(In)efficient Investment in Financial Market Infrastructure: The Role of Governance Structures

Thorsten V. Koepl

Department of Economics
Queen’s University

August 27, 2012

Abstract
This paper establishes a link between investment in financial market infrastructure and its governance structure as either a user-oriented or a for-profit organization. The infrastructure is used to ensure that capital can be allocated through financial trades. I show that only a user-oriented infrastructure that allocates costs appropriately across different user groups can achieve efficiency. If costs can not be allocated appropriately, underinvestment occurs even if the infrastructure is user-oriented. Hence, user-orientation is a necessary, but not a sufficient condition for efficiency. To the contrary, an infrastructure that operates for-profit and is free to set its access fees will always lead to overinvestment. Moreover, requiring free access to the infrastructure does not necessarily lead to efficiency and overinvestment can still occur. These results should caution policy makers and regulators in two ways. First, efficiency does not necessarily result from regulating access to a for-profit infrastructure. Second, one can not interpret measures such as trading volume as an accurate indicator of efficiency in financial market infrastructure.

Keywords: Financial Market Infrastructure, For-Profit, User-Oriented, Limited Enforcement, Over- and Underinvestment

JEL Classification: G20, G28, G34

*I would like to thank the WWZ of the University of Basel for their hospitality where some of this research was done. I acknowledge financial support from SSHRC SRG 410-2010-0610.
1 Introduction

Financial markets rely heavily on infrastructure such as exchanges, clearing and payment and settlement systems to enable trade among its participants. Without such institutions and a general legal framework to monitor and enforce the obligations arising from trade, it is hard to imagine efficient financial transactions. Until recently, this infrastructure used to be either set up as “user-oriented” or as a publicly funded entity. Central to this approach is that the infrastructure is designed and run in the best interest of the general public or the direct users of financial markets. But nowadays, many elements of this infrastructure have either been turned form user-oriented to for-profit organizations, or have seen the emergence of privately funded and operated alternatives.

In this paper, I set out to understand the consequences of these developments for efficiency in financial markets. Are governance structures important for efficient investment into infrastructure? And if so, what are necessary and sufficient conditions to have an efficient level at which such infrastructure is provided? Interestingly, I find that both governance structures do not necessarily lead to efficiency. A user-oriented infrastructure tends to suffer from a holdup problem, where some participants have no incentive to participate in its financing as they expect not to recover their investment costs. Hence, systems with such a governance structure tend to suffer from too little investment. With a for-profit infrastructure, the problem is rather the reverse; whenever the owners also profit directly from using the infrastructure, they may invest too much into the infrastructure.

To understand these insights, I identify a user-oriented infrastructure as one that benefits the average user, while a for-profit infrastructure operates for the benefit of its owners which are a subset of the users. Then, I look at a framework that is close to Koeppl, Monnet and Quintin (2011) where limited enforcement hinders the exchange of capital among people or, equivalently, investors. Here, (limited) enforcement corresponds to a well-known concept in economics and is understood as the (in)ability to force a counterparty to satisfy the obligations arising from financial trades. Even though enforcement is a priori limited, one can invest into a costly technology to increase enforcement.

---

1This is why I refrain from using the term “user-owned” which I will discuss in more detail in the next section. Most infrastructure that has demutualized has some share of equity – if not a dominant one – held by their major users (see for example Pirrong (2000) on evidence for stock exchanges).

2Classic references on this concept include Sappington (1983) and Kocherlakota (1996).

3This approach was first used by Koeppl (2007).
In my framework, what then matters for efficiency is not only the exchange of capital, but also the costs for the necessary investment into enforcement that enables such exchange. Moreover, it is also important how these costs are distributed across the participants in the exchange of capital. I interpret this investment as infrastructure investment and view how the costs are levied across trading participants as a fee structure for accessing the infrastructure. Efficiency requires that one maximizes the average user’s benefit which automatically implies that user-orientation is necessary for efficiency. But more importantly, efficiency also requires a particular fee structure where some participants have to pay more than others.

This fee structure is crucial for understanding why the inefficiencies arise. First, with user-orientation, once the investment has occurred, it is beneficial to increase the fees on some participants to redistribute the gains from financial exchange further. Then, a holdup problem arises, where some investors have no incentive to participate in setting up the infrastructure in the first place. Second, with for-profit orientation, owners always have the incentives to extract surplus from charging high access fees. This raises the marginal benefit from the infrastructure for the owners above the social one causing too much investment. Even more interesting, this overinvestment result can survive, despite regulating access fees to be zero. What is the intuition for this result? Even though owners now pay fully for the infrastructure, their private benefit can still exceed the social one, which is the relevant measure in a user-oriented infrastructure. I provide some examples, where this is case precisely, whenever some users are charged fees above cost to subsidize access for other users.

