

Risk Rationing and Wealth Effects in Credit Markets

Abstract

By shrinking the available menu of loan contracts, asymmetric information can result in two types of non-price rationing in credit markets. The first is conventional quantity rationing. The second is ‘risk rationing.’ Risk rationed agents are able to borrow, but only under relatively high collateral contracts that offer them lower expected well-being than a safe, reservation rental activity. Like quantity rationed agents, credit markets do not perform well for the risk rationed. While the incidence of conventional quantity rationing is straightforward (low wealth agents who cannot meet minimum endogenous collateral requirements are quantity rationed), the incidence of risk rationing is less straightforward. Increases in financial wealth, holding productive wealth constant, counter intuitively result in the poor becoming entrepreneurs and the wealthy becoming workers. While this counterintuitive puzzle has been found in the literature on wealth effects in principal-agent models, we show that a more intuitive pattern of risk rationing results if we consider increases in productive wealth. Empirical evidence drawn from four country studies corroborates the implications of the analysis, showing that agents with low levels of productive wealth are risk rationed, and that their input and output levels mimic those of low productivity quantity rationed firms.

JEL Classification: D81, D82, O12.

The authors thank Paul Mitchell as well as seminar participants at Berkeley, Cornell, Maryland, and Wisconsin.

Stephen Boucher
Department of Agricultural and Resource Economics
University of California - Davis
One Shields Avenue
Davis, CA 95616
USA
Email: boucher@primal.ucdavis.edu

Michael R. Carter
Department of Agricultural and Applied Economics
University of Wisconsin-Madison
Madison, WI 53706
USA
Email: mrcarter@wisc.edu

Catherine Guirking
Department of Agricultural and Resource Economics
University of California - Davis
One Shields Avenue
Davis, CA 95616
USA
Email: guirking@primal.ucdavis.edu

1 Introduction

In a competitive world of symmetric information and costless enforcement, credit contracts could be written conditional on borrower behavior. Borrowers would then have access to loans under any interest rate-collateral combination that would yield lenders a zero expected profit. However, as a large literature has shown, information asymmetries and enforcement costs make such conditional contracting infeasible and restrict the set of available contracts, eliminating as incentive incompatible high interest rate, low collateral contracts.¹ This contraction of contract space can result in quantity rationing in which potential borrowers who lack the wealth to fully collateralize loans are involuntarily excluded from the credit market and thus prevented from undertaking higher return projects.

A principal contribution of this paper is to show that the contraction of contract space induced by asymmetric information can result in another form of non-price rationing, one that we label “risk rationing.” Risk rationing occurs when lenders, constrained by asymmetric information, shift so much contractual risk to the borrower that the borrower voluntarily withdraws from the credit market even when she or he has the collateral wealth needed to qualify for a loan contract.² The private and social costs of risk rationing are similar to those of more conventional quantity rationing. Like quantity-rationed individuals, risk rationed individuals will retreat to lower expected return activities and occupations.³ *From a development policy perspective, however, the implications of risk rationing are quite different from those of quantity rationing. For example, a land titling*

¹ Recent summaries of this literature include: Ghosh, Mookherjee, and Ray (2000); Udry and Conning (2005); and Dowd (1992).

² Like an interest rate increase, an increase in contractual risk will also help equilibrate the loan market by reducing demand and is thus a form of non-price rationing.

³ The insight here is consistent with the observation by Bell and Clemenz (2006) that low-wealth producers are very wary of collateral loss. They show that lenders interested in obtaining collateral through default can manipulate contract terms to frighten away safer borrowers, obtaining a riskier (and more desperate) clientele likely to face distress and forfeit their collateral assets to the lender.

program may enhance agents' ability to use their assets as collateral and thereby reduce the incidence of quantity rationing. The same policy will be ineffective in combating risk rationing because it does not enhance agents' willingness to offer collateral. The potential for risk rationing thus suggests that even in the presence of well defined and transferable property rights, the development of credit markets will be constrained by weak insurance markets.

In addition to establishing the existence of risk rationing, this paper also explores its incidence: Are higher or lower wealth agents the ones who are risk rationed? If it is the latter, then costs of asymmetric information will be borne primarily by low wealth agents who would suffer from both conventional quantity rationing as well as from risk rationing.

The incidence of conventional quantity rationing, and its impact on occupation choice, has been extensively studied with models that assume risk neutral agents. Under this assumption, Banerjee and Newman (1993), Ghatak and Jian (2002), and Aghion and Bolton (1997) all find that it is the poorest agents who are excluded from the credit market and who end up pursuing lower return occupations and activities.⁴ While these studies all further develop the implications of such non-price rationing for the evolution of the distribution of wealth, our analysis here steps back and asks what happens if we relax the assumption of risk neutrality. As is obvious, when agents are risk averse, the availability of an expected income enhancing contract is no longer sufficient for agents to choose the entrepreneurial activity. If the credit contract requires agents to bear too much risk, then they will instead prefer to voluntarily exclude themselves from the credit market and the entrepreneurial activity. The incidence of risk rationing will thus depend on the amount of risk an agent is *willing to bear* compared to the amount of risk that agent is *forced to bear* by the optimal contract appropriate for the agent's level of wealth. As this paper explores, it is

⁴ The intuition behind wealth biased quantity rationing in these papers is similar. Since the entrepreneurial project requires a fixed capital requirement, poorer agents must borrow more. Under project success, loan repayment is increasing and borrower income decreasing in loan size so that poorer agents have greater incentive to either shirk or voluntarily default. As a result, the lender may refuse to offer any contract to the poorest agents.

non-trivial to determine whether contractual risk increases faster or slower than willingness to bear risk as wealth increases.⁵

Indeed, this question about the incidence of risk rationing parallels a puzzle in the more general principal-agent literature on risk-bearing and entrepreneurship in a world in which entrepreneurial effort is unobservable and non-contractible. As analyzed by Newman (1995), and subsequently extended by Thiele and Wambach (1999), this literature asks how the wealth of a risk averse agent affects the terms on which the agent can contract to share risk with the capital market (the principal) and become an entrepreneur.⁶ Newman obtains the seemingly counter-intuitive result that under plausible assumptions about the nature of preferences, optimal contractual risk will increase so much with agent wealth, that wealthier agents will choose *not* to become entrepreneurs even when absolute risk aversion is decreasing in wealth. In contrast to the conventional Knightian theory of entrepreneurship, Newman's results imply that the poor, not the rich, will become the capitalist entrepreneurs, despite the latter's intrinsically greater capacity to bear risk.

In this paper we distinguish between purely financial wealth versus productive wealth, where the latter determines the scale of the entrepreneurial activity. We show that Newman's logic unambiguously holds only for *financial* wealth. Agents with greater financial wealth are indeed the most likely to suffer risk rationing in credit markets. While it is possible to overturn this result by manipulating the nature of preferences (for example, by making the marginal disutility of high effort decrease with wealth), we here propose an alternative approach that directly speaks to the counter-intuitive nature of Newman's result. In particular, Newman's results most likely do not

⁵ Greater agent wealth insulates the agent's consumption against bad outcomes, reducing the effectiveness of any given incentive structure and hence requiring that additional risk (and incentives) be passed to the wealthier agent to insure that the agent voluntarily provides high effort.

⁶ Mookherjee (1997) analyzes a related question, namely how the optimal effort level varies with the risk averse agent's wealth. In addition to allowing for continuous effort level, Mookherjee's analysis differs from ours in that he assumes the principal has full market power and that the principal's participation constraint will never bind so that a contract making the agent at least as well off as his reservation activity will always be available.

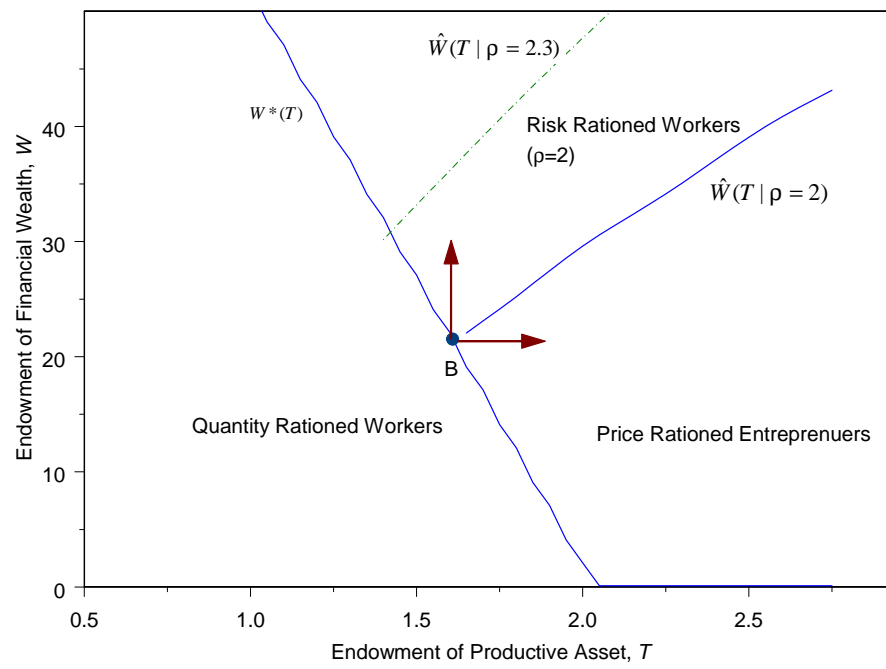
hold with respect to increases in *productive* wealth.

Figure 1 can be used to display our basic results. Financial wealth is displayed on the vertical axis. On the horizontal axis is productive wealth: land in the case of agriculture; factories in the case of industry. Imagine that an agent located at point B in the wealth space is indifferent between the entrepreneurial activity financed with a credit contract and the non-entrepreneurial, reservation activity. In our model, a move straight north from B , that it is an increase in financial wealth, will generate risk rationing of the wealthy under the empirically plausible assumption identified by Thiele and Wambach, namely that $P < 3A$, where P and A are the agent's degree of prudence and absolute risk aversion. However, under this same preference assumption, a move straight west from B can also generate risk rationing of those who are poor in productive assets. Agents to the east of B (larger landowners, or those who have pre-committed or sunk more of their wealth into factories) will not be risk rationed and will instead become the entrepreneurs. While the incidence of non-price rationing under risk aversion is more nuanced than that obtained by the literature spawned by Bannerjee and Newman (1993), we continue to find that the entrepreneurs will be located in a cone of increasing wealth. Initial wealth and activity choice continue to be tightly linked, to use the language of Eswaran and Kotwal (1990).

