M \). Differentiating B’s utility with respect to \( \omega_M \) in this case yields

\[
\frac{dU_B}{d\omega_M} = 1 > \frac{Q'(\omega_E + \omega_M) [Q'(I) - (1 + \phi)] + \phi}{\phi}.
\]

Hence, when \( \omega_E < \omega_E^m \) and \( \omega_M < \omega_M^m \), E borrows from B and a bank-financed M with \( \omega_E + L_E + \omega_M + L_M = I' \).

Part (ii): When \( \omega_E < \omega_E^m \) and \( \omega_M \in [\omega_M^m, I^* - \omega_E] \) the only difference from Part (i) is that M’s debt capacity has improved, allowing B to extend the entire loan to M as this saves the incentive rent otherwise shared with E.
Part (iii): When $\omega_E < \omega_E^m$ and $\omega_M \geq I^* - \omega_E$, $M$ is able to self finance first-best investment and the same outcome as described in Part (ii), Lemma A6 is obtained.

**Proof of Proposition 4**

Proof. Part (i): Differentiating equations (10) and (11) in the main text with respect to $L_E$, $\bar{L}_M$ and $\omega_E$ using Cramer’s rule I obtain

$$\frac{dL_E}{d\omega_E} = \frac{\phi [Q'(I) - \phi] - (1 - \alpha) [Q'(I) - 1]}{\phi [1 + \phi - \alpha Q'(I)]} > 0$$

and

$$\frac{d\bar{L}_M}{d\omega_E} = \frac{(1 - \alpha) [Q'(I) - 1]}{\phi [1 + \phi - \alpha Q'(I)]} > 0,$$

where the inequalities follow from Lemma A2, $Q'(I) \geq 1$, and $\phi < 1$. The proof that $dL_E/d\omega_M > 0$ and $d\bar{L}_M/d\omega_M > 0$ is provided in Lemma A5.

Part (ii): All proofs are provided in Lemma A7.

Part (iii): Differentiating equations (16) and (17) in the main text with respect to $L_M$ and $\omega_E$ using Cramer’s rule I obtain

$$\frac{dL_M}{d\omega_E} = \frac{(1 - \alpha) [Q'(\omega_E + \omega_M) - 1]}{\alpha} > 0,$$

where the inequality follows from $Q'(I) \geq 1$. The proof that $dL_M/d\omega_M > 0$ is provided in Lemma A7.

**Proof of Proposition 5**

Proof. Part (i): Under competitive banking, the relevant equations are given by (A2) and (A4). Denote the critical $\omega_E$ that satisfies (A4) by $\tilde{\omega}_E^c$. Comparison yields $\omega_E^c > \tilde{\omega}_E^c > 0$, where the last inequality follows from the assumption that $\phi > \phi(\omega_M)$. Under monopoly banking, two investment levels $I^*$ [given by equation (15) in the main text] and $I^*$ need to be verified. Starting with $I'$ and combining (A3) and (A6), yields $Q(\omega_E^m) = \alpha Q(\omega_E^m + B) + (1 - \alpha) \omega_E^m - \alpha B + \phi B$. As the critical threshold $\omega_E^m$ decreases in $\alpha$, it follows from concavity that $\omega_E^m > \omega_E^m > \omega_E^m$. The proof when $I = I^*$ is analogous and omitted.

Part (ii): First, note that $\tilde{\omega}_M^c$ as defined by (A8) decreases in $\omega_E$. In particular, allow $\omega_E$ to increase up to the point at which $\phi (\omega_E + L_E) - \alpha Q(\omega_E + \omega_M) - (1 - \alpha) \omega_E + \alpha \omega_M = 0$ for $\omega_E + L_E + \omega_M = I^*$, or $\phi (I_+ - \omega_M) - \alpha Q(\omega_E + \omega_M) - (1 - \alpha) \omega_E + \alpha \omega_M = 0$. Denote the critical $\omega_M$ that satisfies this last equality by $\tilde{\omega}_M^c$. From the previous argument it follows that $\tilde{\omega}_M^c < \tilde{\omega}_M^c$. Hence, to show that $\bar{\omega}_M^c < \tilde{\omega}_M^c$, it suffices to verify that $\tilde{\omega}_M^c < \tilde{\omega}_M^c$.

