UNIVERSITÄT BERN

Faculty of Economics and Social Sciences

Department of Economics

Finance and Competition

Harris Dellas
Ana Fernandes

07-03

February 2007

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach 8573
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Finance and Competition*

Harris Dellas[†] Ana Fernandes[‡]

Department of Economics
University of Bern

February 3, 2007

Abstract

Financial constraints are often thought as representing a barrier to entry for new firms, thus potentially limiting competition in product markets. We investigate the relationship between finance and product market competition in the context of a general equilibrium, two-sector model. The analysis highlights the role played by firm heterogeneity as well as by the level and distribution of wealth. Financial development may lead to lower markups (and thus to more competitive markets) in financially dependent sectors, even when it reduces the number of firms and increases standard market concentration indexes. The analysis implies that incumbency is not a sufficient condition for determining opposition to financial liberalization. It also implies that, for a given level of imperfect financial development, poorer countries will tend to have less competitive product markets.

JEL Codes: L1, E2.

Keywords: financial development, liberalization, market structure, product market competition.

*Preliminary. We thank Luís Cabral, Jeff Campbell, Winand Emons, Bob Hunt, Narayana Kocherlakota, Bob Lucas, Gerd Muehlheusser, Ed Prescott and Sérgio Rebelo for helpful comments.

[†]Corresponding author: VWI, Schanzeneckstrasse 1, CH-3012 Bern, Switzerland. Tel: +41(0)31-631-3989, Fax: +41(0)31-631-3992, email: `harris.dellas@vwi.unibe.ch`, Homepage: `http://www.vwi.unibe.ch/amakro/` http://www.vwi.unibe.ch/amakro

[‡]ana.fernandes@vwi.unibe.ch, http://anafernandes.name.

1 Introduction

Is financial development conducive to greater product market competition in financially dependent sectors? In particular, does it matter for the equilibrium number and size of firms, as well as for the markups these firms charge? Haber (2000) performs a comparative study of the cotton industry in Brazil and Mexico during a period of liberalization of the financial sector in both countries, namely from 1880-1930. He argues that financial liberalization resulted in an improvement in the competitive environment and a decrease in concentration indexes, in particular in Brazil, the country that underwent the most effective financial market reform. Rajan and Zingales (2003) contend that an important element in the development of financial markets – or lack thereof – is the ability of incumbent firms in financially dependent sectors to limit the development of those markets. Incumbents' motivations for opposing financial development would stem from the fact that a more developed financial market breeds competitive behavior and reduces the profitability of established firms. Finally, Cetorelli and Strahan (2006) find that higher concentration in local U.S. banking markets increases the hurdles to potential entrants in nonfinancial sectors. While financial development and greater competition in the financial sector are not synonymous, they do share the important property of being associated with greater fund availability to prospective entrepreneurs.

These examples from the literature are representative of a more general notion that financial development is somehow favorable to competition. This belief probably arises from the fact that financial constraints represent a barrier to undertaking financially dependent activities and may thus limit entry and –potentially– competition in the corresponding product/service markets. Interestingly, although various functions of the financial system have been the object of intense study (see Levine, 1997), to our knowledge, no formal analysis of the relationship between finance and product market competition or market structure has been carried out. The goal of the present paper is to fill in this gap. An additional goal is to shed light on the political economy issues regarding the forces that favor or oppose financial liberalization/development.

The financial sector performs a variety of roles,[1] such as the facilitation of risk hedging, the exertion of corporate control and the allocation of savings to projects. In this paper we focus on the financial sector in its role of allocating funds from multiple savers toward investment projects. The analysis is done in the context of a general equilibrium, two-sector model, with heterogeneous agents. The latter may differ with respect to their productivity levels as well as to their wealth holdings. One sector (sector 1) is financially dependent in the sense that its output depends on the amount of capital used, and that capital may need to be borrowed from the financial system. It is also imperfectly competitive, with firms competing *a la* Cournot. In the presence of unfettered financial markets, entrepreneurs are able to invest their optimal amount. When financial markets are not available (or restricted), output is limited by the entrepreneur's private wealth. In the other sector (sector 2), productivity is independent of capital and the market structure is perfectly competitive. There is free entry in both sectors.

We use this model to compute the general equilibrium with and without financial markets.[2] In particular, we are interested in providing answers to the following questions: How does financial deepening affect the quantity and price of the goods produced in capital dependent sectors? What happens to the number of firms in these sectors when financial imperfections are eliminated? What happens to the degree of product market *competition*, as measured by standard indicators of market shares and markups? What is the role played by the level and distribution of wealth?

As expected, we find that financial markets make the financially dependent sector expand and its price drop. This seems consistent with the finding of Rajan and Zingales (1998) that industries more dependent on external finance grow relatively faster in more financially developed countries. It is an implication of the fact that financial development expands *asymmetrically* the economy's production possibilities frontier. The implications for the number of firms, firm size, market concentration indexes and markups are more intricate. Financial development may, in general, lead to both entry

[1]See Levine (1997) for a comprehensive survey.

[2]One could easily consider intermediate cases of financial market imperfections without affecting the main results.

and exit in the financially dependent sector. Producers who owned their presence in that sector more to their wealth than to their ability will be driven out. At the same time, depending on the level and distribution of wealth and ability, new firms may be created by able but previously constrained agents. The net effect depends on the level and distribution of wealth. In general, relatively wealthy countries will tend to experience a decrease in the number of firms. Average firm size will go up and so may standard concentration measures. The effect on the number of firms is ambiguous for low income countries.

What happens to markups? The answer again depends on the level of wealth as well as on the distribution of ability and wealth. If per capita wealth is relatively low and wealth and ability are not correlated, then the development of financial markets lowers markups. If, on the other hand, the level of wealth is relatively high then concentration increases, both in terms of number of firms and market shares for the largest firms. Notwithstanding the reduction in the number of firms, the effect on markups cannot be fully signed, but it is conceivable, if not likely, that markups decrease in this economy too.

The model thus reveals the existence of interesting, potentially conflicting patterns in the behavior of commonly used measures of "competition." Changes in the number of firms (or changes in concentration indexes) and the size of markups may be positively related – rather than negatively related as is commonly presumed. One should then be careful in interpreting an increase in concentration as being necessarily harmful, at least in the presence of heterogenous producers and additional distortions (such as financial imperfections). But if markups go down, then our analysis suggests that financial development (or liberalization) has positive economic effects that go beyond those that have been identified in the existing literature regarding the improvement in the allocation of resources.