The remainder of this paper is organized as follows. The next section lays out a model of entrepreneurial behavior under uncertainty and describes the structure of credit contracts. Section 3 explores the implications of asymmetric information on the existence and terms of the optimal credit contract and demonstrates the potential for both quantity and risk rationing. Section 4 takes up the comparative statics of non-price rationing and wealth, examining the impact of both financial and productive wealth on the incidence of quantity and risk rationing. Section 5 presents descriptive statistical results from surveys of households and enterprises in four countries. In all four surveys, risk rationing is quantitatively important (affecting twelve to seventeen percent of

2 2₃

Figure 1: Risk rationing and activity choice



1.pdf

households) and tends to affect lower wealth households (whose economic behavior mimics that of quantity rationed agents). Section 6 concludes the paper.

2 Key Assumptions and Model Structure

Agents enjoy three endowments: financial wealth, W ; productive wealth, T ; and one unit of labor. Financial wealth is liquid and fully collateralizable. For simplicity, we assume agents earn a certain net rate of return equal to zero on financial wealth. Productive wealth, in contrast, is illiquid and cannot be used as collateral. While strong, this assumption clarifies the key incentive effects of productive wealth, as will be discussed later. We will refer to productive wealth as land and the productive/entrepreneurial activity as farming, but it could also refer to machinery or other productive infrastructure. Due to indivisibilities, if the agent chooses to farm, she must produce on her entire land endowment. Farming requires a fixed investment per unit-land, k . We restrict attention to agents with endowments such that $W < Tk$, i.e. to those who lack the capacity to self-finance production and thus must borrow in order to farm.⁷ To capture uncertainty, we assume that gross farm revenues per unit land are x_g if the state of nature is “good” and x_b if the state of nature is “bad” with $x_g > k > x_b$.

2.1 The Agent’s Preferences and Choices

An agent’s well being depends on both her end of period financial wealth (consumption) and the effort she exerts. We assume the following additively separable utility function:

$$U(I_j, e) = u(I_j) - d(e) \tag{1}$$

where I_j is the agent’s end of period financial wealth in state j and is composed of initial financial wealth plus the net income from the chosen activity and e is the effort level which can be either

⁷ An earlier version of this paper included agents that could self-finance and showed that some agents that would seek the insurance of the first-best contract instead self-financed under asymmetric information. As this is a secondary point, we exclude agents that can self-finance from the analysis.

high ($e = H$) or low ($e = L$). The disutility of effort, d , is increasing in effort so that $d(H) > d(L)$. We also assume that all agents have access to a minimum income level yielding finite utility which is exogenously guaranteed to the agent by social or other mechanisms. This assumption is required to establish quantity rationing but not risk rationing.⁸

The agent's primary decision is what to do with her land. The agent's reservation activity is to rent out the land at rental rate τ and hire out her labor at wage rate ϖ . We further assume that the reservation labor contract requires that the agent exert high effort.⁹ The agent's utility under the reservation activity is thus: $U^R = u(T\tau + \varpi + W) - d(H)$. If the agent has access to a credit contract, she must decide whether to farm her own land or undertake the reservation rental activity. If she decides to farm, she also must decide how much effort to put into farming. In addition to lowering the agent's utility, high effort raises the probability of the "good" state of nature. Let ϕ^e be the probability of the good state of nature under effort, e , so that $\phi^H > \phi^L$. We further assume that the impact of effort on profitability is sufficiently strong that under high effort farming is more profitable than the the reservation activity, while under low effort farming earns a negative rate of return. Letting \bar{x}^H and \bar{x}^L represent expected gross revenues per unit land under high and low effort and r denote the opportunity cost of lenders' funds, we make the following two assumptions about returns to effort and to the different activities:

$$\bar{x}^H - rk > \tau + \varpi/T > 0 \tag{A1}$$

$$\bar{x}^L - rk < 0 \tag{A2}$$

The implications of these two assumptions will be discussed below.

⁸ The consumption minimum prevents the lender from offering contracts which drive the agent's utility under failure towards negative infinity. If the lender could do so, then there would always exist incentive compatible contracts and quantity rationing would never occur.

⁹ This assumption – which implies that the agent's effort is the same under the reservation activity and farming with the optimal credit contract – simplifies the ensuing analytics. It is consistent with a labor contract specifying easily monitored tasks or piece rate employment in which high effort is optimally chosen by the agent.

2.2 Credit Contracts as State Contingent Payments

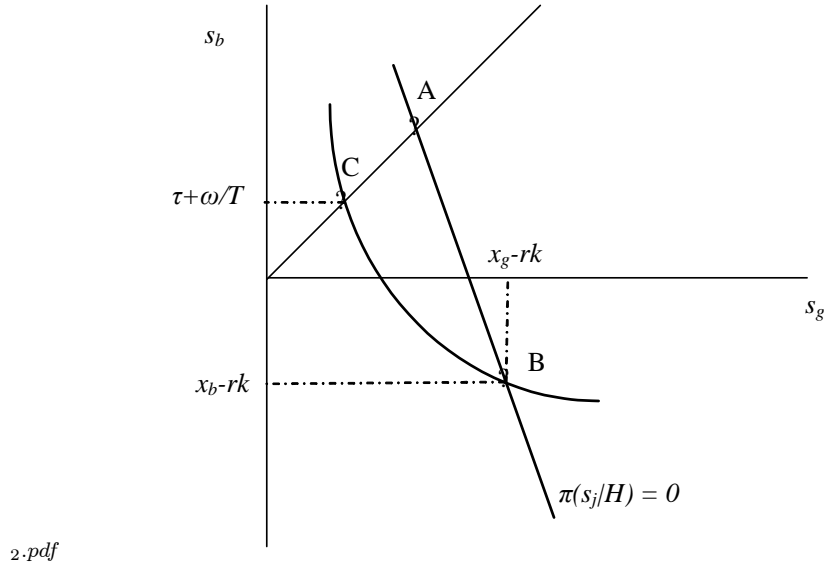
We assume the loan market is competitive and lenders are risk neutral with an opportunity cost of capital equal to r . Assumption A1 implies that under high effort the entrepreneurial activity yields a higher expected income than the safe reservation activity. Assumption A2 implies that low effort yields negative expected income under the entrepreneurial activity so that any loan contract that is offered will need to require, or induce, high effort. We also assume that if the farm project is financed, the lender provides the entire capital amount, Tk , and the farmer does not use any of her own financial wealth. A loan contract takes the form (s_g, s_b) , where s_g and s_b are the borrower's payoff per unit area financed under each state. *Thus, as in Conning (1999), the loan contract specifies how the project returns are divided between the borrower and lender under each state.* Since asymmetric information prevents lenders from specifying the borrower's effort level, lenders must choose payoffs that are incentive compatible – or that induce the borrower to choose high effort.

Much of the intuition behind the rationing results can be gleaned diagrammatically. Figure 2 portrays various potential contracts. The horizontal and vertical axes represent the borrower's payoff under good and bad states of nature respectively. The line labelled $\pi(s_j|H) = 0$ is the lender's zero expected profit contour. It gives the locus of contracts which—conditional on high borrower effort—yield zero expected profit to the lender. The slope of this contour is $-\phi^H/(1-\phi^H)$, the ratio of success to failure probabilities. Since loan contracts completely divide the farm surplus between the lender and the borrower, the contracts on this contour also yield the borrower a constant expected income – although not constant expected utility. Contracts to the northeast of this contour yield decreasing expected profits for the lender and increasing expected income for the borrower. The opposite occurs for contracts to the Southwest of this contour.

Next consider the 45-degree, or full insurance, line. Any contract on it guarantees complete

Figure 2: Borrower and lender indifference curves

2 Final



consumption smoothing for the borrower. The contract at point A , for example, is the full insurance contract yielding zero lender profits. The contract at point B , in contrast, is the full liability contract. Under contract B , the borrower bears the full farming risk while the lender's profit is certain. The scope for risk sharing via credit contracts is clear. Movements along the lender's zero profit contour from A to B represent a shifting of risk from lender and to borrower.

The risk neutral lender is indifferent to the shifting of contractual risk. In contrast the risk averse borrower is not indifferent to these movements. Holding constant high effort, the borrower's indifference curves are convex to the origin because the rate at which she is willing to trade consumption across states depends on how 'smoothed' consumption is. At a point like A or C on the full insurance line, consumption is perfectly smooth, so the borrower is willing to trade consumption across states at the same rate as the risk neutral lender $-\phi^H/(1-\phi^H)$. At a point such as B , consumption in the bad state is relatively scarce so the borrower is only willing to give up a little bit of it in order to increase consumption in the good state. Movements away from the 45-degree line along an expected income contour – such as from A to B – make the risk averse borrower worse

off.

