

## 7. ROLE OF CONTRACT INCOMPLETENESS AND RELATIONSHIP SPECIFIC INVESTMENTS

This chapter analyses the research question from the perspective of the second specific objective, which states:

*To understand the role of Relationship-Specific Investments and contract incompleteness in the observed types of relationships between ESEs (autonomous hospitals) and purchasers.*

The previous chapter analysed the observed types of relationships between ESEs and their purchasers. It was shown that cooperation varied in these relationships and that a plausible explanation had to do with the ESEs' level of total revenues within a context of lack of exit options. This chapter will analyse how transaction costs contribute to shape the observed relationships that were described in chapter 6. Transaction costs, as shown in chapter 2, stem basically from contract incompleteness and RSI, both of which will be analysed in this chapter. These transaction costs will be analysed in terms of costs of contracting and the overall perception of contract incompleteness.

On the lines of these theoretical considerations, the propositions to be tested in this chapter, regarding contracts both with the SOH and ARSs, are:

- **Proposition 1:** There are relevant RSI in the relationships between ESEs and purchasers (SOH and ARSs).
- **Proposition 2:** The presence of RSI increases the transaction costs of relationships observed between ESEs and purchasers (SOH and ARSs).
- **Proposition 3:** Contract incompleteness pervades the relationships between ESEs and purchasers (SOH and ARSs).
- **Proposition 4:** Contract incompleteness is by itself a source of transaction costs of relationships observed between ESEs and purchasers (SOH and ARSs).
- **Proposition 5:** The transaction costs of these relationships, stemming both from RSI and contract incompleteness, result into long-term relationships

This chapter is organised in three sections: findings, discussion and conclusions. After a brief analysis of objective functions of the actors, the findings section is organised along

three main lines, following the framework proposed in figure 2.2.: RSI, contract incompleteness and the transaction costs derived from these two sources. Contract incompleteness is analysed in terms of two main areas: uncertainty and complexity of products, which have their own headings when the text is too extended. Each of these three aspects is analysed for the three contracts with the SOH (uninsured, PAB and ambulances) and for the contracts with ARSs.

## **Findings**

### 7.1.Objective functions of the actors

A key general point that has to be raised, before a more detailed analysis is done of transaction costs, is the objective function of the actors. As shown in chapter 5, although the description and analysis of objective functions were not explicitly considered during the design of this study, they emerged as a key element during data analysis and relevant quotes were found in the interview materials that shed light on this issue. As explained in chapter 5, the flexibility of case-study methods allows for this reshaping of research questions in the search for the best explanation for the findings, without losing the focus on the original research question and design.

It is assumed that, given that the SOH is a public organisation, its objective function is not to make profits. When asked about its agenda, the Secretary of Health answered:

*“...If you ask me what my concern is, I say: sustainability,”*

He also revealed a mix of aggregate welfare and social legitimacy (understood as responsiveness and acceptance of the hospital within the community) in the objective function, and some tradeoffs:

*“...The ultimate [objective] function for all of us is to guarantee (...) good-quality services and improvements in population’s health status. But at the individual and enterprise level, we define a trade-off between economic sustainability and social legitimacy.”*

Regarding ARSs, it is assumed that the for-profit ones behave as the standard theory of the firm, i.e., they have a clear and single-minded objective function: profits. However, regarding nonprofit ARSs, it could be assumed that they are more concerned about optimising good quality services or access to enrolees, but these objective functions were not explicitly ascertained. An indirect reference to objective functions of ARSs can be found in this quote from the manager of IIC:

*“...[ARSs] have (...) some particular features. ARSI [a non-profit ARS] is more serious and organised and gives the impression to be more committed to (...) have its patients well cared for, to keep them healthy. (...) ARSIII [a for-profit ARS] has been improving (...) of course, it is a private entity, it knows what she goes for, because it has its economic terms clearly stated.”*

Regarding ESEs, although they cannot privatise profits and although autonomous they keep their public nature, they were concerned for generating surpluses all the time. However, some managers, as will be seen below, openly manifested that social welfare was their objective function. Others though, expressed concern for sustainability of the ESE. It could also be argued that ESEs' objective function is to maximise revenues, because it allows them to increase output as a source of manager satisfaction. This perception rests on the following findings: 1) Managers were all the time boasting about the growth in revenues they achieved thanks to their efforts to attract ARS contracts, to increase the SOH's billing cap, or to attract other sources of revenues. 2) They mentioned increases in capacity and output as a proof of their competency. 3) Because they were not allowed to privatise surpluses, they had no interest in increasing profitability but rather in generating surpluses to invest them in more capacity expansions, which ultimately led to higher revenues. 4) Under the FFS payment, they strongly induced demand, whereas under FGPP they induced final activities but restricted intermediate activities (this will be clearly shown below and in chapter 8).