The model also has implications that are relevant for two other literatures. One regards the relationship between wealth (income) and productivity. For a given level of imperfect financial development, our analysis indicates that poorer countries will be more likely to have less competitive markets (as measured by markups) than richer countries, and hence lower *production* per capita. This seems consistent with both anecdotal evidence and the arguments of Parente and Prescott (1998) about market structure in less developed countries. The other literature regards the political economy of financial liberalization. A popular view is that the development of financial markets is hindered by the power of incumbents. This includes both incumbent financial institutions that are concerned about competition in the financial market, as well as incumbent firms that fear that a more competitive financial system will finance entrants into their sectors. The model shows that incumbency may not be sufficient to characterize preferences towards financial liberalization. There may be incumbents who will support liberalization (efficient but undercapitalized producers), as well as incumbent who may object to it (efficient and well capitalized firms). It is the fact that financial markets tend to favor disproportionately the most efficient but poorly capitalized producers that can account for such divisions within the class of incumbent firms.

The paper proceeds as follows. In the next section, we present our model. Section 2 describes the model, results are presented in section 3, section 4 provides some generalizations, and section 5 concludes.

2 The Model

2.1 The Environment

Preliminaries The economy is populated by a finite number of individuals N. Individuals differ with respect to their ability levels and private wealth holdings. Let $\mathcal{A}$ and $\mathcal{K}$ denote the corresponding sets of ability and wealth. We have $\mathcal{A} = \{a_j\}_{j=1}^{N}$ and $\mathcal{K} = \left\{\bar{k}_l\right\}_{l=1}^{N}$. Individual i is defined by a pair $\left(a_i, \bar{k}_i\right) \in \mathcal{A} \times \mathcal{K}$, $i = 1, \ldots, N$. $\bar{K}$ denotes the economy's global endowment of capital, given by $\sum_i \bar{k}_i$. Ability and wealth holdings are publicly observed and common knowledge. $F(a, k)$ denotes the joint cdf of ability and wealth in the economy.[3]

[3]We make the assumption of finite N since it seems more in line with deviations from price-taking behavior. However, in the remainder of the paper, we effectively treat the economy as continuous.

Production There are two goods produced in this economy. Good 1 requires capital as an input. If individual i works in sector 1, its output q_i is

$$q_i = a_i k_i^{\beta}, \tag{1}$$

with $\beta \in (0,1)$. k_i is the amount of capital individual i invests in production.

Output in sector 2 is independent of ability and, moreover, it does not require the use of any capital. If individual i chooses to work in sector 2, he produces A units of good 2.

In the presence of financial markets, individuals may borrow capital if their desired scale of operations exceeds their individual capital holdings. Without financial markets, individual investment is constrained to satisfy:

$$k_i \leq \bar{k}_i. \tag{2}$$

Since the ability to borrow affects the scale of individual production in sector 1 but not in sector 2, sector 1 is said to be a financially dependent sector.

Consumption Individuals have utility defined over two goods. Given consumption vector (c_1, c_2), total utility is

$$u(c_1, c_2) = \log(c_1) + \gamma \log(c_2),$$

for $\gamma > 0$. Let p be the relative price of good 1 in terms of good 2.

Wealth endowments, $\bar{k}_l \in \mathcal{K}$, are expressed in units of good 2. The income, l_i, of an individual i who chooses to operate in sector 1 is:

$$l_i = \bar{k}_i + (pq_i(k_i) - k_i).$$

and that of one who operates in sector 2:

$$l_i = A + \bar{k}_i.$$

Individuals buy (or sell) the difference between $\bar{k}_i$ and k_i at the price of one. If they choose to work in sector 1, quantity choice q_i will solve a standard problem of profit maximization *a la* Cournot, spelled out below.

The timing is as follows. Individuals first select the sector in which they wish to work and engage in production. Then, markets for both goods open and consumption takes place.

Aggregate Demand The budget constraint for i is:

$$pc_i^1 + c_i^2 = l_i.$$

Given income level l_i, demand for goods 1 and 2 is:

$$c_i^1 = \frac{1}{p}\frac{l_i}{1+\gamma}, \quad c_i^2 = \frac{\gamma}{1+\gamma} l_i.$$

Since Engel curves are straight lines from the origin, we have a representative agent economy. Aggregate demand depends only on aggregate income (the sum of individual income across individuals) and not on how it is distributed. Define

$$I \equiv \sum_{i=1}^{N} l_i,$$

so that I is aggregate income. Aggregate demand for good j, denoted C^j, is then

$$C^1 = \frac{1}{p}\frac{I}{1+\gamma}, \tag{3}$$

$$C^2 = \frac{\gamma}{1+\gamma} I. \tag{4}$$

The inverted demand curve for good 1 is:

$$p = \frac{1}{C^1}\frac{I}{1+\gamma}. \tag{5}$$

The log preference format dictates that the (absolute value of the) elasticity of demand of good 1 with respect to its relative price p equal 1. For later use, we note that the relative demand schedule of good 1 in terms of good 2 is:

$$\frac{C^1}{C^2} = \frac{\frac{1}{p}\frac{I}{1+\gamma}}{\frac{\gamma}{1+\gamma} I} = \frac{1}{p}\frac{1}{\gamma}. \tag{6}$$

Relative demand of good 1 does not depend on aggregate income and is a negative function of the relative price p.

Next, we examine the economy without financial constraints.

2.2 Financially Unconstrained Economy

We first describe how a firm that has chosen sector 1 selects its optimal level of production. We then describe how firms choose their sector of activity.

2.2.1 Optimal Choice of Level of Production in Sector 1

Consider individual i, whose ability endowment is a_i. Profits from operating in sector 1 are:

$$\pi_i^1 = p\left(Q_1\right)q_i - k_i = p\left(Q_1\right)q_i - \left(\frac{q_i}{a_i}\right)^{\frac{1}{\beta}}, \tag{7}$$

where q_i is the quantity produced by individual i, k_i the amount of capital used, Q_1 the total output of good 1 produced, and $p\left(Q\right)$ the inverse demand curve for good 1. We assume that sector 1 is characterized by quantity competition *a la* Cournot, while sector 2 is perfectly competitive. Let Q_{1_i} denote the output of good 1 by producers other than i. Then,

$$Q_1 = q_i + Q_{1\_i}.$$

Under Cournot competition, each firm chooses output q_i taking the quantities of the remaining firms Q_{1_i} as given. The first-order condition for firm i is:

$$p\left(1-\frac{q_i}{Q_1}\right) = MC\left(q_i\right), \tag{8}$$

where $MC\left(q_i\right)$ indicates firm i's marginal cost. We thus obtain the familiar result that the price to marginal cost ratio, the markup, equals $\left(1-q_i/Q_1\right)^{-1}$. Using the expression for the production function to relate q_i and k_i, and inserting into (8) yields:

$$p\left(1-\frac{q_i}{Q_1}\right) = \frac{1}{\beta}\left(\frac{1}{a_i}\right)^{\frac{1}{\beta}}\left(q_i\right)^{\frac{1-\beta}{\beta}}. \tag{9}$$

Clearly, $MC\left(q_i\right)$ is strictly increasing in q_i.

Equation (9) defines firm i's optimal quantity q_i in terms of the relative price p, total market output Q_1, and level of efficiency, a_i. Optimal quantity q_i is strictly increasing in ability. Holding p and Q_1 constant, more able firms will have greater market shares and higher markups than less able ones. Since the marginal cost declines with ability, more able firms are more profitable than less able ones.