3 Optimal Loan Contracts and the Potential for Non-price Rationing

This section examines the properties of the optimal contract in the presence of moral hazard. Given the assumption of a competitive loan market, we use a principal-agent framework in which the optimal contract maximizes the agent's expected utility subject to the principal's participation constraint and the agent's incentive compatibility constraint. *Conning (1999) uses a similar setup to model optimal contract design with monitored lending and a risk neutral borrower.* The payoffs of the optimal contract solve the following program:

$$\underset{s_g, s_b}{Max} \quad Eu(W + Ts_j | e = H) \tag{2}$$

subject to :

$$\pi(s_j | H) \equiv \phi^H(x_g - s_g) + (1 - \phi^H)(x_b - s_b) - rk \geq 0 \tag{3}$$

$$[u(W + Ts_g) - u(W + Ts_b)](\phi^H - \phi^L) \geq d(H) - d(L) \tag{4}$$

$$-s_j \leq W/T; \quad j = g, b \tag{5}$$

Equation 3 is the lender's *participation constraint* and requires that contracts, conditional on high agent effort, yield non-negative lender profits. Equation 4 is the agent's *incentive compatibility constraint* (ICC). The left hand side gives the change in the agent's expected utility while the right hand side gives the disutility cost of choosing high instead of low effort. A contract is incentive compatible if the expected utility gain outweighs the disutility cost of high effort. Finally, equation 5 gives the agent's wealth or liability constraint. Note that the agent's payoff is not restricted to be non-negative. A negative payoff requires the borrower to hand over some of her financial wealth and thus is equivalent to a collateral requirement.¹⁰

¹⁰ We do not explicitly write the agent's participation or reservation utility condition as a constraint on the maximization problem. Approaching the problem this way permits us to first characterize the contract(s) that would be made

Consider first the solution to this problem ignoring the incentive compatibility constraint. Note that if entrepreneurial effort were contractible (*i.e.*, observable and enforceable), then this constraint could be ignored. Combining the first order necessary conditions for the above maximization problem yields:

$$\frac{\frac{\partial Eu}{\partial I_g}}{\frac{\partial Eu}{\partial I_b}} = -\frac{\phi^H}{1 - \phi^H} \quad (6)$$

The above expression confirms that the first-best contract equates the marginal rates of substitution of state contingent consumption across borrower (left hand side) and lender (right hand side). Since the borrower's MRS equals $\frac{\phi^H}{1 - \phi^H} \frac{u'(I_g)}{u'(I_b)}$, the first best contract sets $s_g = s_b$ and equalizes the borrower's consumption across states. In Figure 3, the first best (contractible effort) contract would be at point *A* – exhibiting the familiar tangency condition between the borrower's indifference curve and lender's zero profit contour. In the absence of asymmetric information, credit contracts could serve the dual role of providing both liquidity and efficiently distributing risk. In this case the risk neutral lender provides full insurance to the risk averse borrower.

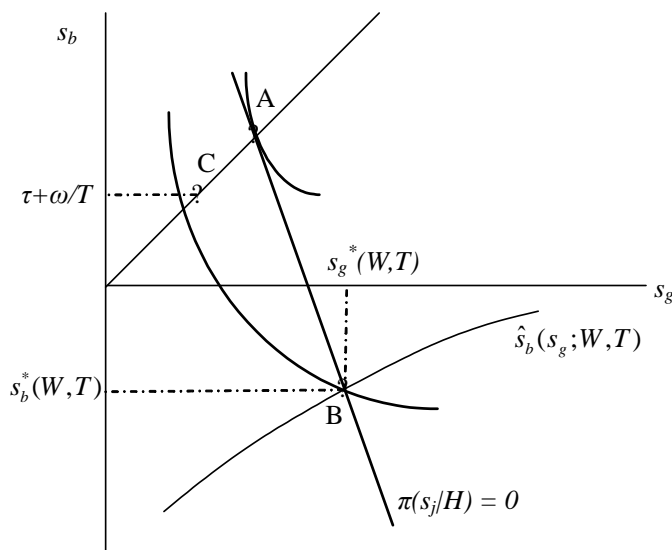
Suppose now that asymmetric information renders it impossible to enforce loan contracts written conditional on agent effort. In this case, the contract at *A* will not be available because of moral hazard. With her consumption completely shielded from farm risk, the agent would have no incentive to apply high effort. Inspection of the ICC (Equation 4) reveals that incentive compatible contracts require $s_g > s_b$. The lender motivates the borrower to apply high effort by offering contracts that reward her in the good state and punish her in the bad state.

Let $\hat{s}_b(s_g; W, T)$ – which we call the incentive compatibility boundary (ICB) – denote, for a given payoff in the good state, the payoff in the bad state such that the ICC binds. To reduce notational clutter, we will suppress the conditioning arguments W and T . Total differentiation of

available by a competitive banking system in conformity with the lender's participation constraint, the ICC and the liability constraint. We then ask whether or not any of the available contracts also fulfill the borrower's participation constraint. While appropriate for the questions at hand in this paper, analysis of other constraint configurations would be valuable for answering other questions (*e.g.*, what sort of subsidy would have to be offered lenders to induce them to offer a contract that was incentive compatible and met the borrower's liability and reservation constraints).

Figure 3: The potential for risk rationing

Final



3.pdf

the ICB yields:

$$\hat{s}'_b = \frac{u'(I_g)}{u'(I_b)} \quad (7)$$

The ICB is thus upward sloping with a slope less than unity. Concavity of the utility function implies that a \$1 increase in the payoff under the good state requires a less than proportionate increase in the payoff under the bad state. More draconian payoff combinations that lie below the ICB are incentive compatible. Those that lie above the ICB are not. Note that the ICB thus eliminates low collateral, high interest rate loans from the menu of contracts that competitive lenders will offer.

In fact, if the constrained optimal contract resulting from the optimization program defined by equations 2 - 5 exists, it will be unique and characterized by simultaneously binding ICC and LPC. That both constraints bind is intuitive. If the LPC did not bind, the lender could slightly increase s_g so that the resulting contract would continue to satisfy the ICC and make the borrower strictly better off. Similarly, if the ICC did not bind, the lender could offer a contract that marginally

increases s_b while decreasing s_g at a ratio of $\frac{\phi^H}{1-\phi^H}$. This shift, which would hold the lender's profit constant, would reduce the borrower's risk and again make her strictly better off. As illustrated in Figure 3, if it exists, the constrained optimal contract occurs at the intersection of the LPC and ICC at point B with corresponding payoffs of $(s_g^*(W, T), s_b^*(W, T))$.

The restriction of the optimal contract to the intersection of the LPC and ICB creates the potential for two sorts of non-price rationing. The first is conventional quantity rationing. *Quantity Rationing* occurs when (1) The agent would be offered and demand a credit contract in the symmetric information world; but, (2) The agent lacks sufficient wealth to collateralize the contract at the LPC-ICB intersection (*i.e.*, $W < -Ts_b^*(W, T)$). In this case, the feasible contract set will be empty and the lender will not make any contract available to the agent.

The second sort of non-price rationing that can potentially exist is what we have labelled risk rationing. *Risk Rationing* occurs when (1) The agent would be offered and demand a credit contract in the symmetric information world; (2) The agent is offered a financially feasible contract in the asymmetric information world (*i.e.*, $W \geq -Ts_b^*(W, T)$); but, (3) The agent chooses not to accept the offered contract in the asymmetric information world, preferring the safe, reservation activity.

Figure 3 can be used to depict the idea of risk rationing. Assume that the ICB is drawn for an agent with financial wealth $W > -Ts_b^*(W, T)$. The censoring of available contracts that results from asymmetric information is evident. All contracts between the full insurance line and the ICB are removed from the feasible contract set. While contracts between A and B yield higher expected utility for the borrower, the lender will not make them available as the agent has no way to commit to applying high effort. Clearly the agent would prefer to undertake the entrepreneurial activity with loan contract A to the reservation activity with its certain payoff at point C . Indeed, as can be seen the agent would accept a large number of loan contracts that lie between A and the

constrained optimal contract B . However, as drawn, the expected utility under contract B , with its sharply negative payoff in bad states of the world, is less than the expected utility associated with the reservation activity. Such an agent would rationally choose the low-returning reservation activity in preference to the entrepreneurial activity and is thus risk rationed by the definition above.

While the concept of risk rationing can thus be easily illustrated, proof of its existence, and its incidence with respect to wealth, is less straightforward and is the topic of the next section.

4 Wealth and Non-Price Rationing under Asymmetric Information

In the previous section we showed that the constrained optimal contract lies at the intersection of the LPC and ICC. The existence of the ICC – and the resulting censoring of the menu of available loan contracts – creates the potential for both quantity rationing and risk rationing. This section will show that both of these forms of non-price rationing can exist and will explore the relation between non-price rationing and both financial and productive wealth.

4.1 Quantity rationing of the poor

Feasibility of a contract for an agent endowed with financial wealth W and productive wealth T , requires that $s_b^*(W, T) > -W/T$ so that the agent has sufficient financial wealth to meet the collateral requirement. Equivalently, a sufficient condition for a positive credit supply is that the contract that requires the agent to pledge her entire financial wealth as collateral is both incentive compatible and yields non-negative lender profits. If this “full-wealth-pledge” contract cannot satisfy both of these constraints, then the feasible contract set will be empty and the agent will be quantity rationed. Proposition 1 states the conditions under which quantity rationing will occur and identifies its wealth bias.

Proposition 1 (Wealth Biased Quantity Rationing) *Assume all agents have financial wealth*

of at least \underline{W} and define $u(0)$ as the agent's utility when her state contingent payoff equals the negative of her financial wealth ($s_b^*(W, T) = -\frac{W}{T}$). Then if, for a given value of T :

$$u\left(\frac{T(\bar{x}^H - rk) + \underline{W}}{\phi^H}\right) < \frac{d(H) - d(L)}{\phi^H - \phi^L} + u(0) \quad (8)$$

then: a) There will exist a unique $W^*(T)$ such that agents with financial wealth less than $W^*(T)$ will have an empty feasible contract set and will be quantity rationed. Agents with financial wealth greater than or equal to $W^*(T)$ will have a non-empty feasible contract set. b) Holding W constant at $W^*(T)$, agents with productive wealth less than T will be quantity rationed while those with greater productive wealth will not. c) $\partial W^*(T)/\partial T < 0$, so that the minimum financial wealth required for access to a contract is decreasing in productive wealth. (Proof: See Appendix A)

While the complete proof of this proposition is detailed in the appendix, the intuition behind it can be explained. Consider whether the agent with the lowest financial wealth can qualify for a loan if she pledges her entire financial wealth, \underline{W} , as collateral. Note that under this full-wealth-pledge contract, $s_b = -\underline{W}/T$. For this value of s_b , the lender's participation constraint then defines the *maximum* payout that can be made to the borrower in the good state of the world without violating the lender's non-negative profit condition. Denote this maximum as $s_g^{\max}(\underline{W}|T)$. Similarly, the incentive compatibility constraint defines the *minimum* incentive compatible payout that can be made to the borrower in the good state of the world when $s_b = -\underline{W}/T$. Denote this minimum payout as $s_g^{\min}(\underline{W}|T)$. Payouts below this level will destroy incentives for the borrower to choose high effort.