Next, observe that $\tilde{\omega}_M^c$ decreases in $\alpha$. Hence, combining the expression for $\tilde{\omega}_M^c$ as defined above with the expression for $\tilde{\omega}_M^c$ as given by (A5) and allowing $\alpha \to 1$, yields $I^* [Q(I^*) - I^* - Q(\omega_E + \omega_M) + \omega_E + \omega_M] + \tilde{\omega}_M^c [Q(\omega_E + \omega_M) - \tilde{\omega}_M^c] - \tilde{\omega}_M^c [Q(1^*) - 1^* + \omega_E] = 0$. Let $[Q(I^*) - I^* - Q(\omega_E + \omega_M) + \omega_E + \omega_M] = \Phi$, where $\Phi > 0$ by concavity and $Q'(I) \geq 1$. Suppose first that $\tilde{\omega}_M^c = \tilde{\omega}_M^c$. This implies that $(I^* - \omega_E^m) \Phi = 0$. But this
equality contradicts \( I^* > \tilde{\omega}_M^m \) and \( \Phi > 0 \). Suppose then that \( \omega_M^m = \tilde{\omega}_M^c + \epsilon \). This yields 
\[
(I^* - \tilde{\omega}_M^m) > \Phi + \epsilon \left[ (Q(\omega_E + \tilde{\omega}_M^m) - \tilde{\omega}_M^m) > 0, \right.
\]
which again generates a contradiction since 
\[
Q(\omega_E + \tilde{\omega}_M^m) > \tilde{\omega}_M^m.
\]
It follows that \( \tilde{\omega}_M^m < \tilde{\omega}_M^m \), establishing the claim. ■

**Proof of Lemma 1**

[The competitive (monopoly) outcome is denoted by superscript \( c \) (m).]

**Proof.** Part (i): There are three distinct cases to consider. First, when \( \omega_M < \omega_M^m, I^c > I^m \) follows from Lemma A2, equation (15) in the main text, and concavity. Combining equations (10) and (11) and (13) and (14) in the main text, yields 
\[
Q(I^c) - I^c = \Phi I^c - Q(I^m) - D = \tilde{\omega}_M^m, \quad \text{respectively. Subtracting } L^m_E \quad \text{from } L^c_E \quad \text{using E's incentive constraints given by equations (10) and (13) yields } \alpha \left(Q(I^c) - I^c - (Q(I^m) - D)\right) = \alpha (I^c - I^m) > 0 \text{ and hence } L^c_E > L^m_E.
\]
Next, when \( \omega_M \in [\omega_M^m, I^* - \omega_E] \), the claim is trivial as monopoly bank lending to E ceases, hence \( L^c_E > L^m_E = 0 \). Third, when \( \omega_M \geq I^* - \omega_E \), the competitive outcome is obtained regardless of bank market structure, hence \( L^c_E = L^m_E \).

Part (ii): I begin by showing the existence and the uniqueness of \( \omega_M \). From Lemma A5, \( dL^m_M/d\omega_M < 0 \) when \( \omega_M \in (\omega_M^c, \omega_M^c) \). In addition, from Lemma A7, \( dL^m_M/d\omega_M > 0 \) when \( \omega_M \in (\omega_M^m, \omega_M^c) \). By continuity and Proposition 5, there exists a unique threshold \( \omega_M = \omega_M(\phi) \) for \( \omega_M \in (\omega_M^c, \omega_M^c) \) at which \( L^c_M = L^m_M \). Having established the existence and the uniqueness of \( \omega_M \), there are four distinct cases to consider. First, when \( \omega_M < \omega_M^m \), the proof is analogous to the proof of Part (i) resulting in \( L^c_M > L^m_M \). Second, suppose \( L^m_M > L^c_M \) when \( \omega_M \in [\omega_M^m, \omega_M] \). This implies that \( \omega_M < \omega_M^c \), which contradicts Proposition 5 and so \( L^c_M > L^m_M \). Third, when \( \omega_M \in [\omega_M, \omega_M] \) I have from Lemma A8 that \( L^m_M \geq L^c_M \). Fourth, when \( \omega_M \in [\omega_M^c, I^* - \omega_E] \), the claim is trivial as competitive bank lending to M ceases, hence \( L^c_M > L^m_M = 0 \). ■

**Proof of Proposition 6**

(Let \( U_E^i \) and \( U_M^i \) denote E’s and M’s respective utility.)