In conclusion, my work has shown that an efficient infrastructure is necessarily user-oriented, but also needs to recover its costs with a particular distribution of its costs via user access fees. If this latter requirement cannot be fulfilled, inefficiency will result independent of the governance structure. In particular, whenever these fees involve a subsidy for one group of users to another, there is an incentive for some users to run the infrastructure for profit, even if they are regulated to charge zero access fees to users. To recover the set-up costs, the owners have then an incentive to overinvest into the infrastructure. The consequences of this result for public policy are twofold. First, it is not clear that regulating fees in for-profit financial market infrastructure will restore efficiency. Second, large transactions volumes enabled through an infrastructure when it is run for profit is not necessarily an appropriate measure to judge efficiency.

Of course there are many other aspects that determine the efficiency of financial markets.
infrastructure such as competition, innovation or transaction costs and details of the fee structure. But recently, the Committee on Payments and Settlement Systems (CPSS (2010)) has recognized the importance of governance structures for financial market infrastructures.\footnote{Hart and Moore (1996) were among the first interested in the question whether demutualization of financial exchanges matters for financial markets. Furthermore, governance structure might be an important factor that can hinder the efficient consolidation of fragmented infrastructure. For example, Pirrong (2007) points out that there can be incentives to underinvest when merging different vertical layers of infrastructure such as exchanges and clearing and settlement systems.}

I have confirmed here that governance structure indeed matters and, most remarkably, have found that different governance structures can lead to both, too much as well as too little investment.\footnote{The main elements that lead to my under- and overinvestment results have appeared in the economics literature before. One example is the incomplete contracting literature that gets too little or too much investment as an efficient reaction to holdup problems. Another example is the literature on patents that can generate overinvestment into R&D, through races between profit-oriented competitors where private benefits outweigh social ones.} An open question remains whether there is empirical support for these results. Serifsoy and Tyrell (2006) have looked into the issue somewhat and found evidence that exchanges have increased investment into non-core business areas after demutualization. Otherwise, empirical studies have mostly focused on the role of governance for operational efficiency such as transaction costs, but not on the investment side.\footnote{See for example Krishnamurthy, et al. (2003) or Serifsoy (2007).} Here, results point to infrastructure being more efficiently operated after demutualization which provides somewhat a counterpoint to my overinvestment result.

## 2 Efficient Infrastructure Investment and Financing

### 2.1 The Environment

The economy consist of two groups of people each of measure 1. Each member of the groups is endowed with an initial endowment of capital $a_1 = a - \delta$ and $a_2 = a + \delta$, respectively with $\delta \in (0, a)$. For convenience, we call the two groups of people, the poor and the rich. People have also access to a decreasing returns to scale technology that transforms capital into a consumption good and is for simplicity described by the production function $k^\alpha$, where $\alpha \in (0, 1)$. Their preferences are defined over the consumption good and described by a standard utility function that is assumed to be strictly decreasing and strictly concave.
Since production exhibits decreasing returns to scale and endowments differ across people, there is an incentive to trade capital before production. I assume, however, that there is limited enforcement. No obligations from trading capital can be enforced, unless some resources are invested into an enforcement technology. People can default on any transfer they owe after production, in which case they incur a penalty \( \eta \in [0, \infty) \) in consumption equivalent units. Establishing this penalty requires an (economy-wide) investment of capital equal to \( g(\eta) \) prior to production, where \( g \) is a strictly convex function with \( g(0) = g'(0) = 0 \). In this sense, enforcement is non-rivalrous, as its use by some person does not preclude its use by some other person. It is costly, however, as one has to invest capital that is not any longer available for producing the consumption good. But there are benefits from this technology, as it enables a better allocation of capital.\(^7\)

I interpret this setting in a very specific way. Financial exchange requires institutions that are costly to build and operate. These are not only the entire legal system that enables the enforcement of contracts. They also comprise physical infrastructure that is vital for the functioning of financial markets. Among these are exchanges, clearinghouses and payment and settlement systems which offer technologies to carry out and enforce financial transactions. Participants in this infrastructure differ greatly in their financial resources as well as in the function they take on. The heterogeneous groups in my framework can then be thought of as differentiating between larger, institutional investors or lenders and smaller retail investors or debtors. When looking at efficiency, I will use a high level of abstraction throughout, referring to a technological choice for achieving enforcement in the economy. Whenever I look at a decentralized equilibrium with financial markets transactions, I will alter my interpretation to one of financing an infrastructure that enables competitive exchange on financial markets. This still leaves open how I introduce the notion of governance. I look at two different governance structures. A user-oriented structure is designed to benefit all users of the infrastructure whereas an infrastructure that is run for-profit maximizes the returns for its owners.\(^8\) I will start out characterizing efficiency by looking at a planner that maximizes the welfare of an

\(^7\)For technical details on the material presented in this section, please refer Koeppl, Monnet and Quintin (2011) who analyze the efficient enforcement choice extensively. One could include other elements into the framework such as some fixed cost in the sense that \( \lim_{\eta \to 0} g(\eta) \geq 0 \) or other specifications of the penalty. None of the results would be materially affected.