If $s_g^{\max}(\underline{W}|T) \geq s_g^{\min}(\underline{W}|T)$, then there is at least one full wealth contract that is both incentive compatible and provides non-negative profits to the lender. However, if $s_g^{\max}(\underline{W}|T) < s_g^{\min}(\underline{W}|T)$, then the smallest payment that can be made to insure the incentive compatibility of the full wealth contract is too high and violates the lender's non-negative profit condition. In this case, the borrower will not be able to secure a loan even when pledging her full wealth as collateral. Graphically, $s_g^{\max}(\underline{W}|T) < s_g^{\min}(\underline{W}|T)$ means that the ICB cuts the Lenders Participation Constraint below $s_g^{\max}(\underline{W}|T)$. Note that since the ICB is upward sloping, less than full wealth contracts (*i.e.*, those specifying $s_b > -\underline{W}/T$) will offer a payout to the borrower in excess of $s_g^{\min}(\underline{W}|T)$. All such

contracts would offer even lower profits to the lender than the full wealth contract and will necessarily violate the non-negative profit condition. In this case, there will be no financially feasible contract that competitive lenders can offer the agent, who will by definition be quantity rationed.

As shown formally in the appendix, the full wealth contract cannot fulfill both the incentive compatibility and the lender participation constraints for the financially poorest agent when the inequality in equation 8 holds. This inequality can be rewritten as

$$u\left(Ts_g^{\max}(\underline{W}|T) + \underline{W}\right) < \frac{d(H) - d(L)}{\phi^H - \phi^L} + u(0) \quad (9)$$

and says that the full-wealth-pledge contract cannot fulfill both the zero profit and incentive compatibility constraints if the borrower's utility in the good state of the world (evaluated at $s_g^{\max}(\underline{W}|T)$), is too small to offset the opportunity cost of high effort. Note that whether or not this condition holds depends on the parameters of the problem. For example, if $u(0)$ is infinitely negative, then there will never be quantity rationing. However, as mentioned above, we assume that all agents enjoy a safety net that prevents them from suffering infinite loss in the event that they forfeit all their collateral wealth, meaning that quantity rationing is possible.

As detailed in Appendix A, if the lowest wealth agent is quantity rationed, then a large enough increase in financial wealth will always lead to the disappearance of quantity rationing.¹¹ As can be seen by inspecting the left-hand side of the inequality in equation 8, greater financial wealth will always increase $u\left(\frac{T(\bar{x}^H - rk) + \underline{W}}{\phi^H}\right)$, while it leaves the term $\frac{d(H) - d(L)}{\phi^H - \phi^L} + u(0)$ unchanged. There will thus always exist a threshold wealth level, $W^*(T)$ such that $u\left(\frac{T(\bar{x}^H - rk) + \underline{W}}{\phi^H}\right) = \frac{d(H) - d(L)}{\phi^H - \phi^L} + u(0)$.

By the same logic, any agent with financial wealth in excess of this threshold will not be quantity rationed and there will be at least one contract offered to the agent. Intuitively, this result holds because the agent's ability to offer more collateral in the bad state of the world allows the lender

¹¹ There will be some relatively wealthy agents who do not face quantity rationing as long as the following condition holds: $u\left(\frac{T(\bar{X}^H - rK) + \bar{W}}{\phi^H}\right) \geq \frac{d(H) - d(L)}{\phi^H - \phi^L} + u(0)$, where $\bar{W} = kT$ is the largest financial wealth held by any agent with productive asset level of T .

to offer a higher payoff in the good state of the world without violating the zero profit constraint. The full-wealth-pledge contract will thus be both incentive compatible and will not violate the non-negative profit condition for agents with wealth in excess of $W^*(T)$. As expected, for given T , quantity-rationing is thus biased against financially poor agents.

Less clear, however, is the direction of quantity rationing with respect to productive wealth, T . Consider a marginally quantity rationed agent who enjoys financial endowment $W^*(T)$. An increase in T dilutes the agent's available (financial) collateral per dollar borrowed (recall that production requires k units of borrowing per-unit T). The maximum payout to the borrower per-unit T that is consistent with non-negative lender profits, $s_g^{\max}(W|T)$, decreases with T , holding financial wealth fixed at $W^*(T)$.¹² This decrease would, other things equal, make it more difficult to ensure incentive compatibility, as can be seen from equation 9.

However, the marginal increase in T also creates an offsetting incentive effect as high entrepreneurial effort now yields a larger payoff as it now effects the payout on more than T units of productive capital. Indeed, as can be seen in the left-hand side of equation 8, the incentive effect always offsets the collateral reduction effect as a larger value of T unambiguously increases the returns to high efforts under the full wealth pledge contract.¹³ The increase in T has no effect on the right-hand-side of equation 8, and hence an increase in T for the marginally quantity-rationed agent will always ensure the availability of a loan contract.

Taken together, these results imply that $\partial W^*/\partial T < 0$. That is, the minimum financial wealth required to avoid quantity rationing is decreasing in productive wealth. As shown in Figure 1, the $W^*(T)$ locus is downward sloping and quantity rationing is thus biased against agents poorly

¹² As defined by the lender's non-negative profit condition, $s_g^{\max}(W|T) = \frac{\bar{X}^H - rK}{\phi^H} + \frac{1 - \phi^H}{\phi^H} \frac{W}{T}$. Note that this term is strictly decreasing in T due to the collateral dilution effect.

¹³ This result holds because $\bar{X}^H > rK$, meaning that incremental increases in project size create additional surplus beyond capital costs that can be distributed to the agent.

endowed with both financial and productive assets. Our model thus generates the same pattern of wealth biased quantity rationing that obtains under agent risk neutrality in the models of Banerjee and Newman, Aghion and Bolton, and Ghatak and Jiang.

4.2 Risk rationing and financial wealth

This section has several tasks. First, it will show that for any given level of productive assets, a sufficient increase in the drudgery of high effort will always suffice to insure that there will exist a financial wealth level, $\widehat{W}(T)$, such that the agent endowed with $\widehat{W}(T)$ is just indifferent at the optimal contract between the reservation and the entrepreneurial activities (*i.e.*, that agent is marginally risk-rationed). Assuming that high effort is sufficiently undesirable so that the marginally risk rationed agent indeed exists, this section then explores the incidence of risk rationing, asking whether it is agents with wealth greater than or less than $\widehat{W}(T)$ who will be risk rationed. This question is structurally similar to the one analyzed by Newman (1995) and especially Thiele and Wambach (1999), who examine how a risk neutral firm owner's cost of hiring a risk averse manager varies with the manager's financial wealth. Our analytical strategy for examining the wealth bias of risk rationing draws on the approach used by Thiele and Wambach. Like them, we obtain a counter-intuitive result about the impact of wealth. In our case, we find that it is the financially wealthy who will be risk rationed. Finally, this section will show the conditions under which risk rationing is economically relevant in the sense that the potentially risk rationed are not also quantity rationed.

Turning first to the existence of risk rationing, it is relatively straightforward to show that we can always find parameter values such that the marginally risk rationed agent exists. To see this, consider Figure 4, which portrays the indifference curve through the reservation activity equivalent contract for an agent of arbitrary financial wealth, W , and productive wealth, T . The point $(\tilde{s}_g, \tilde{s}_b)$ is the contract making this agent indifferent between the reservation activity and farming.

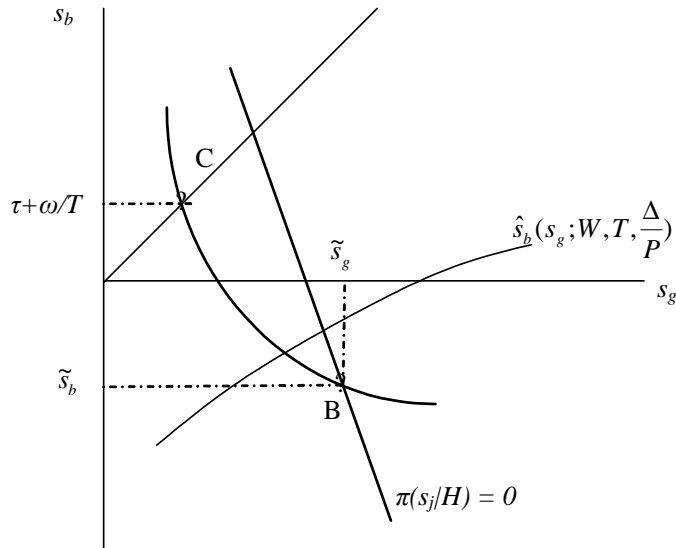
As drawn, this contract is strictly incentive compatible. It is easy to show, however, that we can “pick” parameters to convert this agent into the marginally risk rationed agent. To see this, let $\Delta \equiv d(H) - d(L)$ and $P \equiv \phi^H - \phi^L$ and explicitly write the incentive compatibility boundary, $\widehat{s}_b(s_g)$ as:

$$\widehat{s}_b = \frac{u^{-1} \left[u(W + Ts_g) - \frac{\Delta}{P} \right] - W}{T} \quad (10)$$

Since u^{-1} is an increasing function it is easy to see that by increasing or decreasing the term $\frac{\Delta}{P}$, the incentive compatibility boundary shifts down or up. Consider $\widehat{s}_b(\widetilde{s}_g)$, the maximum incentive compatible payoff under the bad state when the payoff in the good state is \widetilde{s}_g . At one extreme, if we let $\Delta = 0$, *i.e.*, we make low effort just as painful as high effort and thereby eliminate the incentive problem, then $\widehat{s}_b(\widetilde{s}_g) = \widetilde{s}_g$ so that, as to be expected, contracts on the full insurance line would be available. In contrast, if we make Δ large, we can drive $\widehat{s}_b(\widetilde{s}_g)$ to arbitrarily small (large negative) values. Since the agent’s indifference curve is independent of $\frac{\Delta}{P}$, we can always find parameter values to make any agent indifferent between her optimal contract and the reservation activity so that $\widehat{W}(T)$ will always exist. In the analysis to follow, we assume that $\underline{W} < \widehat{W}(T)$.

Figure 4: A closer look at incentive compatibility

Final



4.pdf

We turn now to the question of incidence: conditional on having access to a contract, will the financially wealthy or financially poor suffer risk rationing? At first glance, it would seem intuitive that those agents who are more sensitive to risk would be more likely to be risk rationed. Thus under decreasing absolute risk aversion (DARA), we might expect the relatively poor agents – given their greater willingness to pay for insurance – to be the first to retreat from the risk of the entrepreneurial activity. Indeed, if contract terms were exogenous to borrower wealth, this would certainly occur. Contract terms are not, however, independent of borrower wealth.