Equation (9) allows us to determine the optimal quantity produced by firm i, $q_i^\star$. It can be written as:

$$q_i^* = q_i^* \left(\overset{+}{p}, \overset{+}{Q_1}, \overset{+}{a_i} \right). \tag{10}$$

The first-order condition (8) can also be written in terms of capital:

$$\beta a_i \left(p + p'\right) = k_i^{1-\beta}.$$

It follows that more able firms choose a larger scale of production. Together with the properties of q_i in (10), we have:

$$k_i^* = k_i^* \left(\overset{+}{p}, \overset{+}{Q_1}, \overset{+}{a_i} \right). \tag{11}$$

Figure **??** displays the optimal investment schedule $k_i^* (\cdot)$.

2.2.2 Optimal Choice of Sector of Activity

Individuals choose to work in the sector that generates the highest income. Individual i will choose sector 1 if

$$\max_{q_i} \pi_i^1 \left(q_i; a_i, p, Q_1\right) \geq A. \tag{12}$$

The profit function is an increasing function of ability. Under appropriate assumptions on the distribution of ability and the parameters of the model, there exists a level of ability, $\tilde{a}$, such that

$$\pi \left(q^* \left(p, \tilde{a}, Q_1\right); \tilde{a}, p, Q_1\right) = A. \tag{13}$$

The ability level $\tilde{a}$ is the threshold determining the separation of entrepreneurs into activities: those whose ability exceeds $\tilde{a}$ work in sector 1 whereas the remaining work in sector 2. This threshold is implicitly defined by equation (13) as a function of p and Q_1:

$$\tilde{a} = \tilde{a} \left(p, Q_1\right). \tag{14}$$

Once the choice of activity has been made and production completed, the aggregate supply of good 1, Q_1, is given by

$$Q_1 = \sum_{a_i \geq \tilde{a}(Q_1, p)}^{\bar{a}} q_i^* \left(\overset{+}{p}, \overset{+}{a_i}, \overset{+}{Q_1} \right). \tag{15}$$

Entrepreneurs engaged in the production of good 1 use good 2 as an input. Total input demand for good 1 production, K^*, is:

$$K^* = \sum_{a_i \geq \tilde{a}(Q_1,p)}^{\bar{a}} \left(\frac{q_i^*}{a_i}\right)^{\frac{1}{\beta}}. \tag{16}$$

Let Q_2 denote the total quantity of good 2 produced in the economy:

$$Q_2 = \sum_{a_i \leq \tilde{a}(Q_1,p)} A. \tag{17}$$

We define K as the difference between the initial endowment of capital, $\bar{K}$ and the amount of capital used up in production in sector 1, K^*:

$$K = \bar{K} - K^*.$$

The quantity of good 2 available for consumption is then simply

$$C_2 = Q_2 + K. \tag{18}$$

2.2.3 The Equilibrium

Definition 1 *An equilibrium in the financially unconstrained economy is a triple* $\left(\tilde{a}, p, \{q_i\}_{i=1}^N\right)$, *with* $Q_1 = \sum_{i=1}^N q_i$, *such that:*

i) q_i *satisfies (8) for* $a_i \geq \tilde{a}$ *and is equal to zero for* $a_i < \tilde{a}$*;*

ii) $C_1 = Q_1$, $C_2 = Q_2 + K$*;*

iii) Equation (6) is satisfied.

In order to construct the equilibrium we proceed as follows. First we use (13) to solve for $\tilde{a} = \tilde{a}(p, Q_1)$ (see figure 1). Substituting this expression into (15), (17) and (16) yields Q_1, Q_2 and K^*, respectively, as functions of p. Substituting Q_1 for C_1 and $Q_2 + K$ for C_2 in (6) determines p. In our economy, an unconstrained equilibrium exists and is unique.[4]

[4] Since marginal cost is increasing, the production of greater quantities of good 1 requires increasing amounts of capital as well as an increasing number of entrepreneurs. Both forces show that, as C_1 increases, an increasingly greater amount of C_2 must be foregone to generate an extra unit of C_1. This shows that the consumption possibilities frontier of the economy is strictly concave. The $\log(\cdot)$ preference format generates strictly convex indifference curves and, therefore, the tangency point is unique.

2.3 Financially constrained economy

We now discuss the determination of the equilibrium in this economy in the *absence* of financial markets.[5] The superscript c is used to indicate equilibrium values in the constrained economy. Let q_i^c denote individual i 's output in the constrained environment. If this entrepreneur cannot use external funds, then his production in sector 1 cannot exceed the level that could be financed by his own initial capital stock, $\bar{k}_i$:

$$q_i^c = min\{a_i\bar{k}_i^{\beta}, a_i(k_i^*)^{\beta}\} \tag{19}$$

where k_i^* is determined by equation (11). Unlike the financially unconstrained economy, the choice of activity now depends on both individual ability and individual wealth. Let us consider the minimum amount of capital, k_{min} that makes an individual of ability level a indifferent between the two sectors. $k_{\min}(a,p)$ is thus determined by the equation:

$$pa_ik_{\min}^{\beta} - k_{\min} = A. \tag{20}$$

Equation (20) defines a relationship between ability and capital. It can be verified that $k_{\min}(a,p)$ is decreasing and convex in ability. Since a higher price raises profits for given levels of ability and capital, an increase in p shifts the $k_{\min}(\cdot)$ schedule downwards.

The threshold ability level for the choice of activity in the financially constrained economy, $\tilde{a}^c$, as well as the corresponding amount of capital, $k_{\min}(\tilde{a},p)$, are then determined by combining equations (11) and (20). All individuals with ability $a \geq \tilde{a}^c$ who have wealth $k \geq k_{min}(a)$ will operate in sector 1. The rest, those who are not simultaneously able and rich enough to operate in sector 1, will work in sector 2.

The introduction of financial constraints will make at least some of the entrepreneurs financially constrained. Some of them will afford (have sufficient funds) to stay in sector 1 but will operate below their optimal capacity. The rest will move to sector 2. Figures 2 and 3 offer two *examples* of a financially constrained equilibrium under the assumption that all individuals have a common level of initial wealth, $\bar{k}_i = \bar{k}$. And that this common

[5]While for reasons of simplicity we study an economy without any asset trade, our analysis is applicable to more general environments with asset markets but with restricted asset trade.

level is such that in Figure 2 some entrepreneurs are able to achieve their optimal capacity (those with $a \geq a'$). While in Figure 3 none of them can achieve their optimal scale of production.