\(^8\)Note that I deliberately do not use the term “user-owned”, as ownership and the objective of the infrastructure are different issues. For example, the general public could own the infrastructure, but it is designed to benefit its users. Similarly, some users can own the infrastructure, but run it for their own profit.
average user. Hence, the efficient allocation arises from considering a user-oriented infrastructure. Later on, in Section 4, I will analyze how the allocation changes if the planner maximizes only the welfare of one group – the rich. Thus, in my interpretation, this corresponds to a for-profit infrastructure where the rich can be viewed as the owners.

Finally, whenever I look at an equilibrium in financial markets, I have to address the issue who covers the costs for setting up the infrastructure. Since I do not allow for outside funding, only the users can contribute and I interpret these contributions as access fees. Note that I do not place any restrictions on these fees a priori. I will choose them in such a fashion that a particular equilibrium will be realized in financial markets. For example, when I analyze a for-profit infrastructure, the rich who own this infrastructure can extract surplus from the other users via such fees to the extent that they appropriate all benefits. After this, I look at the opposite extreme where the for-profit infrastructure cannot levy any access fees, so that the owners need to finance the investment themselves. In what follows, proofs of all the results are either derived in the text or relegated to the appendix.

2.2 Efficient Allocations

I will first describe the efficient allocation of resources in this economy. Assume a social welfare function that puts equal weight on the two groups in the population. There is a unique efficient allocation for the economy given by an enforcement level $\eta^*$ and capital levels $(k_1^*, k_2^*)$. This allocation solves the problem

$$
\max_{\eta, k_1, k_2, t_1, t_2} k_1^\alpha - t_1 + k_2^\alpha - t_2 
$$

subject to

$$
k_1 + k_2 + g(\eta) \leq 2a \quad (2.2)
$$

$$
k_1^\alpha - t_1 \geq a_i^\alpha \text{ for } i = 1, 2 \quad (2.3)
$$

$$
t_1 \leq \eta \text{ for } i = 1, 2 \quad (2.4)
$$

$$
t_1 + t_2 = 0. \quad (2.5)
$$

Note that it is equivalent here to maximize total output, consumption or average utility. The reason is that we allow for consumption transfers and the restrictions on such transfers and participation ensure that maximal output also yields maximal utility.
The first constraint is a resource constraint that takes into account that the endowments are either used in the production of the consumption good or in the enforcement technology. The second constraint takes into account that each group (after transfers) receives at least the value of staying in autarky. This ensures the participation of the two groups in the efficient allocation. The last two constraints spell out that only transfer up to the level of $\eta$ can be enforced and that these transfers have to sum to zero. Note that transfers $t_i$ can be positive or negative.

The optimal allocation has the feature that enforcement is strictly positive ($\eta^* > 0$) and is described by\(^{10}\)

$$g'(\eta^*) = \frac{1}{\alpha k_2^{\alpha-1}} - \frac{1}{\alpha k_1^{\alpha-1}}. \quad (2.6)$$

The interpretation of this condition is straightforward. For efficiency, one needs to equate the marginal cost of the enforcement technology with its marginal benefit which arises from a better capital allocation as expressed by reducing the wedge between marginal products of capital. Finally, the optimal capital levels for production are pinned down by the resource constraints and the fact that the planner will redistribute as much capital as possible. This implies that the participation constraint of the rich is binding. The efficient allocation is then fully described by two more equations,

$$k_1^{*\alpha} + k_2^{*\alpha} + g(\eta^*) = 2a \quad (2.7)$$
$$k_2^{*\alpha} + \eta^* = (a + \delta)^\alpha. \quad (2.8)$$

An important feature of optimal capital input is that they are never equated in an efficient allocations. Saving some of the costs on enforcement is of first order, while reducing the wedge between marginal products is of second-order. The transfers of capital are then constrained by the enforcement level through keeping the rich participating, or

$$a_1^\alpha < k_1^{*\alpha} - \eta^* < k_2^{*\alpha} + \eta^* = a_2^\alpha. \quad (2.9)$$

The next proposition summarizes the efficient allocation.

**Proposition 2.1.** The efficient level of enforcement is positive ($\eta^* > 0$), but the marginal

\(^{10}\)The first-order condition is expressed in units of capital input and, thus, leads to an inverse Euler equation akin to the one in the literature on optimal Mirrleesian taxation (see e.g. Kocherlakota (2005)).
products of capital at the efficient allocation are not equalized \((k_1^* < k_2^*)\). The planner allocates all surplus from investing in the enforcement technology to the poor.