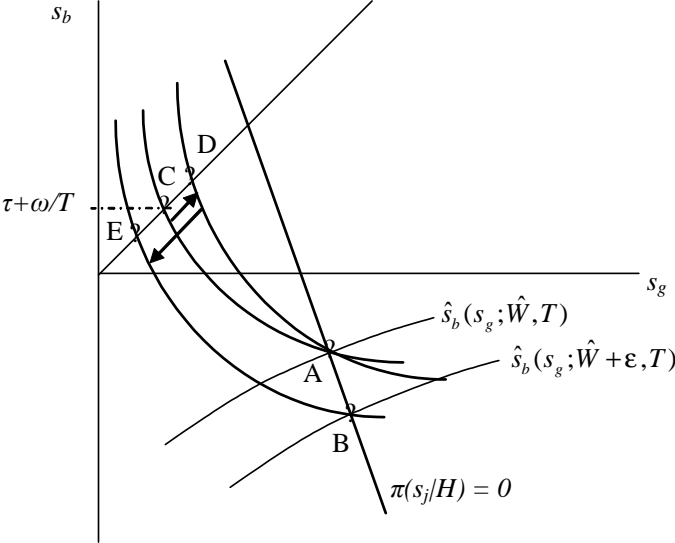
The endogeneity of contract terms to borrower wealth is easily seen by inspecting the ICC given by equation 4. Lenders make contracts incentive compatible by driving a wedge between the borrower’s payoffs, and thus consumption, across states of nature. Due to decreasing marginal utility of consumption, a constant differential in contractual payoffs, $s_g - s_b$, translates into a declining utility differential, $u(W + Ts_g) - u(W + Ts_b)$, as agent wealth increases. This implies that financially wealthier agents – who are less sensitive to a given contractual risk – must face riskier contracts than poorer agents in order to maintain incentive compatibility.

The impact of an increase in the agent’s financial wealth can be decomposed into two offsetting effects. Consider the agent with financial wealth \widehat{W} , who is indifferent between her optimal contract and the reservation activity. The *risk aversion effect* states that if we hold contract terms constant and give this marginal agent an additional dollar of financial wealth, she would strictly prefer farming with this contract. The *incentive effect*, which works in the opposite direction, implies that if the marginal agent was offered the optimal contract of a slightly wealthier agent, she would strictly prefer the reservation activity because of the additional risk required to make the wealthier person’s contract incentive compatible.

These two effects are shown in Figure 5. The marginally risk-rationed agent is indifferent between her optimal contract at A and the reservation activity at C . Note that $\tau + \varpi/T$ is the

certainty equivalent associated with the optimal contract for this agent. Under DARA, an agent's indifference curve through any contract becomes steeper as her financial wealth increases. Thus as the marginal agent is given an $\varepsilon > 0$ increase in wealth, her indifference curve through the original contract at A will cross the 45-degree line at a point like D , to the northeast of C . As the certainty equivalent of a given contract is increasing in borrower wealth, our slightly wealthier agent would strictly prefer the original contract to the reservation activity. The risk aversion effect is thus given by the increase in the certainty equivalent of the original optimal contract – represented by the move from C to D . The contract at A would not induce high effort and thus would not be available to the wealthier borrower. The increase in wealth causes the ICB to shift down, resulting in the new optimal contract at B . This is the incentive effect. As the new optimal contract is riskier, the wealthier agent's certainty equivalent falls – as represented by the move from D to E .

Figure 5: Decomposition of wealth effect on the optimal contract



As drawn in Figure 5, the incentive effect dominates the risk aversion effect so that risk rationing would affect the financially wealthiest agents who would then retreat to the low return, but certain reservation activity while poorer agents would accept the contract and undertake the risky

entrepreneurial activity. Of course we could also draw the figure such that the opposite result holds and the financially poor are risk rationed. Ultimately, the net outcome of these two effects depends on the nature of agent preferences and, more specifically, on the higher order curvature of the utility of consumption.

To explore the incidence of risk rationing with respect to financial wealth, define the utility of the marginally risk rationed agent under the reservation activity as $V^R(\widehat{W}; T)$ and the expected utility of that same agent under the entrepreneurial activity with the optimal contract as $V(\widehat{W}; T)$.¹⁴ Since $V(\widehat{W}; T) = V^R(\widehat{W}; T)$, the incidence of risk rationing will be determined by the sign of the following expression:

$$\Delta^W(\widehat{W}; T) \equiv V_W(\widehat{W}; T) - V_W^R(\widehat{W}; T) \quad (11)$$

where the W subscripts indicate derivatives taken with respect to financial wealth. If this expression is positive, then expected utility under the endogenous optimal contract exceeds that of the reservation activity as financial wealth increases and the financially poor will be risk rationed. If $\Delta^W(\widehat{W}; T) < 0$, then the financially wealthy will be risk rationed.

As shown in the appendix, with use of the envelope theorem we can write:

$$V_W = \frac{u'(W + Ts_b^*)u'(W + Ts_g^*)}{\phi^H u'(W + Ts_b^*) + (1 - \phi^H)u'(W + Ts_g^*)} \quad (12)$$

so that $\Delta^W(\widehat{W}; T)$ becomes:

$$\Delta^W(\widehat{W}; T) \equiv \frac{u'(\widehat{W} + Ts_b^*)u'(\widehat{W} + Ts_g^*)}{\phi^H u'(\widehat{W} + Ts_b^*) + (1 - \phi^H)u'(\widehat{W} + Ts_g^*)} - u'(\tau T + \varpi + \widehat{W}). \quad (13)$$

It turns out this somewhat forbidding looking expression can be signed as the following proposition details:

Proposition 2 (Risk Rationing and Financial Wealth) *Hold farm size fixed at T and assume that agent preferences are described by DARA. Let A and P denote respectively the coefficients of absolute risk aversion and prudence. If $P > 3A$ then any agent with financial wealth*

¹⁴ i.e. $V^R(\widehat{W}; T) \equiv U(\widehat{W} + Tq + w, H)$.

greater than \widehat{W} will strictly prefer the entrepreneurial activity financed with their optimal contract, while agents with financial wealth less than \widehat{W} will prefer the low return, certain reservation activity. Similarly, if $P < 3A$ then any agent with financial wealth greater than \widehat{W} will strictly prefer the reservation activity while agents with financial wealth less than \widehat{W} will prefer the entrepreneurial activity under their optimal contract. (Proof: see Appendix B)

Under proposition 2, risk rationing can thus be biased either for or against the financially wealthy. Without additional assumptions about agent preferences, however, it is not clear whether we should expect the rich or the poor to be risk rationed. In general, the relative size of P and A depends on the functional form of $u(\cdot)$ and on the level of income at which they are evaluated. We can gain some insights, however, by considering the class of constant relative risk averse (CRRA) preferences which implies a one-to-one mapping between the degree of relative risk aversion and the ratio P/A . Letting γ denote the coefficient of relative risk aversion, it is straightforward to show that $\gamma < 1/2$ is equivalent to $P > 3A$. If we believe that preferences are adequately described by CRRA preferences, we might be more inclined to expect risk rationing of the rich since most empirical studies, such as those cited in Gollier (2001), suggest that plausible values for γ lie between 1 and 4.

The existence of \widehat{W} , however, does not imply that risk rationing is economically relevant. Economically relevant risk rationing depends both upon the direction of risk rationing – as described in proposition 2 – and the relative size of the two marginal wealth levels: W^* and \widehat{W} . For a given farm size, there are four possible cases, corresponding to whether risk rationing is biased against the relatively poor ($P > 3A$) or the relatively rich ($P < 3A$) and the relative sizes of W^* and \widehat{W} . If it is biased against the rich, then risk rationing will occur independently of the relative size of W^* and \widehat{W} . In this case, if $\widehat{W} > W^*$, then all agents with financial wealth greater than \widehat{W} will be risk rationed, while if $\widehat{W} < W^*$, then only agents with financial wealth greater than W^* will be risk rationed.¹⁵ If, instead, it is biased against the poor then risk rationing will only occur if

¹⁵ Agents with financial wealth such that $\widehat{W} < W < W^*$ are ‘doubly-rationed’ as they neither have access to a

$\widehat{W} > W^*$.¹⁶ In this case, agents with intermediate wealth ($W^* < W < \widehat{W}$) are risk rationed. The following proposition summarizes these ideas and provides a sufficient condition for the existence of economically relevant risk rationing.

Proposition 3 (Economically relevant risk rationing) *Let $\overline{W}(T) \equiv Tk$ denote the maximum endowment of financial wealth for an agent with productive wealth, T . Assume that equation 8 holds so that the marginally quantity rationed agent exists within the relevant wealth spectrum. Then, if $P < 3A$, some relatively wealthy agents will always be risk rationed. If instead $P > 3A$, then some relatively poor will be risk rationed if the following equation holds:*

$$\frac{\phi^H \Delta}{P} + u(0) < u(W^* + \tau T + \varpi) < u(\overline{W} + \tau T + \varpi) < \phi^H u\left(\frac{T(\bar{x}^H - rk) + \overline{W}}{\phi^H}\right) + (1 - \phi^H)u(0) \quad (14)$$

A proof of the first part of the proposition was sketched in the discussion above. A proof for the second part of the proposition is provided in Appendix C.

In summary, returning to Figure 1, the most plausible assumptions about the nature of preferences suggest that risk rationing will occur as financial wealth increases and we move straight north from point A. The relatively poor will, however, bear the cost of quantity rationing.

4.3 Risk rationing and productive wealth

While the analytics behind risk rationing of the financially rich are clear, the result itself feels unsatisfactory. As discussed by Newman (1995), it is rather hard to accept the result that poor workers should undertake risky investment projects and hire-in the wealthy as wage workers, or that the rich rent out their land or factories to the poor. Are there ways to “overturn” this counter-intuitive result? One option is to relax the assumption of separability of effort and income in the agent’s preferences. In their labor market application, Thiele and Wambach (1999) pursue this strategy numerically and show that – for plausible coefficients of relative risk aversion – risk rationing of the poor can obtain if the disutility of effort is decreasing in income.¹⁷ In this

contract nor would they want the contract at the intersection of the ICC and ICB if it were available. We give priority to the supply-side restriction and call these agents quantity rationed.