**Proof.** Part (i): First, from Lemma 1 I have that \( L^c_E > L^m_E \). Hence, for \( \omega_M < \omega_M^m \), \( U_E^i = \phi (\omega_E + L^c_E) > \phi (\omega_E + L^m_E) = U^m_E \) and for \( \omega_M \in [\omega_M^m, I^* - \omega_E] \), \( U_E^i = \phi (\omega_E + L^c_E) > \phi (\omega_E + L^m_E) > a Q(\omega_E + \omega_M) + (1 - a) \omega_E - a \omega_M = U^m_E \). Next, when \( \omega_M < \omega_M^c \), \( U_E^c = \phi (\omega_M + L^c_M) > \phi (\omega_M + L^m_M) = U^m_M \). When \( \omega_M \in [\omega_M^c, I^* - \omega_E] \), \( U_E^c = \omega_M < (1 - a) (Q(\omega_E + \omega_M) - \omega_E) + a \omega_M = U^m_M \).

Part (ii): Denote isolation by \( U^i_E \) and coexistence by \( U^mc_E \). For \( \omega_M < \omega_M^m \), \( U^m_i = \phi I > \phi (I - \omega_M - L_M) = U^mc_E \) [with \( I' \) given by equation (15) in the main text] and for \( \omega_M \in [\omega_M^m, I^* - \omega_E] \), \( U^m_i = \phi I' > \phi (I' - \omega_M - L_M) > a Q(\omega_E + \omega_M) + (1 - a) \omega_E - a \omega_M = U^mc_E \). ■

**Proof of Proposition 7**

**Proof.** When \( \omega_M \in (\tilde{\omega}_M, I^* - \omega_E) \), \( B^m / I^m - B^c / I^c = (B^m I^c - B^c I^m) / I^m I^c > 0 \), since \( B^m > B^c \) from Lemma 1 and \( I^c \geq I^m \). When \( \omega_M < \tilde{\omega}_M \), \( B^m / I^m - B^c / I^c \) is indeterminate, as \( B^m < B^c \) from Lemma 1, while \( I^c > I^m \). ■
Proof of Proposition 8

Proof. Differentiating the ratio of informal credit to investment, $B/I$, with respect to $\phi$, yields $i = \left[ dL_M / d\phi \times I - dI / d\phi \times (\omega_M + L_M) \right] / I^2$. Part (i): The relevant equations under competitive banking are given by (12) in the main text and the first-order condition $Q'(I) - 1 = 0$. Differentiating (12) and the first-order condition with respect to $L_M$, $I$, and $\phi$ using Cramer’s rule I obtain

$$\frac{dL_M}{d\phi} = \frac{\omega_E + L_E}{1 + \phi} > 0$$

and

$$\frac{dI}{d\phi} = 0.$$ 

Inserting $dL_M / d\phi$ and $dI / d\phi$ into $i$ yields $(\omega_E + L_E) / I (1 + \phi) > 0$. Similarly, the relevant equations with a monopoly bank are either given by (16) and (17) in the main text or by (16) in the main text and the first-order condition $Q'(I) - 1 = 0$. Starting with the former case, differentiating (16) and (17) with respect to $L_M$, $I$, and $\phi$ using Cramer’s rule I obtain

$$\frac{dL_M}{d\phi} = \frac{dI}{d\phi} = -\frac{\omega_M + L_M}{\phi} < 0.$$ 

Inserting $dL_M / d\phi$ and $dI / d\phi$ into $i$ gives $-(\omega_M + L_M) \omega_E / \phi I^2 < 0$. Differentiating (16) and the first-order condition with respect to $L_M$, $I$, and $\phi$ using Cramer’s rule I obtain

$$\frac{dL_M}{d\phi} = \frac{dI}{d\phi} = 0$$

and $i = 0$. □

Proof of Proposition 9

Proof. Part (i): It suffices to show that $R^c/B^c - 1 > 0$ for $\omega_M < \omega^c_M$, as $R^m/B^m > R^c/B^c$ (established in Part (ii) below). Hence, from (11) in the main text, I have that $R^c/B^c - 1 = [\phi(\omega_M + L^c_M) - \omega_M] / (\omega_M + L^c_M) > 0$, where the inequality follows from $\omega_M < \omega^c_M$.

Part (ii): I first demonstrate that $R^m/B^m > R^c/B^c$ and then show that $R^m/B^m$ increases under credit market segmentation. First, when $\omega_M < \omega^m_M$, let $\alpha \to 1$. This gives $R^m/B^m - R^c/B^c = (R^m B^c - R^c B^m) / B^m B^c = (D_M + \omega_M - \omega_M - L^m_M) / B^m B^c = 0$. Then let $\omega_M \to \omega^m_M$. This gives $\{B^c [Q'(I^m) - Q'(\omega_E + L^m_M)] - B^m [Q'(I^c) - Q'(\omega_E + L^c_E)]\} / B^m B^c$. Applying the mean-value theorem yields $B^m B^c [Q'(\epsilon) - Q'(\delta)] / B^m B^c$, where $Q'(\epsilon) \in (Q'(I^m), Q'(\omega_E + L^m_M))$ and $Q'(\delta) \in (Q'(I^c), Q'(\omega_E + L^c_E))$. From Lemma A2 and equation (15) in the main text, I have that $Q'(I^m) > Q'(\omega_E + L^c_E)$ and hence $R^m/B^m > R^c/B^c$. Next, when $\omega_M \in [\omega^m_M, \omega^c_M]$ and proceeding in analogous fashion by taking limits, I have again that $R^m/B^m > R^c/B^c$. Finally, when $\omega_M \in (\omega^c_M, I^*-\omega_E)$, $R^c/B^c = 1$ and the claim follows trivially.