2.3 Efficient Governance

I turn next to the question who should pay for the investment into an efficient market infrastructure. Denote the contribution of group \(i\) for financing \(g(\eta)\) by \(\kappa_i\). Consider a decentralized financial markets equilibrium with level \(\eta^*\) for enforcement. I assume that people can be excluded from trading on the market, if they do not pay their contribution \(\kappa_i\). One can understand this assumption as follows. A planner invests in the enforcement technology and raises the costs for it from people through tax contributions. Since I require that everyone will benefit from the planner’s investment taking into account taxes, there is an incentive for everyone to contribute. The planner, then, needs to only avoid a free-riding problem, where individuals do not pay the fees, but still trade on the market. Hence, the fee can be understood as paying a voluntary access fee to the technology or capital market as I will elaborate further below.\(^{11}\)

Given the infrastructure, claims can be enforced in the market up to a level of payment equal to \(\eta^*\). This implies a borrowing constraint given by

\[
(k_i - (a_i - \kappa_i)) R \leq \eta^*. \tag{2.10}
\]

Here, \(k_i - (a_i - \kappa_i)\) is the borrowing level by each group – taking into account that contributions to finance the infrastructure must come from endowments – and \(R\) is the equilibrium interest rate. For the optimal capital allocation to be the outcome in the financial markets equilibrium, it can be shown that contributions need to be set equal to

\[
\kappa_i = a_i - (k_i^* - \frac{t_i^*}{R}). \tag{2.11}
\]

Here, the equilibrium interest rate is determined by the rich people, i.e. \(R = \alpha k_2^{\alpha - 1}\). The intuition is that the rich people are lenders, hence, unconstrained and therefore need to be indifferent between lending and not. Taking into account that \(t_1^* = -t_2^* = \eta^*\), one can verify

\(^{11}\)Note that the assumption is then not inconsistent with the notion of limited enforcement.
that the contributions $\kappa_i$ are sufficient to finance $g(\eta^*)$.

I interpret this result as follows. The technology is used to clear and settle trades and, thus, enables trading in the first place. The contributions $\kappa_i$ can be seen as access fees to the infrastructure. The result then says, that it is efficient to run the infrastructure as a user-oriented one that recovers its costs through access fees. The costs levied on the individual participants respect their incentives to use the infrastructure – in other words, they are individually rational for people wanting access to the market. Importantly, however, in this model, it can be the case that $\kappa_i < 0$, i.e., that some investors receive a subsidy to trade on financial markets.\footnote{In general, this depends on the relative curvature of the production function and the cost function. For example with $\alpha = 1/2$ and $g(\eta) = \eta^2$ people are just indifferent between any level of $\eta$. Increasing curvature on either function causes $\kappa_i$ to be negative for the poor.} This shows that an efficient market infrastructure can require explicit subsidies to some participants. In the model, the rich might finance all of the investment in enforcement plus a capital subsidy to the poor. Through competitive trading of capital in financial markets, they however recoup all their expenses. In conclusion, the optimal governance structure of the infrastructure is user-oriented where not all users necessarily pay uniform or even positive fees.\footnote{Of course, the efficient allocation can be also achieved through a combination of ex-ante contributions to set up the infrastructure and taxes on the return on lending. These would correspond to initial fees for financing the infrastructure, as well as ex-post access fees for trading.}

3 Underinvestment in User-oriented Infrastructure

In the efficient allocation, the rich get reimbursed for their contribution of capital after production, so that they are exactly equally well off as in autarky. This requires commitment on the side of the planner. I assume now that the planner cannot guarantee to reimburse people for their initial contribution to finance the enforcement technology.\footnote{Importantly, I still maintain the assumption that ex-post – i.e. after capital gets allocated across people for production – the planner cannot seize the returns. If he could do so, the only feasible allocation would be autarky, so that no investment in infrastructure would take place.}

This leads to a two-step problem that can be treated as two subperiods. In the first subperiod, the planner invests into enforcement, $\eta$, and receives contributions of capital, $\kappa_1$ and $\kappa_2$, to finance it, where $\kappa_1 + \kappa_2 = g(\eta)$. In the second subperiod, he allocates capital and its return across people. When doing so, he needs to respect that people indeed have an incentive to
participate. This captures the notion that people cannot be forced to engage in the exchange of capital once the investment into enforcement has taken place. To be clear, what the planner cannot do is to commit to reimburse people for their initial contributions $\kappa_i$. Given the decisions $(\eta, \kappa_1, \kappa_2)$ in the first period, the planner maximizes total output in the second period, or\textsuperscript{15}

$$
\max_{k_i, t_i} \sum_{i=1}^{2} k_i^\alpha - t_i \\
\text{subject to}
$$

\begin{align*}
& k_1 + k_2 = 2a - g(\eta) \\
& t_1 + t_2 \geq 0 \\
& t_i \leq \eta \text{ for } i = 1, 2 \\
& k_i^\alpha - t_i \geq (a_i - \kappa_i)^\alpha \text{ for } i = 1, 2.
\end{align*}