¹⁶ If risk rationing is biased against the poor and $\widehat{W} < W^*$ then all agents with a positive supply of credit would accept their contracts and risk rationing would not occur.

¹⁷ In earlier versions of this paper, we also derived a similar result.

case, since the ‘cost’ of high effort is decreasing in income, wealthier agents need a smaller utility differential across states to maintain incentive compatibility. In terms of the language used here, an increase in the agent’s financial wealth weakens the incentive effect, making it easier for the risk aversion effect to dominate.

The results of Section 4.2 are in part counter-intuitive because they seem to create the vision of lax wealthy entrepreneurs letting their many factories (or hectares of farm land) languish under the low effort of slack management. Clearly the opportunity cost of such slack management rises with the number of factories and hectares of farmland, making it seem less likely that the wealthy really would be lax entrepreneurs. This intuition does not of course really question the results in the prior section, which were about increases in financial wealth, holding the level of productive wealth fixed. The intuition does, however, suggest that the incidence of risk rationing may be fundamentally different with respect to productive wealth. Indeed, in the case of inegalitarian agricultural economies (such as those of Latin America), questions about financial market efficiency concern land poor peasant households versus land abundant wealthy households (and are not about financially wealthy peasant households versus financially poor peasant households).

The key question addressed thus becomes the following: Under the same preferences, is it possible that the comparative statics of productive wealth yield the opposite result? Will the relatively land poor be risk rationed while the land wealthy will choose to participate in the credit market and fully exploit their productive asset (land)?

To explore this question, we proceed in a similar fashion to the previous section. We now hold financial wealth fixed and let \hat{T} denote the land size such that the agent is indifferent between the two activities:

$$V^R(\hat{T}; W) = V(\hat{T}; W), \tag{15}$$

where as before V^R is the utility of the reservation activity and V is the expected utility of the

entrepreneurial activity under the optimal loan contract. Analogous to the prior section, those poor in productive assets will be risk rationed if $\Delta^T(\hat{T}; W) \equiv V_T(\hat{T}; W) - V_T^R(\hat{T}; W) > 0$, whereas the productive asset rich will be risk rationed if the opposite sign holds. Following the same logic used in the proof of proposition 2, it can be shown that under the endogenous optimal contract:

$$V_T = (\bar{x}^H - rk) \frac{u'(W + \hat{T}s_g^*)u'(W + \hat{T}s_b^*)}{(1 - \phi^H)u'(W + \hat{T}s_g^*) + \phi^H u'(W + \hat{T}s_b^*)} \quad (16)$$

and that

$$V_T^R(\hat{T}; W) = \tau u'(\tau \hat{T} + \varpi + W). \quad (17)$$

Assembling these terms, $\Delta^T(\hat{T}; W)$ can be rewritten as:

$$\Delta^T(\hat{T}; W) \equiv \rho \frac{u'(W + \hat{T}s_g^*)u'(W + \hat{T}s_b^*)}{(1 - \phi^H)u'(W + \hat{T}s_g^*) + \phi^H u'(W + \hat{T}s_b^*)} - u'(\tau \hat{T} + \varpi + W). \quad (18)$$

where $\rho = \frac{\bar{x}^H - rk}{\tau}$ is the ratio of the expected marginal returns to the productive asset when used entrepreneurially relative to the returns when the asset is rented out to others. As before, $\Delta^T(\hat{T}; W) > 0$ will imply risk rationing of the poor, while the opposite sign will imply that those rich in productive assets will rent out their assets and become workers, while the productive asset poor will become the entrepreneurs.

Note that this expression is identical to equation 13 except that first term (which captures marginal expected utility returns to the entrepreneurial activity) is multiplied by ρ . Under assumption (A1), $\rho > 1$ and represents precisely the increased incentives for the productive asset rich to use their assets entrepreneurially rather than renting them out. This additional term makes it more likely that $\Delta^T(\hat{T}; W)$ from equation (18) is positive and that the land poor will be risk rationed. Consistent with the intuition discussed above, the larger are the relative returns to high effort entrepreneurialism, the more likely that the productive asset rich will have adequate incentives to supply high effort and the less need for high risk, draconian credit terms to induce high effort. However, without imposing additional structure on preferences, we cannot derive a neat analytic

condition equivalent to the $P > (<)3A$ conditions of Proposition 2 that are necessary and sufficient for risk rationing of the land rich or the land poor. From equation 18, however, we do know that the more empirically plausible condition $-P < 3A$ is necessary but not sufficient for risk rationing of the land rich, while the less plausible condition $-P > 3A$ is sufficient but no longer necessary for risk rationing of the land poor. Thus under the empirically more plausible assumption that $P < 3A$, both types of risk rationing with respect to productive wealth could occur.

4.4 Numerical analysis of the incidence of non-price rationing

The analysis in the prior sections has identified conditions under which risk and quantity rationing will exist. The incidence of quantity rationing with respect to financial and productive wealth is also clearly identified analytically. Under reasonable assumptions about the nature of preferences, the financially wealthy will be risk rationed. However, the incidence of risk rationing with respect to productive wealth depends on the more subtle interplay of a number of parameters and no simple analytical expression exists that indicates whether it is the productive asset poor or rich who will be risk rationed.

To gain better purchase on the incidence or risk rationing, and its interaction with quantity rationing, Figure 1 reports the results of a numerical analysis, mapping out the regions of the endowment space that are characterized by the different types of price and non-price rationing. Appendix D reports the full set of parameter values used for the simulation.¹⁸ We assume that agents have constant relative risk aversion and for the analysis, we set the coefficient of relative risk aversion to 1.1 (meaning that $P < 3A$). Note that this value falls into the empirically plausible range and is also in the range identified in Proposition 2 under which the financially wealthy will be risk rationed. Other parameters are set such that both risk and quantity rationing occur over the

¹⁸ Consistent with note 6 above, parameters have been chosen such that all agents (except those in the northwest corner of Figure 1) lack the financial wealth to self-finance the production process.

illustrated portion of the endowment space (recall from the earlier analysis that high effort must be sufficiently onerous in order for non-price rationing to exist).

The solid lines in Figure 1 represent the non-price rationing boundaries for the case when the critical parameter ρ is set equal to 2, meaning that expected marginal returns to productive wealth are twice as high when operated entrepreneurially than when rented out. The downward sloping line is the quantity rationing locus, $W^*(T)$. There is no contract available to agents with endowment locations below that line, and hence that portion of the endowment space is characterized by quantity rationing. Above that line, competitive loan contracts are available. Under this numerical specification, the incentive effects of large endowments of productive assets are strong enough that there is no quantity rationing of agents who have at least 1.65 units of productive wealth.

The upward sloping solid line marks the risk rationing boundary, $\widehat{W}(T)$. Agents above that line are risk rationed, preferring the reservation activity to the risky entrepreneurial activity financed by the optimal contract. Agents below that line will be price rationed and accept the optimal contract and undertake the entrepreneurial activity. As can be seen, the risk rationing boundary is upward sloping. Starting with a marginally risk-rationed agent at a point like A , this slope indicates that, for this parameter set, agents with greater productive wealth will become the entrepreneurs. With $\rho = 2$, the incentive effect of having an additional unit of productive wealth is strong enough that the optimal contract becomes less onerous for agents with additional productive wealth. Agents with less productive wealth than A will undertake the reservation activity, renting out their resources and becoming wage workers.

As can be seen in Figure 1, the quantity rationing and risk rationing boundaries create a cone of agents who will become the entrepreneurs. The cone widens as distance from the origin increases, indicating that it is indeed the wealthy who become the entrepreneurs. As can also be seen, if ρ

increases to 2.5, strengthening incentive effects for taking high entrepreneurial effort, then the risk rationing boundary shifts up and the entrepreneurial cone widens further as shown by the broken line in Figure 1. While the positioning of these boundaries is of course an artifact of the specific numerical values chosen, they do illustrate the forces that promote wealth-biased risk rationing identified analytically in the prior section.

5 The Frequency and Incidence of Risk Rationing: Empirical Evidence

Risk rationing, like quantity rationing, prevents individuals from efficiently exploiting their endowments of productive assets. Any empirical evaluation of the performance of financial markets should thus account for both forms non-price rationing. If a non-trivial fraction of economic agents is risk rationed, then failure to account for risk rationing would lead to an understatement of the frequency of non-price rationing and an overly optimistic evaluation of the health of the financial system. Econometric analyses of the determinants and impacts of credit rationing should also incorporate risk rationing. This empirical literature has struggled with the fundamental problem of distinguishing individuals with zero loan demand from quantity rationed individuals. While the econometrics of unobserved regime switching offer one approach to this problem (*e.g.*, (?), (?)), an alternative approach is to obtain direct indicators of positive loan demand via constraint elicitation questions that inquire whether firms without loans applied for them, and if not, why not (*e.g.*, (?), (?), (?)). Both econometric approaches may generate biased estimates if risk rationing is important empirically but not accounted for.

Table 1 reports data on risk-rationing from four enterprise surveys, one of rural farm and non-farm enterprises in Guatemala, and the other three of agricultural enterprises in Peru, Honduras and Nicaragua.¹⁹ Each survey was designed to directly elicit the firm's (or farm's) credit rationing

¹⁹ (?) describe the Guatemala survey, while (?) does the same for the Peru survey and (?) provide an overview

status, including the potential for risk rationing. Firms reported as price rationed in the table include both firms that borrowed and those that chose not to because they did not need capital or found the cost of capital to be too high. Non-price rationed firms are those that indicated that they would have liked to borrow money at the going rate of interest, but that they either could not qualify for a loan (*i.e.*, were quantity rationed),²⁰ or were afraid to take one because of the risk of collateral loss (risk rationed). As can be seen, risk rationed enterprises constitute between 12% and 17% of all surveyed enterprises, and between 20% and 40% of all non-price rationed firms. Failure to account for risk rationed firms as non-price rationed would clearly have a major effect on the analysis of the efficiency of credit markets under asymmetric information.