## 7.2.RSI regarding relationships with the SOH

### 7.2.1. Contracts for services to the uninsured

The most relevant RSI that was found in all ESEs was dedicated capacity. Given that the SOH is by far the largest purchaser of services of all ESEs, it is clear that they

depend on this contract to keep an adequate stream of revenues that avoids the costs of idle capacity. However, the implications of idle capacity are different from what could be expected from the literature on TCE for two main reasons: 1) investments in dedicated capacity were done before the PPS, and the *ex-post* bilateral monopoly situation was only strengthened by these decisions made in previous periods. 2) After the PPS, investments in dedicated capacity, particularly capital and buildings, were undertaken by the SOH, based on their estimation of local demand needs. Given the bilateral monopoly between SOH and ESE, it was unlikely that the SOH used its bargaining position vis à vis a party that had a dedicated-capacity type of RSI to hold it up and extract a quasi rent, if that decision would lead to a political stalemate that would backfire on the SOH itself.

In addition, ESEs' surpluses were not expected to be large enough to finance capacity expansions, but only to improve shorter-term performance. It did not mean that if an ESE accumulated surpluses large enough to increase capacity or acquire new equipment, it was not allowed to do so. For example, IIC built a new unit for administrative offices and IIB bought a gamma-chamber (see below). But given that most of the times surpluses were not reported to be this large, the SOH still kept the obligation to provide the necessary funds for capacity expansion, either buildings or technology equipment. Thus, it was found that the presence of dedicated-capacity RSI did not imply a risk of hold-up and did not entail transaction costs as expected from the literature from other sectors of the economy.

An illustrative case of dedicated capacity was found at IIB, where the manager decided to buy a gamma-chamber using accumulated surpluses. The SOH estimated it did not need such a service for the uninsured, because it was being provided somewhere else. However, the manager purchased the equipment on the expectation that he would sell the service to ARSs and other purchasers. This expectation proved optimistic, as reported by an SOH officer:

*“He has had the gamma-chamber for two years and we have not purchased from him [this service] because it is cheaper outside.”*

Regarding physical RSI, it was found that all contracts with the SOH were plagued by RSI, but these were not large enough to cause a risk of hold-up. In addition, ESEs

incurred the investment because there was no option but to invest, otherwise they would be unable to contract with the SOH. Examples of these physical RSI are certain specific software programs to issue bills to the SOH.

#### 7.2.2. Contracts for PAB services

The PAB contracts were reported by all ESE managers to very frequently involve some specific equipment that had no alternative uses with other clients. However, this equipment was not costly enough to imply a high risk of hold-up, and most of the times the ESE subcontracted those activities that required a very specific type of equipment, in order to avoid the acquisition of such equipment, as illustrated by the manager of IIF:

*“...sometimes those screening campaigns involve certain equipment. (...) we do not invest [in the equipment] if we cannot guarantee that it will be used beyond that [campaign]. So when the issue is too tied-in we rent it or we [sub-]contract personnel who provide the equipment.”*

This situation begs the question as to why the SOH did not contract these PAB services directly with the subcontractors the ESE uses. It could be argued that the ESE role was rather that of a middleman who did not add value to the process. In fact, the managers of IE, IIC and IIC recognised that this reasoning was true. But the explanation was that the specific equipment was not the only input, and that most of these ESEs had fixed labor costs that had to be paid for through the PAB contracts. This is illustrated by a quote from the manager of IJ:

*“...My major concern, if I decided not to contract for PAB services, would be with the environment specialists, because they perform a very specific activity.”*

ESE IK was a special case where the largest share of revenues were generated through PAB contracts, because of the political implications of the area in the armed conflict. In addition, it was the sole provider in the rural area where it was located. Accordingly, the manager of the ESE invested in PAB activities, training of personnel and equipment, without any consideration of further transaction costs. These were certainly human and dedicated capacity types of RSI, but the manager did not perceive the risk of hold up as a discouraging factor to invest in them.

### 7.2.3. Contracts for ambulances

It could be argued that an ambulance is an example of a dedicated-capacity type of RSI, because if the CRU decides not to contract a given ambulance with an ESE, it would be left idle. However, as ambulances were donated by the national government and some managers argued that the market for ambulance services in Bogota was highly competitive (i.e., low entry and exit barriers, unconcentrated demand and supply and high substitutability among suppliers), ambulances and their equipment were less of a RSI. The dedicated-capacity argument stems from the fact that the SOH takes advantage of the set of ambulances the ESEs have, to operate them jointly and exploit economies of scope. But, as the manager of IIB put it:

*“...we are exploring the possibility that, if at a given moment they do not contract one ambulance, we can offer it to the private sector at competitive prices...”*

In addition, ESEs had to invest in equipment for the ambulance that could also be considered RSI. However, this equipment was considered by all managers a minimum entry requirement, so ESEs had no other choice but to invest in them.