In this problem, the participation constraints are relaxed for any $\kappa_i > 0$. It is still the case that for any optimal choice of enforcement in the first subperiod, the enforcement constraint must be binding for the poor agent, i.e. $t_1 = \eta$. If not, one could simply lower $\eta$ to begin with which increases utility. Furthermore, it must be the case that for any given level $\eta$ we have $k_2 \geq k_1$. This implies that the ex-post participation constraint for the rich must be binding, or

$$
k_2^\alpha + \eta = (a + \delta - \kappa_2)^\alpha \tag{3.6}
$$

which is less than the value of autarky for $\kappa_2 > 0$.

As a consequence, rich agents will have no incentives to contribute to the enforcement technology at all in the first subperiod. The poor agents will have to finance the investment in enforcement alone and the planner faces a \textit{financing constraint}

$$
g(\eta) \leq (a - \delta). \tag{3.7}
$$

The next result shows that this constraint will be binding as long as the initial endowment of capital is sufficiently unequally distributed, i.e., $\delta$ is sufficiently close to $a$. The intuition for this result is straightforward. The higher the dispersion in endowments, the higher is the

\textsuperscript{15}Again, it is equivalent here to maximize total output or the average utility across people.
optimal investment in enforcement. Hence, increasing $\delta$ tightens the constraint from both sides. This can be summarized below.

**Proposition 3.1.** Suppose the planner cannot commit to reimburse capital contributions to finance enforcement. Then he faces a financing constraint, as only the poor have an incentive to finance the investment in enforcement.

Furthermore, the planner chooses an enforcement level $\eta$ that is strictly lower than the socially optimal one whenever inequality is large enough, i.e. $\eta < \eta^*$ for all $\delta \in (\bar{\delta}, a]$.

How does this planning problem relate to a decentralized economy where capital is exchanged through competitive financial markets and the planner can tax financial returns? One can think again of an infrastructure being set up initially among its users where individual users are free to contribute. Once the infrastructure has been put into place, trading on competitive financial markets takes place and taxes are announced on returns from financial transactions. It is clear, that with appropriately set taxes, a planner can achieve his preferred allocation also in a competitive equilibrium where capital is exchanged through financial markets. Proposition 3.1 establishes then that not all users will have an incentive to contribute to financing the infrastructure. The rich anticipate that they will never recoup their share of financing, as they face a tax that makes them indifferent between (i) lending out their capital and (ii) not participating in the market. As their contribution $\kappa_i$ is sunk, they are held up by the planner when facing taxes on their return from lending which prevents them from recovering their initial investment cost.

This result is intriguing, but not too surprising. If users decide to set up an infrastructure, incentives to do so might vary. In my model, there is a group of investors or lenders that is crucial to the market, but has no incentive to put the necessary market infrastructure into place. The reason is a holdup problem. Once the infrastructure investment is sunk for these investors, the gains from financial trades might not be sufficient for this group to recover the initial investment. Hence, they will not participate in creating this infrastructure. Recall that the infrastructure is run as to maximize the utility across all users; i.e. it is user-oriented. There is then an incentive to redistribute benefits from financial exchange to a particular group of users, the poor. Here, the rich investors anticipate high fees – this is how I interpret the taxes – when they use the infrastructure. Of course, the remaining investors can still
set-up the infrastructure and pay for it themselves. However, since they bear all the costs and are limited in the fees they can charge other users, they will underinvest.

Again, here it is important to understand the difference between corporate governance and ownership. The poor finance the infrastructure entirely, but it is not necessary that they own the infrastructure. Indeed, even if the poor were the sole owners, ran the infrastructure for-profit and charged access fees from non-owners, the results would not change, as a user-oriented infrastructure already extracts as much trading surplus from the rich as possible. The bottomline to take away from this section is then that a user-oriented infrastructure needs still to ensure that all users participate in financing the infrastructure in order to achieve efficient investment. But this may be difficult to ensure, if access fees are expected to vary across users ex-post.