The Equilibrium

Definition 2 *An equilibrium in the financially constrained economy is a triple* $\left(\tilde{a}^c, p^c, \{q_i^c\}_{i=1}^N\right)$, *with* $Q_1^c = \sum_{i=1}^N q_i^c = \sum_{i=1}^N a_i (k_i^c)^\beta$, *such that:*

ia. $k_i{}^c = \min\{k_i^\star, \bar{k}\}$ *if* $k_i{}^c \geq k_{min}$ *and zero otherwise*

ib. $k_i^\star$ *satisfies (11) and* k_{min} *satisfies (20)*

ii. $C_1 = Q_1$, $C_2 = Q_2 + K$

iii. Equation (6*) is satisfied.*

Conditions (ia-ib) state that a firm will operate in sector 1 if it is profitable and affordable to do so. If it operates in this sector, it will either produce its optimal quantity –if unconstrained– or, if constrained, the quantity allowed by its capital stock.

3 The effects of financial markets

3.1 Finance and the Allocation of Resources

We now study the implications of the elimination of financial markets for the allocation of resources, the number of firms in sector 1, equilibrium markups and the relative price p. We start by assuming that the initial level of wealth is the same for all agents, $\bar{k}_i = \bar{k}$ for all i, and that the financial constraint binds for at least some agents. We consider two cases, one with high per capita wealth and the other with low per capital wealth. In the first one, the common level of *per capita* wealth $\bar{k}$ exceeds the demand for capital by the marginal firm in the unconstrained economy. That is,

$$k_i^* \left(p^u, Q_1^u, \tilde{a}^u\right) < \bar{k}. \tag{21}$$

This is the "high" per capita wealth situation depicted in Figure 2. In the second case (the low wealth case, Figure 3), the level of per capita wealth is below the optimal investment scale of the marginal firm in the unconstrained economy, $k_i^* \left(p^u, Q_1^u, \tilde{a}^u\right) > \bar{k}$.

These two cases will provide helpful insights concerning the effects of imperfect financial intermediation. We discuss later the implications of using a more general specification of the wealth distribution. Note that, under the current wealth distribution, the equilibrium number of firms in sector 1 varies monotonically – and inversely – with $\tilde{a}$: higher values of this threshold are associated with a lower number of good 1 producers.

We establish that, independently of the level of wealth, the equilibrium relative price of the financially dependent good, p, is higher in the absence of financial markets (Proposition 1) while its quantity, Q_1 is lower (Proposition 3). The change in the number of firms is ambiguous and depends on the level of initial –per capita– wealth (Proposition 4). If this level is sufficiently high (a notion to be made more precise in Proposition 4), then the elimination of financial markets brings about an *expansion* in the number of firms operating in the financially dependent sector, while firm size shrinks. On the other hand, if initial wealth is low, then the number of firms in this sector decreases. In the next subsection we examine the implications of these changes for markups. Interestingly, we find that markups may increase –making the market less competitive– irrespective of what happens to the number of firms.

Let $\left(p^u, \tilde{a}^u, \{q_i^u\}_{i=1}^N\right)$ denote the equilibrium in the unconstrained economy, and $\left(p^c, \tilde{a}^c, \{q_i^c\}_{i=1}^N\right)$ denote the equilibrium in the constrained economy. The proofs of the following propositions can be found in the appendix.

Proposition 1 *The financially constrained economy has a higher relative price of the financially dependent good.*

Proposition 2 *In the high wealth economy, the introduction of financial constraints results in firm entry into sector 1: $\tilde{a}^c < \tilde{a}^u$.*

Proposition 3 *Production of the financially dependent good is lower in the constrained economy.*

Proposition 4 *Let $k_{\min}(\tilde{a}^u, p^c)$ represent the minimum scale required by firm $\tilde{a}^u$ (the threshold firm in the unconstrained equilibrium) in order to find it profitable to remain* *in sector 1 (equation 20). There will be firm exit following the introduction of financial constraints if $\bar{k} < k_{\min}(\tilde{a}^u, p^c)$.*

In a sufficiently wealthy economy, that is, in an economy where the condition in Proposition 4 is violated, the imposition of financial constraints will be associated with an *increase* in the number of firms in the financially dependent sector. In a low wealth economy, on the other hand, firm exit will take place following the elimination of financial markets.[6] In the latter case, all the firms that remain in sector 1 are financially constrained.

Proposition 5 *If the introduction of financial constraints leads to firm exit, then all firms operating in the financially dependent sector will be financially constrained.*

In the case of entry, either all or only some of the remaining firms are financially constrained. Without further assumptions it is not possible to offer any general results on these properties of the constrained equilibrium.

These propositions establish that the elimination of financial markets leads to a higher price and a lower quantity for the financially dependent good. This lower level of output may be produced by a smaller or larger number of firms, depending on the level (and also distribution) of wealth in the economy. Can anything be said about the behavior of various concentration indexes, as well as of markups?

3.2 Finance and Competition

Markups The firms operating in the financially constrained sector after the shut-down of financial markets can be partitioned into three groups: new entrants (NE), unconstrained incumbent firms (UI), and constrained incumbent firms (CI). In the low wealth economy, only the last group is present. In the high wealth economy, all groups

[6] Note that our usage of the concept of high or low wealth is related but does not completely correspond to customary usage. This is because our classification into high or low wealth status relies on $k_{\min}(\tilde{a}^u, p^c) - \bar{k}$, rather than $\bar{k}$ alone. As such, it depends also on the distribution of ability as well as on the various determinants of the relative price (productivity, preferences and so on). Nonetheless, for countries that differ mostly along the wealth dimension, our usage and the customary one coincide.

are present. For the first two groups, markups are determined from the firm's first-order condition, equation (8). They are given by:

$$\frac{p}{MC_i} = \psi_i = (1 - \frac{q_i}{Q_1})^{-1}. \tag{22}$$

For the financially constrained CI group, the markup is:

$$\frac{p}{MC_i} = \frac{p}{MC\left(q(\overline{k})\right)}. \tag{23}$$

Proposition 6 *The introduction of financial constraints leads to an increase in the market shares of the financially unconstrained incumbent firms (UI).*

Proposition 7 *Incumbent firms charge a higher mark up when they are financially constrained.*

Proposition 6 implies that markups unambiguously increase in a low wealth, financially constrained economy when there is firm exit. This is because the firms that remain in the market are those that had the highest market shares – and thus the higher markups – in the unconstrained economy. And these firms now have even higher markups because they face a higher price and have lower marginal costs (due to the fact that they produce less).

The situation is more complicated when there are new entrants (NE). This group consists of firms whose markups may be below those of the marginal firm in the economy with asset markets. The new entrants are firms with a higher marginal cost than the (previously) marginal firm. But they also face a higher price, and so the net effect is ambiguous. In the low wealth case, since all firms producing good 1 are financially constrained and thus produce the same quantity, it is possibly to show that a standard measure of markups such as the Lerner index increases. In the high wealth economy, however, without information on the exact distribution of output across firms, it is not possible to sign the change in that index. Nevertheless, it can be established that an "unweighted" version of the Lerner index (the unweighted *sum* of markups across all firms) is lower in the unconstrained economy.