31
I now determine how $R^m / B^m$ changes as result of segmentation. To do this, I evaluate $R^m / B^m$ at $\omega_M = \omega_M^C$ and compare with $\hat{R}^m / \hat{B}^m$ at $\omega_M = \omega_M^C - \varepsilon$. Here, $L_E = \delta, L_M = L_M (\omega_M^C) - \gamma$, where $\varepsilon, \delta,$ and $\gamma$ are small, strictly positive, and satisfy $\delta = \varepsilon + \gamma$, as investment is constant. First, let $\alpha \rightarrow 1$. This gives $R^m / B^m = \frac{\hat{R}^m}{\hat{B}^m} = \{B^m [\phi I + \varepsilon - Q (\omega_E + \omega_M)] + D_E - \delta [Q (I) - Q (\omega_E + \omega_M)]\} / B^m \hat{B}^m$, with $\phi I = \phi (\omega_E + \delta) + \phi (\omega_M + L_M) = \phi (\omega_E + \delta) + \omega_M + \omega_M = Q (\omega_E + \omega_M)$, where the inequality follows from $\omega_E < \omega_M^C$. Hence, for $\delta$ sufficiently small, $R^m / B^m > \frac{\hat{R}^m}{\hat{B}^m}$. Next, let $\alpha \rightarrow \hat{\alpha}$. This gives $\{B^m [Q (I) - Q (\omega_E + \delta)] - \delta [Q (I) - Q (\omega_E)]\} / B^m \hat{B}^m$. Again, for $\delta$ sufficiently small, $R^m / B^m > \frac{\hat{R}^m}{\hat{B}^m}$. ■

Proof of Proposition 10

Proof. Part (i): First, when $\omega_M < \omega_M^C$, $r^g = (1 - \alpha) [Q (I^C) - I^C] / B^C > 0$. Next, when $\omega_M \geq \omega_M^C$, $r^g = (1 - \alpha) [Q (I^C) - I^C] / B^C = 0$, as $\alpha = 1$.

Part (ii): Evaluating $r^g$ yields $\{[Q (I) - \alpha Q (\omega_E + \omega_M) - (1 - \alpha) \omega_E + \alpha \omega_M] L_M^m - (\omega_M + L_M^m) [Q (I) - Q (\omega_E + \omega_M)]\} / (\omega_M + L_M^m) L_M^m$. First, let $\alpha \rightarrow 1$. This gives $r^g = -\omega_M [Q (I) - Q (\omega_E + \omega_M)] L_M^m / (\omega_M + L_M^m) L_M^m < 0$, where the inequality follows from concavity and $Q' (I^C) \geq 1$. Then let $\alpha \rightarrow \hat{\alpha}$. Here $r^g = \{[Q (\omega_E + \omega_M) - Q (\omega_E)] L_M^m - \omega_M [Q (I) - Q (\omega_E + \omega_M)]\} / (\omega_M + L_M^m) L_M^m$. Applying the mean-value theorem yields $r^g = \omega_M L_M^m [Q' (\varepsilon) - Q' (\delta)] / (\omega_M + L_M^m) L_M^m$, where $Q' (\varepsilon) \in (Q' (\omega_E + \omega_M), Q' (\omega_E))$ and $Q' (\delta) \in (Q' (\omega_E + \omega_M), Q' (\omega_E + \omega_M))$. Hence, $r^g = \omega_M [Q' (\varepsilon) - Q' (\delta)] / (\omega_M + L_M^m) > 0$. By continuity there exists a threshold $\hat{\alpha} \in (\hat{\alpha}, 1)$ for $\omega_M \in [\omega_M^C, I^* - \omega_E]$ at which $r^g = 0$. ■
References


Conning, Jonathan. 2001. Mixing and Matching Loans: Credit Rationing and Spillover in a Rural Credit Market in Chile. mimeo Williams College.