4 Overinvestment in For-profit Infrastructure

4.1 No Fee Regulation

I consider now a different problem, where the planner does not maximize overall output, but only the consumption of the rich people. The new problem is then given by

$$\max_{\eta, k_1, k_2, t_1, t_2} \quad k_2^\alpha - t_2$$
subject to

$$k_1 + k_2 + g(\eta) \leq 2a$$

$$k_1^\alpha - t_1 \geq (a - \delta)^\alpha$$

$$t_i \leq \eta \text{ for } i = 1, 2$$

$$t_1 + t_2 = 0.$$

There are two changes with respect to the planner’s problem for an efficient allocation. The objective function only considers consumption for people with high endowment; and consequently, there is no participation constraint for this group of people. For any choice of $\eta$, it is optimal to transfers as much output to the rich as possible, or $t_2 = -\eta$, while giving as little capital as possible to the poor. This implies that the participation constraint of the
Poor binds, or
\[ k_1^\alpha - \eta = (a - \delta)^\alpha. \]  
(4.6)

Note that the planner’s allocation extracts all the surplus from the poor and allocates it to the rich. Hence, at least implicitly, only the poor pay for the enforcement technology. For the efficient level of enforcement, \( \eta^* \), we have then that \( k_1 < k_1^* \) and, by feasibility, that \( k_2 > k_2^* \). Since the return on enforcement has increased, this implies immediately that the planner would like to invest more than the efficient amount in enforcement.\(^{16}\)

**Proposition 4.1.** *If the planner maximizes the welfare of the rich, the investment in enforcement exceeds the socially optimal level, i.e. \( \bar{\eta} > \eta^* \).*

One can again interpret this result in the context of financial market infrastructure. Now, the infrastructure is run for profit, as it does not benefit the average user, but only the rich people. Consider now again an equilibrium where trading takes place on financial markets after the infrastructure has been set up. The rich investors pay upfront for the investment, but recover some of it by charging an access fee, \( \phi \), from non-owners for trading on financial markets. This implies that trading begins with endowments equal to \( (a - \delta) - \phi \) and \( (a + \delta) - (g(\eta) - \phi) \). To achieve the optimal allocation for the rich, the fee \( \phi \) for accessing markets has to be set equal to
\[ \phi = \bar{k}_1 - (a - \delta) - \frac{\bar{\eta}}{R}, \]  
(4.7)

where \( R = a\bar{k}_2^{\alpha - 1} \). Still, the owners of the infrastructure also benefit directly through financial transactions that they carry out using this infrastructure. Hence, without regulating access costs, it follows that a for-profit infrastructure can have too much investment.

### 4.2 Fee Regulation

One might think that the last result arises only because the rich levy high access fees on the poor. However, I will show next that the rich can have an incentive to overinvest in financial markets, *even if* they have to bear the full cost for this investment. In this case, the rich

\(^{16}\text{It is easy to verify that maximizing the poor’s consumption instead of the rich’s consumption would yield the efficient allocation. Hence, profit-orientation in this case would yield efficient investment into infrastructure.}\)
will have to recover all of their investment indirectly by trading on financial markets. To do so, I express the problem for the rich as a Ramsey equilibrium. Given any \( \eta \), there will be a competitive equilibrium for trading capital. Taking this equilibrium as given, the rich (as a group) decide on a level of enforcement with every member paying the cost \( g(\eta) \). When trading subsequently in financial markets, the poor start then out with endowment \( a - \delta \) and the rich with endowment \( (a + \delta) - g(\eta) \).

This captures a situation, where fees are regulated so that all users need to have access to financial markets at zero cost. This certainly is an extreme type of regulation, but I chose to analyze this scenario in order to demonstrate that my overinvestment result does not depend on surplus being extracted through user fees. Instead, having owners also being users can imply high enough private returns from running the infrastructure that there is too much investment.

As before, it can never be optimal for the rich to end up as borrowers of capital on the market. Hence, they are lenders, thus unconstrained, and their input of capital pins down the interest rate at which the poor can borrow up to the level of enforcement. Hence, given \( \eta \), the equilibrium \((k_1, k_2, R)\) is described by the following equations

\[
\begin{align*}
  k_1 + k_2 &= 2a - g(\eta) \quad (4.8) \\
  [k_1 - (a - \delta)] R &= \eta \quad (4.9) \\
  R &= \alpha k_2^{\alpha-1}. \quad (4.10)
\end{align*}
\]

The first equation is market clearing, while the second one formalizes the endogenous borrowing constraint for the poor with the equilibrium interest rate \( R \) pinned down by the marginal product of capital for the rich – the third equation. Using the last two equations one can formalize a standard implementability condition. The Ramsey problem is then given by

\[
\max_{k_2, \eta} k_2^\alpha + \eta \quad (4.11)
\]

subject to

\[
\begin{align*}
  \frac{\eta}{\alpha} k_2^{1-\alpha} + k_2 &= (a + \delta) - g(\eta) \quad (4.12) \\
  k_2 &\geq a - \frac{g(\eta)}{2}, \quad (4.13)
\end{align*}
\]

where the first constraint is the budget constraint of the rich taking into account market
clearing and the second constraint ensures that the rich indeed lend to the poor \((k_2 \geq k_1)\).