Define μ_i^c for firm i in the constrained economy as:

$$\mu_i \equiv \frac{p^c - MC_i^c}{p^c},$$

and let μ_i^u be defined similarly. Then:

Proposition 8 *A. The unweighted Lerner concentration index in sector 1 is strictly lower in the financially unconstrained economy. That is:*

$$\sum_{a_i \geq \tilde{a}\left(p^c, Q_1^c\right)} \left(\mu_i^c\right) < \sum_{a_i \geq \tilde{a}\left(p^u, Q_1^u\right)} \left(\mu_i^u\right).$$

B. The standard Lerner concentration index in sector 1 is lower in a low wealth, financially unconstrained economy.

Part A means that the simple sum of the markups on the marginal unit of all firms is lower in the financially unconstrained equilibrium. Part B implies that the market share–weighted sum of markups across all firms is also lower in a low wealth economy. Consequently, financial development (liberalization) makes the markets for financially dependent goods in low wealth countries more competitive. Do the changes in firm concentration point in the same direction?

Concentration indexes Let us order the firms by size so that 1 represents the largest firm, 2 the second largest firm and so on. The H_j index of the market share of the j largest firms is defined as:

$$H_j \equiv \sum_{i=1}^{j} \frac{q_i}{Q_1}.$$

Under the two wealth distributions considered above, it is always the case that the largest firms are financially constrained, and thus produce less in the constrained economy. Let r represent the set of financially constrained firms in this economy. We have:

Proposition 9 *The concentration ratio H_j is strictly higher in the financially unconstrained economy for $j \leq r$.*

Note that the group of financially constrained firms does not necessarily contain only firms that produce less and have lost market share as a result of the imposition of financial constraints. It is possible that some of these firms became financially constrained because they expanded their production levels relative to the unconstrained environment. These will be the smallest firms within the set of financially constrained ones. Nevertheless, it is still the case that the market share of the incumbent constrained firms as a whole declines with the introduction of financial constraints (due to the existence of the new entrants), and this implies that the largest firms unambiguously lose market share.

How does financial development affect H_j when $j > r$? A subset of the j firms expands production and gains market share. This subset includes the firms in the UI group and possibly firms in the CI group that expanded their output (though suboptimally). The remaining firms lose market share. These are a subset of the CI firms. While there is a presumption that greater concentration will prevail (because it is the larger firms that expand), without additional assumptions on the structure of the economy it may not be possible to sign the total effect. The only thing that can be said with certainty, in addition to Proposition 8, is that the average market share – the ratio of total output Q_1 to the number of producers – does increase with financial development (liberalization) in the high wealth economy.

Our analysis thus shows that financial development (or liberalization) may affect standard measures of firm concentration and competition differently. A regulator examining developments in sector 1 may witness an increase in the average market share as well as greater concentration among the largest producers. But if he could measure marginal costs, he would also observe that incumbent firms have reduced their markups and that the average markup may have gone down. What conclusion is to be drawn in such a situation?

Concentration and efficiency are usually interpreted as monotonic measures of consumer well-being (i.e. lower concentration and lower markups/Lerner indexes are usually seen as conducive to a larger consumer surplus). However, firm concentration measures are not good indicators of consumer well-being in a second-best world with financial constraints (or, presumably, other distortions). In fact, more able producers produce larger quantities at lower total cost than less able ones. They do so while charging a higher markup: they are able to produce the last unit at a strictly lower marginal cost compared to lesser able entrepreneurs. As such, shifting output toward these firms – and raising concentration – would be in the best interest of consumers provided such added concentration would not result in an endogenous increase in markups. The analysis shows that the removal of financial constraints may bring about such benefits.

4 Additional considerations

More General Distribution of Wealth The joint distribution of wealth and ability considered in the previous section is special because it is degenerate along the wealth dimension (everybody has the same endowment of capital). It can be extended in two distinct ways. It can be made to depend on the level of ability while maintaining the assumption that there is a single level of capital corresponding to any particular level of ability. Or, instead of having the agents being distributed over the line with regard to individual wealth, they can be distributed over a two dimensional plane. That is, several individual of different levels of wealth may share the same level of ability and there is positive mass at all levels of ability. Of course, these two extensions could be combined to have the most general case of a non-degenerate distribution of wealth over the two dimensional plane. The effects of introducing dependence between the distributions of ability and wealth are both obvious and difficult to pin down quantitatively without imposing further structure. Allowing the agents to be distributed over a two dimensional plane is equivalent to combining several capital lines like the $\bar{k}$ line in Figures 2 or 3. Some of these lines may even involve unconstrained agents across the whole ability range. The main difference from the analysis in the previous section is that the imposition of financial constraints may now trigger both entry and exit in market 1. The new entrants will be those entrepreneurs whose ability a is such that $a \in [\tilde{a}^c, \tilde{a}^u]$, and whose wealth places them above the $k_{\min}(\cdot)$ schedule. The firms exiting the market, have ability levels above $\tilde{a}^u$, and wealth levels below the $k_{\min}(\cdot)$ schedule. These are very able but poor entrepreneurs, who need to borrow in order to be in sector 1. Without further assumptions one cannot unambiguously determine the net effect on the number of firms or average markups.

Another difference regards the effects of financial development on the standard H_j indexes. While propositions 1-8 could still obtain under plausible generalizations of the wealth distribution, proposition 9 would not. Consider, for example the case with two wealth classes, $\bar{k}_1$ and $\bar{k}_2$ with $\bar{k}_2 > \bar{k}_1$. Suppose that for agents with $k_i = \bar{k}_2$ the financial constraint is never binding, independent of the level of ability. And for agents with $k_i = \bar{k}_1$ the financial constraint is binding at high levels of ability. This means that in the constrained equilibrium, some of the largest firms are unconstrained and thus have a larger market share relative to the case with unconstrained financial markets. Thus standard concentration may decrease following financial liberalization if the most productive firms have sufficient own funds to achieve their optimal scale of production under the financially constrained environment.

The Case of Homogeneous Ability We have also studied the special case of a homogeneous level of ability, $a_i = a$ for all i. As in the more general case, the pattern of firm entry and exit, the level of mark ups and the change in the concentration indexes varies with the distribution of wealth and cannot be signed unambiguously. As before, to the extent that there exist some firms that are not financially constrained, financial liberalization is likely to decrease concentration indexes.

Political Economy Considerations The model can be used to study the political economy of financial liberalization/development. Which groups in the population would favor and which ones would oppose liberalization?

In general, in the financially constrained economy there are three groups: Two in sector 1, namely, financially unconstrained and financially constrained firms. And one in sector 2. Following financial liberalization, some firms from sector 1 (constrained and unconstrained) will exit while some firms from the sector 2 will enter sector 1. The welfare of some of these groups can be unambiguously signed. For instance, those who remain in sector 2 as well as those who leave sector 2 and enter sector 1 are better off. The former because the purchasing power of their output, A, increases due to the lower p. And the latter because they can always do at least as well as the former group by opting to stay in sector 2. For the other groups the situation is more nuanced because while the change in the level of profits can be signed, the implications for consumption (and utility) are harder to derive. Nonetheless, an interesting implication of our analysis is that incumbency does not necessarily imply opposition to liberalization. One of the reasons for this is that some of the incumbent firms may be financially constrained and hence unable to achieve their optimal scale of production. Consequently, it is to be expected that the firms operating in sector 1 may not speak with a single voice on issues of financial liberalization. While firms (or individuals) outside the sector may have more homogeneous views and in all likelihood favor financial liberalization.