Table 1. Risk and Quantity Rationed Firms

	Peru			Guatemala		
	<i>Non-Price Rationed</i> Quantity	<i>Risk</i>	<i>Price Rationed</i>	<i>Non-Price Rationed</i> Quantity	<i>Risk</i>	<i>Price Rationed</i>
% of Sample	36.7	17.2	46.1	31.1	13.7	55.2
Productive Wealth (\$)	13,336*	9,396*	23,771	21,510*	6,024*	38,972
Input (\$/ha)	451*	454*	868	NA	NA	NA
Income (\$/ha)	653*	593*	919	NA	NA	NA
	Honduras			Nicaragua		
	<i>Non-Price Rationed</i> Quantity	<i>Risk</i>	<i>Price Rationed</i>	<i>Non-Price Rationed</i> Quantity	<i>Risk</i>	<i>Price Rationed</i>
% of Sample	22.8	15.6	61.6	48.0	12.3	39.7
Productive Wealth (\$)	10,523*	11,916*	26,508	4,340*	6,283*	12,299
Input (\$/ha)	128	127	167	19*	15*	31
Income (\$/ha)	81	98	125	71	71	76

*Indicates the mean of the non-price rationed group is different from the price rationed group at the 5% significance level.

Table 1 also displays some additional information on risk-rationed versus other types of firms. We can glean a meaningful idea of the activity choice of risk rationed producers by looking at their use of inputs as well as net-income produced per-unit land in the three farm-based surveys.²¹ As of the Honduras and Nicaragua studies.

²⁰ Most of these firms are what Mushinski (1999) calls pre-emptively rationed as they do not bother to apply for loans, knowing full well that they will not receive them.

²¹ Farmers in the Peru sample are relatively homogeneous in that they all cultivate annual crops on irrigated land. The overall Nicaraguan and Honduran samples are more heterogeneous. In order to have a more meaningful comparison, we exclude households that produced perennial crops such as coffee from these two samples.

can be seen, the risk rationed farm households appear similar to the quantity rationed. The value of variable inputs per hectare used by quantity and risk rationed households ranges from 20% less than price rationed households in Honduras to 50% less in Peru. Net income per hectare is also less for both categories of non-price rationed households than for price rationed households.

The final piece of information displayed in Table 1 is the value of productive assets—including land, structures, and machinery—for each type of household. A similar pattern is evident. In each of the countries, the mean wealth holdings of both risk-rationed and quantity rationed producers are significantly below the mean for price rationed firms, providing at least tentative evidence that – consistent with the model specification underlying Figure 1 – risk rationing more affects the productive asset poor than the wealthy.²²

6 The Economics of Risk Rationing: Wealth, Optimal Contracts and Activity Choice

The theoretical model developed in this paper has shown that by shrinking the available menu of loan contracts, asymmetric information can result in two sorts of wealth-biased, non-price rationing in credit markets. The first is conventional quantity rationing in which a subset of low-wealth agents find that there is no contract that is made available to them because they lack the minimum collateral necessary to secure a loan. The second is what this paper has labelled risk rationing. Risk rationed agents are able to borrow, but only under relatively high collateral contracts that offer them lower expected well-being than does a safe, reservation rental activity. This latter effect is particularly relevant in developing countries where insurance markets are scarce and risk averse agents may seek credit contracts both to overcome liquidity constraints and to obtain insurance against production or price shocks. But when faced with the offer of only a high collateral contract

²² Unfortunately, many households were reluctant to provide information on their holdings of financial wealth so that we cannot provide evidence on the incidence of risk rationing with respect to financial wealth.

that places their asset base at risk, risk rationed agents choose a safer, lower return activity than they would choose in a symmetric information world. Like quantity rationed agents, the risk rationed are a class for whom decentralized credit markets do not perform well.²³

While the incidence of conventional quantity rationing is straightforward, the incidence of risk rationing is less straightforward. One contribution of this paper has been to show that its incidence depends on the type of wealth. In particular this paper shows that Newman's (1995) counter-intuitive finding that the poor not the rich will be the entrepreneurs is true only for financial wealth. The opposite is likely to be the case for agents who enjoy large endowments of productive wealth. While analysis of the wealth portfolio decision is beyond the scope of this paper, the analysis does suggest that the wealthy can render themselves more creditworthy by sinking or precommitting more of their wealth to productive assets. Doing so improves their creditworthiness not by collateralizing their loans (indeed, under the assumptions made here, a shift in wealth composition toward productive assets actually reduces the collateral value of an agent's wealth), but by strengthening the intrinsic incentives for the entrepreneur to choose high effort. Note, however, that this portfolio adjustment strategy will not work for the relatively poor. At least under the parameter set used for the numerical analysis, the less wealthy will be either risk or quantity rationed irrespective of the composition of their portfolio.

The analysis here thus at least opens the door to the possibility that contract truncation created

²³ >From a theoretical perspective, this paper's analysis of optimal credit contracts under risk aversion and asymmetric information suggests several extensions. First, the model could be extended to incorporate the various means by which borrowers and lenders overcome information asymmetries. For example, under monitored lending, the agent's effort level is monitored—either by the lender or by other agents in a group lending scheme—and a penalty is imposed if the agent deviates from the agreed upon effort level. Conning (1996, 1999), for example, has taken initial strides along this line by developing a model which endogenizes the level of monitoring and institutional form under moral hazard for individual credit contracts. Besley and Coate (1995) and Armendáriz de Aghion (1999) develop models of endogenous monitoring under group lending. Extending the model in this direction could help explain the frequently observed coexistence of multiple institutional forms of credit delivery. A second and related theoretical extension would be to reconsider the role and logic of informal or local lenders who are less subject to information asymmetries. These lenders may be able to offer contracts with greater implicit insurance than formal sector contracts. This is consistent with empirical observation that informal lenders rarely require collateral. Even if informal contracts are more expensive in terms of the expected value of loan repayment, agents may prefer them for their implicit insurance.

by asymmetric information weighs most heavily on the less well-off households. As in Eswaran and Kotwal (1990), initial wealth and activity choice become tightly wedded in the analysis here. Fully endogenous quantity and risk rationing lead the initially poor to choose the safe reservation activity. Those with more favorable initial wealth endowments become entrepreneurs. The expected rate of return on wealth and labor resources will thus be positively related to wealth (as in Bardhan, Bowles and Gintis, 1998), and in the face of moral hazard-constrained credit markets, both class structure and income inequality will tend to reproduce themselves over time. Whether loan markets in fact have this severe form of wealth-bias is ultimately an empirical question. But irrespective of the direction of risk rationing, this paper's analysis suggests that empirical studies that fail to take risk rationed agents into account will overestimate the health of the financial system. *Similarly, policy efforts that ignore the role of contractual risk as a deterrent to the demand for loans will likely lead to disappointing results in terms of credit market development.*

Appendix A. Proof of Proposition 1

The proof consists of three steps. First, we describe $W^*(T)$ and show that it is the minimum collateral requirement necessary for a non-empty feasible set so that quantity rationing is biased against the financially poor. Second, we show that increases in productive wealth have the same qualitative effect – namely quantity rationing, if it occurs, will affect the productive wealth poor. Finally, we show that Equation 8 is necessary and sufficient for the existence of the marginally quantity rationed agent, $W^*(T)$, within the relevant range of financial wealth.

Define the following payoffs to the agent in the good state:

$$s_g^{min}(W; T) : [u(W + Ts_g^{min}) - u(0)](\phi^H - \phi^L) = d(H) - d(L) \quad (19a)$$

$$s_g^{max}(W) : \phi^H(x_g - s_g^{max}) + (1 - \phi^H)(x_b + W/T) = rk \quad (19b)$$

s_g^{min} is minimum incentive compatible payoff in the good state when the agent posts her full financial wealth as collateral (i.e., when $s_b = -W/T$). Similarly, s_g^{max} is the payoff to the agent in the good state such that the lender just breaks even when the agent again posts maximum collateral. Holding s_b at $-W/T$, any contract with $s_g < s_g^{min}(W; T)$ would violate the ICC while any contract with $s_g > s_g^{max}(W; T)$ would violate the LPC. Thus feasible, full collateral contracts require $s_g^{min}(W; T) \leq s_g \leq s_g^{max}(W; T)$. Now let $W^*(T)$ be the financial wealth level such that the LPC, ICC, and the agent's wealth constraint all bind and consider a marginal increase in W . From equation 19a, $\frac{\partial s_g^{min}}{\partial W} = -1/T < 0$ and from Equation 19b, $\frac{\partial s_g^{max}}{\partial W} = \frac{1 - \phi^H}{\phi^H} \frac{1}{T} > 0$, so that if $W^*(T)$ exists, then any agent with productive wealth T and financial wealth $W < W^*(T)$ will have an empty feasible contract set and will be quantity rationed. Agents with $W > W^*(T)$ will have access to some contracts and will not be quantity rationed.

Now return to our marginally quantity rationed agent, $W = W^*(T)$ and consider an increase in productive wealth, T . Again, using equations 19a and 19b we have: $\frac{\partial s_g^{min}}{\partial T} = -s_g^{min}/T < 0$ and

$\frac{\partial s_g^{\max}}{\partial T} = - \left(\frac{1-\phi^H}{\phi^H} \right) \frac{W}{T^2} < 0$. Since both the lower and upper bounds of the success payoff decrease, an increase in productive wealth will imply a non-empty feasible contract iff $\left| \frac{\partial s_g^{\min}}{\partial T} \right| > \left| \frac{\partial s_g^{\max}}{\partial T} \right|$ or, equivalently, $s_g^{\min} > \left(\frac{1-\phi^H}{\phi^H} \right) \frac{W}{T}$. Next, rewrite equation 19b as:

$$s_g^{\max}(W^*(T); T) = \frac{\bar{x} - rk}{\phi^H} + \frac{1 - \phi^H}{\phi^H} \frac{W^*}{T}. \quad (20)$$

Since, by definition, the agent with $W^*(T)$ has a single contract available, we know that $s_g^{\min}(W^*(T); T) = s_g^{\max}(W^*(T); T)$ so that:

$$s_g^{\min}(W^*(T); T) = s_g^{\max}(W^*(T); T) > \left(\frac{1 - \phi^H}{\phi^H} \right) \frac{W}{T}$$

which is necessary and sufficient for agents with greater productive wealth to have a non-empty feasible contract set while agents with less productive wealth will be quantity rationed.