As is generally the case in Ramsey-type problems, the constraint set is non-convex. Since the objective function is increasing in both arguments, however, any solution must satisfy the budget constraint for the rich. Hence, the constraint must be binding and the solution – at an interior point – is described by the following necessary conditions\(^{17}\)

\[
g'(\eta) = \frac{\eta}{\alpha k_2} \left( 1 - \frac{\alpha}{\alpha k_2^{\alpha-1}} \right)
\]

\[
g(\eta) = (a + \delta) - \frac{\eta}{\alpha k_2^{1-\alpha}} - k_2.
\]

This implies that investment for the rich into their own production and into enforcement are substitutes, since the latter enables more lending of capital on financial markets. Unfortunately, it is not possible to characterize the solution further analytically. I therefore resort to some numerical examples.

These examples use the functional forms \(k^\alpha\) for production and \(\eta^\xi\) for the cost function, where \(\xi > 1\). I also normalize endowments to the value of \(a = 1\) and set \(\delta = 0.5\). There are three different examples corresponding to different curvatures of the production and the cost function.\(^{18}\) The examples are designed to show the relationship between distributing costs across people in user-oriented infrastructure and the incentives to run the infrastructure for-profit for some group. Quite intuitively, the incentives to run it for profit occur when most of the costs need to be covered by a particular group of investors – or, even more telling, when access to the infrastructure is in fact subsidized for some users. In this case, some users have to pay less when they own the infrastructure, pay for it and operate it on the same scale, since they can avoid paying the subsidy. They have then a larger private return and, thus, an incentive to invest more than is socially efficient.

I report the results of these examples in various graphs, which plot all variables as a function of enforcement \(\eta\). The first row depicts the value of an efficient allocation and the ex-ante (net) contributions for the poor for different, given levels of enforcement \(\eta\). The planner

\(^{17}\)In non-convex problems, a candidate for a solution that satisfies the Kuhn-Tucker conditions is not necessarily a solution to this problem. Here it could also be the case, that the solution is at the boundary of the constraint set, as I verify in some numerical solution below. Then, it is not necessarily the case that the first-order condition holds with equality.

\(^{18}\)To understand these examples better, note that the two sides for the first-order necessary condition on \(\eta\) move in opposite directions only if both \(\xi > <2\) and \(\alpha < (>1)/2\).
will choose the level of $\eta$ that achieves the maximum value $V_{\text{eff}}$ in the left graph. The right graph shows the contribution for the poor to finance a level $\eta$, when the planner maximizes output for this given level of enforcement. If the contributions are negative, the poor receive a subsidy to access the infrastructure – or, equivalently, the rich pay more than the costs $g(\eta)$. The second row shows the rich people’s surplus over autarky for two different allocations, again for different levels of enforcement. Here, I calculate the difference between the value of a for-profit infrastructure without ($V_{\text{rich}}$) and with fee regulation ($V_{\text{Ram}}$), respectively, and autarky. If this value is positive, then the rich have an incentive to run the infrastructure for-profit in those cases. Note that for the Ramsey equilibrium, I only look at values where $k_2 \geq k_1$. Finally, for completeness, in the last row I just compare the level of borrowing in the efficient equilibrium ($b_{\text{opt}}$) and the Ramsey equilibrium ($b_{\text{Ram}}$).

The first example (Figure 1) is the base case where $\alpha = 0.5$ and where the cost function is quadratic ($\xi = 2$). Here, it turns out that ex-ante contributions for the poor are zero independent of $\eta$. This implies that in a user-oriented infrastructure all costs are borne by the rich investors. The middle row confirms Proposition 3.1 that there is an incentive to overinvest if there is no fee regulation. However, it is also clear that with regulated fees, the rich who own the infrastructure are just indifferent between any level of enforcement $\eta$. When they have to rely solely on the returns of lending out capital to recover their investment, they

Figure 1: Case: $\alpha = 0.5$ and $\xi = 2$
just break even for all levels of investment.

Figure 2 exhibits a case with more curvature relative to the baseline case in both, the production function and the enforcement cost function. Here, in a user-oriented infrastructure contributions of the poor are indeed negative; i.e., the poor receive a subsidy to participate in the financial market. As the left graph in the second row shows, the rich are now better off avoiding the transfer, if they invest into the infrastructure by themselves and run it for-profit. Furthermore, even when regulated that they cannot charge for access to the infrastructure, they want to invest more into the infrastructure (higher level of $\eta$) and borrowing in the Ramsey equilibrium exceeds the socially efficient level. In this sense, the capital market works too well compared to efficiency. Indeed, the Ramsey equilibrium is a corner solution, where the infrastructure works so well that the two marginal products of capital get equalized ($k_1 = k_2$).

The final set of graphs (Figure 3) shows the opposite case with less curvature on the two functions. Now, contributions for the poor are strictly positive to finance the user-oriented infrastructure. This implies that both the rich and the poor contribute a positive amount to financing the infrastructure. As the middle row of graphs suggests, there is no advantage for the rich anymore to run the infrastructure for-profit, if fees are regulated to grant free access.
The intuition is again the same. Given that one group has to cover all the costs of the infrastructure, the private benefit is below the social one. Indeed, in this case, a for-profit infrastructure would be never set up, since autarky ($\eta = 0$) dominates any other level of investment.