Empirical Implications In an influential paper, Rajan and Zingales, 1998, have studied the effects of financial development on the growth rate of financially dependent industries. They decompose industry growth into growth in the number of establishments and growth in the average size of existing establishments. They state that their "... estimates suggest that financial development has almost twice the economic effect on the growth of the number of establishments as it has on growth of the average size of establishments[7] The analysis in our paper does not contain a growth mechanism so it cannot be used to deal with ongoing growth. Nonetheless, it can still be used to address the issue of how the increase (growth) in output of financially dependent sectors that results from the amelioration of the financial constraints can be divided into the two

[7]But at the same time they also report that, in explaining the relative growth of industries, the dependence of the young firms was lower than that of the mature firms. This seems to point in the opposite direction.

sources of growth discussed by Rajan and Zingales. As discussed above the number of firms may increase or decrease following the loosening of the financial constraints, depending on, among other things, the level and distribution of wealth. When the number of firms increases, it is quite possible for the model to generate a growth rate in the number of establishments that exceeds the growth rate in the average size of establishments (and vice versa). A similar pattern could be generated not through the level of wealth but through dependence of the distributions of wealth and ability. For instance, assuming that most of the rich are able but there are also many able who are poor would deliver the same pattern.

5 Conclusion

The effects of financial development (or financial liberalization) on the allocation of resources, economic growth and welfare have been extensively studied in the literature. There is one aspect, though, that has been completely ignored, notwithstanding a widely held belief that it is of great importance for economic performance and welfare. Namely, the relationship between finance and competition in product markets. In this paper we have taken a first step in characterizing the effects of the existence of financial constraints on competition. Admittedly, the analysis has been carried out in a framework that has been restricted in order to make it feasible to study such a complicated issue. For instance, the nature of the financial constraints has not been modelled. They correspond closer to unspecified costs of asset trade than to the elaborate agency problems typically discussed in the literature. Dynamics have been abstracted from. And so on.

Nevertheless, a number of novel insights emerge that are likely to also arise in more general environments. First, under a relatively flat distribution of income, in poor countries financial development increases competition in product markets. And it is likely to do the same in rich countries. Second, the standard measures of competition used by policymakers (various firm concentration indexes) may be misleading in the evaluation of the effects –and hence, of the desirability– of policies regarding financial liberalization. Higher concentration may well be associated with an improvement in the competitive environment. This suggests that it may be necessary to carry out such evaluations in models –such as ours– that allow for meaningful heterogeneity in efficiency across firms. Third, not only does the distribution of wealth but also per capita wealth plays an important role in the determination of the degree of competition in product markets. Poorer countries are likely to have less competitive markets and hence stand to benefit more from financial liberalization. And forth, being an incumbent firm (and even a large sized one) does not automatically create a bias in favor of opposition to financial liberalization. Some of the incumbent firms stand to benefit the most from liberalization.

There is a number of demanding but important extensions awaiting. One could involve the incorporation of dynamics, so that both the cross section and time series properties of the distribution of firms in financially dependent sectors could be derived. Another one might involve the endogenization of financial constraints.

References

Cetorelli, Nicola and Strahan, Philip E. (2006) "Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Market" *Journal of Finance* 61:1, 437-461.

Harber, Stephen (2000) "Banks, Financial Markets, and Industrial Development: Lessons from the Economic Histories of Brazil and Mexico." *Center for Research on Economic Development and Policy Reform Working Paper* 79.

Levine, Ross (2005) "Finance and Growth: Theory and Evidence." *in* Handbook of Economic Growth Philippe Aghion and Steven Durlauf eds. (North-Holland) Vol. 1, Part 1, 865-934.

Rajan, Raghuram G. and Luigi Zingales (2003) "The great reversals: the politics of financial development in the twentieth century." *Journal of Financial Economics* 69, 5-50.

6 Appendix

Proposition 1: *The financially constrained economy has a higher relative price of the financially dependent good.*

Proof.

i. Let us shut the financial markets down and hold, for the moment, p fixed. Without borrowed capital, the output of the financially constrained firms will decrease as these firms are now forced to operate below their optimal capacity.

i1.In the high wealth case, the financially unconstrained firms will also reduce their optimal output due to equation 8. Since the price is being held constant, lower quantities imply lower profits and there is firm exit. Thus, Q_1 and C_1 decrease while K and Q_2 and consequently C_2 increases. At p^u condition 6 is no longer satisfied and there is excess demand for good 1.

i2. In the low wealth case, $k_{\min}(\tilde{a}) > \bar{k}$, so there is also firm exit. This reinforces the reduction in Q_1 and leads –as in case (i1) above– to excess demand for good 1.

ii) How can the excess demand for good 1 eliminated? It requires the increase in its relative price, p. An increase in p has two effects: First, it increases the optimal quantity of the unconstrained firms (if there are any) due to equation 8. (Note from 8 that both higher p and higher Q_1 raise the optimal quantity q_i and so the effects reinforce each other.) The higher price has additional effects. Holding quantities constant, profits of incumbent firms increase. In addition, the optimal quantity has increased as well (for unconstrained firms), and this pushes profits even higher. Higher profits induce entry due to equation 20. Both of these changes contribute to a higher C_1 and a lower C_2, thus helping to restore 6.

Proposition 2 In the high wealth economy, the introduction of financial constraints results in firm entry into sector 1: $\tilde{a}^c < \tilde{a}^u$.

Proof. Suppose now, for contradiction, that the new equilibrium value of $\tilde{a}$ could exceed $\tilde{a}^u$. For this to be possible, and given that the new equilibrium price will be higher than before, $p^c > p^u$, it must be the case that the optimal quantities produced by unconstrained incumbents are lower than before. This is a necessary condition for the new $k^*(\cdot)$ and $k_{\min}(\cdot)$ schedules to cross to the right of $\tilde{a}^u$, given that Proposition 1 establishes a lower p in equilibrium.

Note that, if $\tilde{a}^c > \tilde{a}^u$ and optimal quantities are lower, since

$$C_2 = K - \sum_{a \geq \tilde{a}} \left(\frac{q_i}{a_i} \right)^{\frac{1}{\beta}} + \sum_{a < \tilde{a}} A,$$

the equilibrium value of C_2^c will be higher than C_2^u. This is so since there are fewer firms in sector 1 – and therefore more of good 2 is produced – and the remaining firms in sector 1 are fewer and produce less than before – so less good 2 is demanded as an input into good 1 production. From equation (4), it follows that $I^c > I^u$ must hold as well. Finally, it must also be the case that $C_1^c < C_1^u$ since there are fewer producers of good 1 and each produces less than before.