Finally, we take up the existence of W^* . From the above argument, it is clear that if the poorest agent is *not* quantity rationed, then no agent will be quantity rationed. Similarly, if the wealthiest agent *is* quantity rationed, then all agents will be quantity rationed. To demonstrate the existence (and uniqueness) of W^* we need to find a condition such that the poorest agent is quantity rationed and the wealthiest is not. Given the discussion above, this condition holds if and only if the following two inequalities hold:

$$s_g^{\min}(W; T) > s_g^{\max}(W; T) \quad (21)$$

$$s_g^{\min}(\bar{W}; T) < s_g^{\max}(\bar{W}; T) \quad (22)$$

Using the definitions of s_g^{\min} and s_g^{\max} above, it is easy to show that these two inequalities are equivalent to:

$$u \left(\frac{T(\bar{x}^H - rk) + W}{\phi^H} \right) < \frac{d(H) - d(L)}{\phi^H - \phi^L} + u(0) < u \left(\frac{T(\bar{x}^H - rk) + \bar{W}}{\phi^H} \right)$$

which is the necessary and sufficient condition in the proposition.

Appendix B. Proof of Proposition 2

To prove Proposition 2, we need to show that $P > 3A$ is necessary and sufficient for the certainty equivalent of the optimal contract to be increasing in the agent's financial wealth. Hold productive wealth constant and, for notational simplicity, set $T = 1$. Let $V(W)$ be the agent's expected utility from farming with her optimal contract. The agent's certainty equivalent of the optimal contract, $C(W)$, is implicitly defined by the following equation:

$$U(W + C(W), H) = V(W) \quad (23)$$

Totally differentiating and rearranging equation 23 yields:

$$\frac{\partial C(W)}{\partial W} = \frac{V'(W) - u'(W + C(W))}{u'(W + C(W))} \quad (24)$$

In what follows, we will show that $P \geq 3A$ is necessary and sufficient for the shadow value of financial wealth to be greater in farming than in the reservation activity ($V' > u'$) so that the numerator on the right hand side of equation 24 is positive.

The Lagrangian of the formal sector optimization problem is:

$$\begin{aligned} \mathcal{L}(W, \lambda, \mu) = & EU(W + s_j, H) - \lambda \{d(H) - d(L) - [u(W + s_g) - u(W + s_b)](\phi^H - \phi^L)\} \\ & - \mu[-\phi^H(X_g - s_g) - (1 - \phi^H)(X_b - s_b) + rk] \end{aligned} \quad (25)$$

where λ and μ are the multipliers associated with the incentive compatibility and participation constraints. Applying the envelope theorem yields:

$$V'(W) = \phi^H u'(W + s_g^*) + (1 - \phi^H) u'(W + s_b^*) + \lambda^* (\phi^H - \phi^L) [u'(W + s_g^*) - u'(W + s_b^*)] \quad (26)$$

Both the lender's participation and the incentive compatibility constraints are binding at the optimum so that λ^* and μ^* are strictly positive and the first order necessary conditions for an

optimum are:

$$\frac{\partial \mathcal{L}}{\partial S_g} = \phi^H u'(W + s_g^*) + \lambda^*(\phi^H - \phi^L)u'(W + s_g^*) - \mu^* \phi^H = 0 \quad (27a)$$

$$\frac{\partial \mathcal{L}}{\partial S_b} = (1 - \phi^H)u'(W + s_b^*) - \lambda^*(\phi^H - \phi^L)u'(W + s_b^*) - \mu^*(1 - \phi^H) = 0 \quad (27b)$$

Solving equations 27a and 27b for λ^* yields:

$$\lambda^* = \frac{\phi^H(1 - \phi^H)[u'(W + s_b^*) - u'(W + s_g^*)]}{(\phi^H - \phi^L)[\phi^H u'(W + s_b^*) + (1 - \phi^H)u'(W + s_g^*)]} \quad (28)$$

Substituting for λ^* in equation 26 and simplifying yields:

$$V'(W) = \frac{u'(W + s_b^*)u'(W + s_g^*)}{\phi^H u'(W + s_b^*) + (1 - \phi^H)u'(W + s_g^*)}. \quad (29)$$

Thus $V'(W) > u'(W + C(W))$ is equivalent to:

$$\frac{u'(W + s_b^*)u'(W + s_g^*)}{\phi^H u'(W + s_b^*) + (1 - \phi^H)u'(W + s_g^*)} > u'(W + C(W)) \quad (30)$$

Next, assume that the utility function $\frac{1}{u'(\cdot)}$ exhibits greater absolute risk aversion than $u(\cdot)$. By definition of the certainty equivalent:

$$u(W + C(W)) = \phi^H u(W + s_g^*) + (1 - \phi^H)u(W + s_b^*) \quad (31)$$

Note that the certainty equivalent of the marginally risk rationed agent (with financial wealth \widehat{W}) is $q + w$ so that equation 30 becomes equation 13 in the text. If presented with the same contract $-(s_g^*(W), s_b^*(W))$ - an agent with the same wealth, but with utility function $\frac{1}{u'(\cdot)}$ would strictly prefer the certainty equivalent:

$$\frac{1}{u'(W + C(W))} > \frac{\phi^H}{u'(W + s_g^*)} + \frac{1 - \phi^H}{u'(W + s_b^*)} \quad (32)$$

Inverting both sides of this inequality yields the inequality in equation 30.

The final step is to demonstrate that $P > 3A$ is equivalent to an agent with utility function $\frac{1}{u'(\cdot)}$ being more risk averse than an agent with utility u . Using the definition of the coefficient of

absolute risk aversion, $\frac{1}{u'(\cdot)}$ is more risk averse than u if and only if:

$$-\frac{(\frac{1}{u'})''}{(\frac{1}{u'})'} > -\frac{u''}{u'} \leftrightarrow P > 3A \quad (33)$$

Following similar steps it can be shown that an agent with utility function $\frac{1}{u'(\cdot)}$ is *less* risk averse than an agent with utility u if and only if $P < 3A$.

To summarize, we have shown that $P > 3A$ implies that the certainty equivalent of the optimal contract is increasing in agent wealth. Consequently, if an indifferent agent exists, any poorer agent would strictly prefer the certain reservation activity to her optimal formal contract so that - under this preference condition - risk rationing is biased against the poor. *Note that since both the value function, V , and the agent's reservation utility are monotonically increasing in W , this result is a global result so that the indifferent agent, if she exists, is unique. To see this, assume that two indifferent agents - and thus two crossing points - exist. Monotonicity implies that the relative magnitudes of the slopes of the two functions would be inverted at the two crossing points. This cannot occur, however, since we have just seen that at a crossing point the value function is steeper than reservation utility.* A symmetric proof can be used to show that $P < 3A$ implies that risk rationing is biased against richer agents.

Appendix C. Proof of Proposition 3.

Here we show that equation 14 is sufficient for the existence of $\widehat{W}(T) \in [W^*(T), \overline{W}(T)]$ when risk rationing is biased against the financially poor ($P > 3A$). To show this, we need to find conditions such that the marginally quantity rationed agent, W^* , *does not* want the single contract available to her while the wealthiest agent, \overline{W} , *does* want her optimal contract. First, W^* will not want her contract if:

$$u(W^* + \tau T + \varpi) > \phi^H u(W^* + Ts_g^*(W^*)) + (1 - \phi^H)u(W^* + Ts_b^*(W^*)) \quad (34)$$

By definition of the marginally quantity rationed agent, $s_b^*(W^*) = -W^*/T$; and $s_g^*(W^*) = s_g^{\min}(W^*)$, where $s_g^{\min}(W^*)$ is the minimum incentive compatible payoff in the good state under the full-wealth-pledge contract collateral ($s_b = -W/T$) as defined in Appendix B. The contract $(s_g^{\min}(W^*), -W^*/T)$ satisfies the incentive compatibility constraint with equality for agent W^* so that:

$$u(W^* + Ts_g^*(W^*)) = \frac{\Delta}{P} + u(0) \quad (35)$$

And we can rewrite equation 34 as:

$$u(W^* + \tau T + \varpi) > \frac{\phi^H \Delta}{P} + u(0). \quad (36)$$

Second, \bar{W} , will want her optimal contract if:

$$u(\bar{W} + \tau T + \varpi) < \phi^H u(\bar{W} + Ts_g^*(\bar{W})) + (1 - \phi^H)u(\bar{W} + Ts_b^*(\bar{W})). \quad (37)$$

As it implies lower risk, \bar{W} strictly prefers her optimal contract to the full wealth-pledge contract $(s_g^{\max}(\bar{W}), -\bar{W}/T)$. Using the lender's zero profit condition it is easy to show that this contract yields expected utility: $\phi^H u\left(\frac{T(\bar{x}^H - rk) + \bar{W}}{\phi^H}\right) + (1 - \phi^H)u(0)$. Thus a sufficient condition for \bar{W} to prefer her optimal contract to the reservation activity is:

$$u(\bar{W} + \tau T + \varpi) < \phi^H u\left(\frac{T(\bar{x}^H - rk) + \bar{W}}{\phi^H}\right) + (1 - \phi^H)u(0). \quad (38)$$

Combining equations 36 and 38 yields:

$$\frac{\phi^H \Delta}{P} + u(0) < u(W^* + \tau T + \varpi) < u(\bar{W} + \tau T + \varpi) < \phi^H u\left(\frac{T(\bar{x}^H - rk) + \bar{W}}{\phi^H}\right) + (1 - \phi^H)u(0) \quad (39)$$

which is equation 14 in Proposition 3.

Appendix D. Numerical Analysis

Utility:

$$u(I_j, e) = \begin{cases} \left(\frac{1}{1-\gamma}\right) (I_j + c_0)^{(1-\gamma)} - D, & \text{if } e = H \\ \left(\frac{1}{1-\gamma}\right) (I_j + c_0)^{(1-\gamma)}, & \text{otherwise} \end{cases}$$

where $\gamma = 1.1$, and $D = 1.1$ and $c_0 = 10$

Entrepreneurial Activity

Gross entrepreneurial incomes: $x_g = 100$; $x_b = 0$

Success Probabilities: $\phi^H = 80\%$; $\phi^L = 20\%$

Capital investment requirement: $k = 25$

Interest rate: $r = 20\%$

Expected net entrepreneurial income under high effort: $\bar{x}^H - rk = 62$;

Safe Wage/Rental Activity

Certain Wage income $\varpi = 5$;

Rent per-unit productive asset $\tau = 25$ for case where $\rho = 2$;

$\tau = 21.75$ for case where $\rho = 2.3$.

References