To summarize, governance structure is intertwined with the optimal financing of the infrastructure. If one cannot distribute the costs appropriately among the users, underinvestment may occur. Otherwise, private benefits can be large enough in a for-profit infrastructure, so that too much investment can be the outcome – even if access fees are tightly regulated.

5 Conclusion

In this paper, I have shown that governance structures matter for having efficient investment in financial markets infrastructure. Due to the model set-up, the features of the efficient allocation and how outcomes change with different governance structures are quite stark. Nonetheless, some powerful insights arise. Quite naturally, a user-oriented infrastructure is a

\[19\]

Of course, if a for-profit infrastructure can charge access fees, overinvest will still occur, even though this is not clearly visible in the graph.

18
necessary condition for efficiency, as for-profit motives usually do not reflect the interest of the average user. A particular feature of my framework is however, that once the infrastructure is set-up for-profit, there can be more investment than is efficient. This arises because the owners of the infrastructure can conduct more financial transactions using this infrastructure than would be efficient taking into account the costs required to build the infrastructure. In a sense, the private returns from the infrastructure exceed the social ones and financial markets are working too well for the average user. As a policy conclusion, one should not allow a for-profit infrastructure that is operated by a subset of its users. Moreover, I have shown, that such a situation can even arise when no explicit fees are charged for using the infrastructure. Hence, standard measures for how well a particular infrastructure works, such as low fees or high transaction volumes, might not be good indicators for judging efficiency.

For full efficiency, a user-oriented infrastructure requires also a proper distribution of its costs across users. If this is not possible, I have shown that underinvestment can occur. Here, an important limitation of my framework is that I am looking at a static environment. This renders a distinction between initial investment, its costs and its financing and the fees required to operate the infrastructure and amortize its set-up costs all but impossible. More generally, other considerations such as efficient pricing or providing incentives to innovate are important. Notwithstanding, linking efficient infrastructure investment for financial markets to governance structures is a novel and intriguing point of view.

It is also possible to put my analysis into a broader context, even though one needs to be cautious given the very specific nature of my model. Recent events linked to the financial crisis have sparked a discussion on what level financial markets should operate, and claims have appeared that financial markets in many countries work “too well”. All modern economies rely on financial markets to allocate capital and such markets need institutions that ensure that the returns on lending out capital are enforced. The higher these investments are, the better financial markets work in the sense that the volume of lending is higher. US financial markets are usually described as leading other markets in borrowing and lending such as Canada or major European economies. Yet, all the countries that lag behind rely more on direct redistribution of resources – mostly through their tax system. Many commentators have pointed out that the US financial system supported a massive misallocation of financial resources. Even though the current channel in my model works quite differently, it might be possible to use my approach to support such views.
This demonstrates that redistribution and investment in the efficiency of financial markets are intimately linked. Having only well-functioning financial markets is not optimal without the proper level of redistribution – in this sense, they complement each other. More importantly, private incentives to invest in the efficiency of capital markets can diverge from social ones. This would call for regulating efforts to increase borrowing and lending in financial markets – without resorting to an argument that is based on financial stability.

Appendix

Proof of Proposition 3.1

Proof. The first part of the proposition has been already established in the text. The RHS of the financing constraint is decreasing in $\delta$. Since $g(0) = g'(0) = 0$ and $g$ is convex, we have that $\eta^* > 0$ for all levels of inequality $\delta > 0$. Furthermore, Koepl, Monnet and Quintin (2011) show that the optimal level of enforcement without the commitment problem $\eta^*$ is increasing in $\delta$. This implies that for $\delta \to a$, we eventually will violate the financing constraint for $\eta^*$. \hfill \Box

Proof of Proposition 4.1

Proof. The constraint set has not changed and the objective function is still strictly concave. This implies that there is a unique solution with $\eta > 0$ to this problem described by the first-order necessary condition

$$g'(\eta) = \frac{1}{\alpha k_2^{\alpha - 1}} - \frac{1}{\alpha k_1^{\alpha - 1}},$$

the feasibility constraint, but now the participation constraint for the poor.

Also, at $\eta^*$, we have that $k_2 > k_2^* > k_1^* > k_1$, and the marginal benefits of enforcement exceeds the LHS of the first-order condition. Furthermore, fix any $\eta < \eta^*$. Then, the best allocation given this level of $\eta$ would necessarily have a lower $k_1$ and a higher $k_2$, i.e. the marginal benefits increase further. Since $g$ is strictly convex, uniqueness of the solution implies then that the first-order condition is only satisfied for some unique level $\eta > \eta^*$. \hfill \Box
References