Consider now the first-order condition for good 1 production:

$$\begin{aligned} p\left(1 - \frac{q_i}{Q_1}\right) &= MC(q_i) \iff \\ \frac{I}{1+\gamma}\frac{1}{Q_1}\left(1 - \frac{q_i}{Q_1}\right) &= MC(q_i). \end{aligned}$$

Under common regularity conditions,[8] when I is fixed, a lower Q_1 leads firms to increase their output. Therefore, for lower q_i to be optimal, it must be the case that I has declined (lower I and lower Q_1 are still compatible with a higher output price). This, however, contradicts the fact that $C_2^c > C_2^u$. ■

Proposition 3. *Production of the financially dependent good is lower in the constrained economy:* $Q_1^c < Q_1^u$.

[8]These amount to assuming that the best-response function of individual firms is negatively sloped. In our model, that will always be the case for firm i provided its market share satisfies:

$$2\frac{q_i}{Q_1} - 1 < 0.$$

In words, firm i's output must be smaller than half the total output of good 1. When firms are small – an assumption we made earlier in treating the economy as continuous – this assumption will always be met.

Proof. We consider the two cases corresponding to firm entry or exit. Recall that firm exit can only occur when $k(\tilde{a}^u) > \bar{k}$. This means that, in the constrained economy, there is a smaller number of producers left in sector 1, all of which produce strictly less that what they would have produced in the unconstrained economy (because $\bar{k} < k^{\star u}(a)$ for $a > \tilde{a}^u$). Therefore, $Q_1^c < Q_1^u$ follows immediately. Consider now the case when there is entry into sector 1, so that there are fewer producers of good 2 and, as a consequence, Q_2 declines. Assume that $Q_1^c > Q_1^u$. Because (at least) the most efficient firms are constrained, it must be the case that $K^*(Q_1^c) > K^*(Q_1^u)$. Therefore, $K^c < K^u$ and, since $Q_2^c < Q_2^u$ (recall $\tilde{a}^u < \tilde{a}^c$), $C_2^c < C_2^u$ violating the property of equilibrium that $(C_1/C_2)^c < (C_1/C_2)^u$. ■

Proposition 4: *Let $k_{\min}(\tilde{a}^u, p^c)$ represent the minimum scale required by firm $\tilde{a}^u$ (the threshold firm in the unconstrained equilibrium) in order to find it profitable to operate in sector 1 (equation 20). There will be firm exit following the introduction of financial constraints if $\bar{k} < k_{\min}(\tilde{a}^u, p^c)$.*

Proof. Consider Figure 3. A move from the unconstrained to the constrained economy results in a downward shift in the $k_{\min}$ schedule due to $p^c > p^u$ (see equation 20). The shift in the $k^\star$ is ambiguous as the higher p and lower Q_1 pull in opposite directions (equation 10). Do these two curves intersect to the left or the right of $\tilde{a}^u$? Intersection at a value $a > \tilde{a}^u$ would imply that the $\tilde{a}^u$ firm would have a minimum scale that exceeded its optimal scale and would thus choose not to operate in sector 1 even if it could afford to do so. This, however, cannot be true because this firm could always profitably choose to remain in sector 1 and simply produce the same amount it was producing before, namely, $q(\tilde{a}^u)$ as $p^c q(\tilde{a}^u) - C(q(\tilde{a}^u)) - A > p^u q(\tilde{a}^u) - C(q(\tilde{a}^u)) - A = 0$. Hence, the curves must intersect to the left of $\tilde{a}^u$. Let a' and k' denote the values of a and k at this intersection. If $\bar{k} > k'$ then all $a \in [a', \tilde{a}^u)$ enter. Consider now the case when $k' > \bar{k}$, and define the following threshold $\hat{a}$:

$$k_{\min}(p^c, \hat{a}) = \bar{k}.$$

Ability level $\hat{a}$ corresponds to the intersection of the schedule $k_{\min}(p^c, a)$ with the wealth level $\bar{k}$. If this ability level is lower than $\tilde{a}^u$, then entry will occur and all entrepreneurs with ability $a \in [\hat{a}, \tilde{a}^u)$ join sector 1. Exit will occur when $\hat{a} > \tilde{a}^u$. Since $k_{\min}(\cdot)$ is a decreasing schedule, $\hat{a} > \tilde{a}^u$ implies that $k_{\min}(\tilde{a}^u, p^c) > \bar{k}$. ■

Proposition 5: *If the introduction of financial constraints leads to firm exit, then all firms operating in the financially dependent sector are financially constrained.*

Proof. When there is firm exit, the new marginal firm will correspond to the intersection of the $k_{\min}$ schedule and the horizontal line $\bar{k}$. Because the $k_{\min}$ schedule is downward sloping, this intersection occurs to the right of the intersection of $k_{\min}$ and k^*. This means that the new marginal firm is producing using $\bar{k} < k^*$. Since all other producers have higher ability, they would demand even more capital than the marginal firm. Therefore, all good 1 producers are financially constrained.

More generally, it can be seen that the composition of firms in terms of constrained and unconstrained firms depends on the location of $\bar{k}$ in relationship to k'. In particular, if $\bar{k} > k'$ (see proposition 4) then there are some unconstrained firms in sector 1. If $k' > \bar{k}$ then all firms are constrained. ■

Proposition 6. *The introduction of financial constraints leads to an increase in the market shares of the financially unconstrained incumbent (UI) firms.*

Proof. Suppose instead that their market shares went down. This would require production by the unconstrained firms to decline proportionately more than the decline in total output Q_1. Since p has gone up, (8) would no longer be satisfied: its left-hand side went up while the right-hand side decreased (due to the fact that marginal cost is increasing). ■

Proposition 7: *The markups of incumbent firms increase once financial constraints are introduced.*

Proof. Consider first a high wealth economy. In this economy, there are different types of incumbent firms, unconstrained and constrained ones. In the latter group there exist two further sub-categories. The group of firms that produce less than in the unconstrained equilibrium (the most able ones). And the group of firms that produce more than in the unconstrained equilibrium. These are the firms that became financially constrained because they expanded their production in order to take advantage of the higher price (the intermediate ability ones). For the incumbent, unconstrained firms, Proposition 6 implies higher markups. Similarly, for the incumbent, constrained ones that lowered their production, equation (22) implies higher mark ups too. What about the markups of the last group, the constrained firms that increased their production? The optimal mark up for the firms in this group –that is, the markup they would charge if they were not financially constrained– goes up because their optimal market share increases. But because these firms are constrained, they produce less than the optimal amount, face a lower marginal cost and hence a higher markup relative to their optimal one.

The low wealth economy is straightforward since all firms in sector 1 are financially constrained. The proof that incumbents charge higher markups is the same as in the high wealth case. ■

Proposition 8: *A. The unweighted Lerner concentration index in sector 1 is strictly lower in the financially unconstrained economy. That is:*

$$\sum_{a_i \geq \tilde{a}\left(p^c, Q_1^c\right)} \left(\mu_i^u\right) < \sum_{a_i \geq \tilde{a}\left(p^u, Q_1^u\right)} \left(\mu_i^c\right).$$

B. The standard Lerner concentration index in sector 1 is lower in a low wealth, financially unconstrained economy.

Proof. Part A Equation (8) implies that

$$\mu_i = \frac{p - MC_i}{p} = \frac{q_i}{Q_1}$$

for unconstrained firms. For constrained firms, $\mu_i > q_i/Q_1$. Therefore, in a financially unconstrained economy,

$$\mu^u = \sum_{a_i \geq \tilde{a}\left(p^u, Q_1^u\right)} \mu_i^u = \sum_{a_i \geq \tilde{a}\left(p^u, Q_1^u\right)} \frac{q_i}{Q_1} Q_1 = 1.$$

In the constrained economy, however,

$$\mu^u = \sum_{NE} \mu_i^u + \sum_{UI} \mu_i^u + \sum_{CI} \mu_i^u > \sum_{NE} \frac{q_i}{Q_1} + \sum_{UI} \frac{q_i}{Q_1} + \sum_{CI} \frac{q_i}{Q_1} = 1.$$

Therefore, the sum of the markups for all the marginal units produced by all firms is lower in the financially unconstrained equilibrium. ■

Proof. Part B. In the low wealth economy all firms produce the same quantity. Thus the unweighted and standard Lerner indexes coincide. ■

Proposition 9. *The concentration ratio H_j is strictly higher in the financially unconstrained economy for $j \leq n$.*

Proof. It follows immediately from Proposition 6. ■

Figure 1: Equilibrium in the financially unconstrained economy

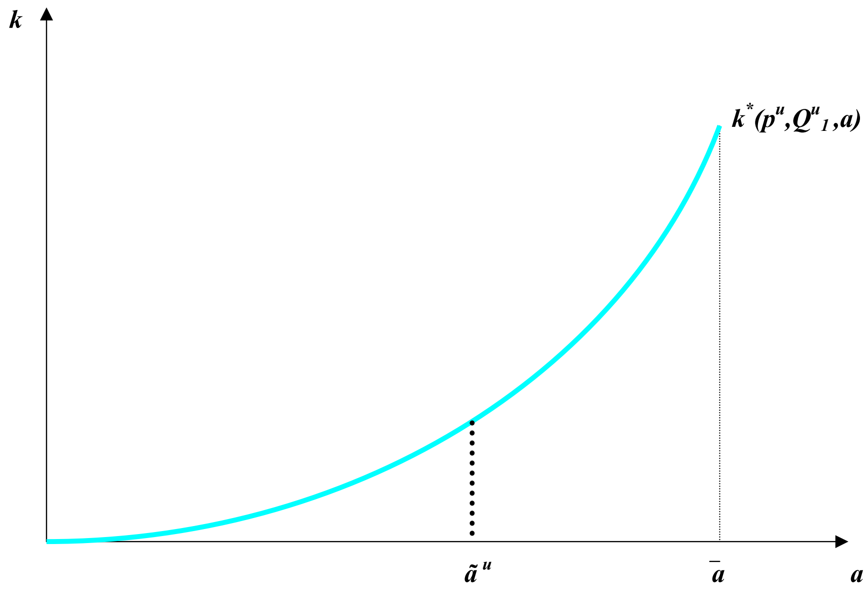

Figure 2: Equilibrium in the financially constrained economy: High wealth

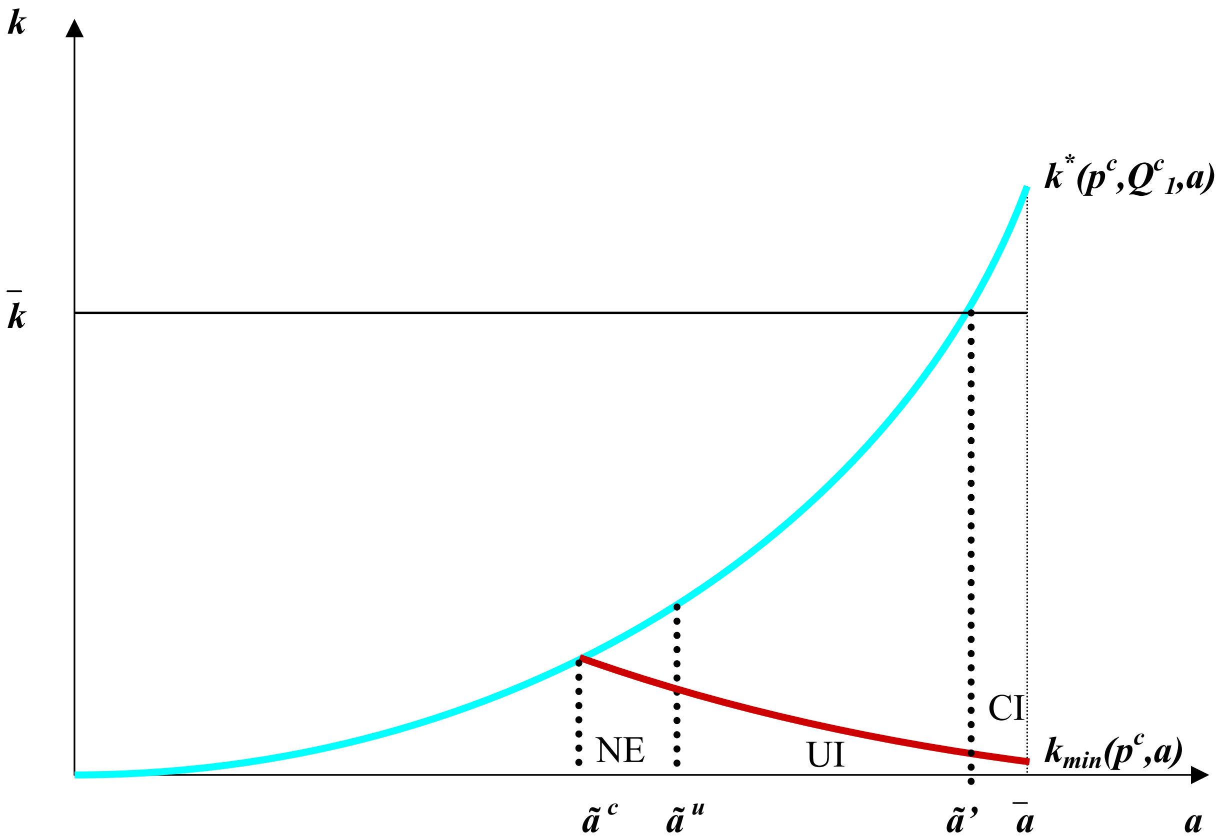

Figure 3: Equilibrium in the financially constrained economy: Low wealth