### 7.3.RSI regarding relationships with ARS

The most clearly identified type of RSI within ARS contracts was human RSI. Although ESE managers and SOH officers did not readily identify it as a RSI during interviews, after some hints they clearly recognised it was. Human RSI in this case referred to the investments an ESE had to incur to improve its capacity to manage the risk of a capitated population. It included a diagnostic process to determine the risk profile of the capitated population and the corresponding interventions aimed at reducing the short/medium-term incidence of events and hence obtain a surplus out of the capitation. These interventions are known as P&P, for (health) promotion and (disease) prevention. The rationale to consider it an RSI lies in the fact that once these investments have been made by the ESE, the ARS can substitute a different population for the one the ESE has invested in, and extract a quasi rent because the benefits of the P&P investment will accrue to the ARS instead of the ESE. According to TCE, this

dynamic would be expected to make the ESE more conservative to make the investment, thus leading to a suboptimal level of investment.

Regarding this RSI, two clearly different approaches were found among ESEs. 1) A group of ESEs, like IIC, IA, IJ, IIIA and IIB, considered it was unwise to invest in P&P, following the dynamic described in the previous paragraph. 2) Another group of ESEs, like IB, IE, IK and IIIC, considered it was wise to invest in P&P, no matter the interventions were paid for or not, because a large share of the population tends to stay in the area, so the benefits of the investment will be reaped in the future, no matter who the purchaser is.

On the side of the ESEs with a negative view of investment in P&P, they perceived an opportunistic behaviour by ARSs, i.e., substituting population once the investment had been done. Thus, they considered it was unwise to invest in reducing medical risk, particularly avoiding those interventions with medium and long-term effects, given the prospect of losing the investment. Thus, they preferred to invest only in those P&P activities that yielded short-term benefits, or to invest the minimum required by law or by the contracts. The manager of IJ said:

*“When a capitation contract is generating some surplus to the ESE due to rationality in use, [the ARS] reduces the capitation rate for the next contract, or they switch the population.” “...we do not incur [investments in P&P] because that would mean giving a healthy population to an ARS without knowing if it will later contract with me (...) and given that I do not have level II or III under my responsibility, then I do not invest in following up situations that would improve health care outcomes, but I restrict myself to curative care.”*

Some level II ESEs also perceived the risk to lose the RSI, like the manager of IA:

*“...It is a pity that there is no guarantee of a minimum continuity in enrolment in the same ARS. For example, if I sign a contract for three years, I commit to invest in P&P because I will reap the benefits. But if I will not reap them, why should I do that? And it costs, it is an effort [whose benefits] you don't see in the short term.”*

On the side of the ESEs with a more positive view of investing in P&P, four main reasons were mentioned to justify their behaviour: first, concerned for the ESE's image among its users, these managers believed they should not turn away patients because that would cause a negative image for the ESE. Second, they perceived a strategic advantage in investing in P&P because they used this investment as a way to enhance patient loyalty. Given that SS beneficiaries were allowed to choose their preferred ARS, securing patient loyalty to the ESE meant that ESEs strengthened their bargaining position vis à vis ARS, and ESEs themselves could even induce patients to select those ARS they preferred to deal with. This allowed these ESEs to avoid the opportunistic behaviour of ARS that substitute patients, and at the same time it helped them to weed out ARS with undesirable behaviours, specifically those that delayed payments excessively. Thus, they did not care about opportunistic behaviour by ARS because they considered these opportunistic ARS would soon be out of their client portfolio. A third reason they gave to invest in P&P was that given that the population tends to remain in the area, and given the expectation to achieve a high level of loyalty, the medium and long-term effects of the investment will be received by the ESE. The fourth reason they mentioned was that they were a social enterprise, not a private business, and generating the positive externality involved in the P&P investment was part of their *raison d'être*. This fourth argument strengthened the other three, making a mix of arguments that prompted them to invest in P&P.

Patient loyalty was recognised by all ESE managers as a major concern, but those in the first group (the ones who invested only the minimum) did not see P&P investment as a strategy to secure patient loyalty. They preferred to appeal to other ways of attracting patients, like working with the community organisations or with the community representatives in the board of directors. For example, the manager of IA appealed to messages like *"this is your hospital"*. The success of this strategy was held unquestioned by the managers, as illustrated by the manager of IA:

*"...We have strong users' associations. The job with the community that is undertaken by the office of social participation emphasises them that it is their hospital"*.

On the side of the ARS, the perception starkly contrasted with that of ESEs. The five ARS officers that were interviewed had the same perception of the importance of investing in P&P and keeping capitated populations with the same ESE for periods long



enough to obtain a return on investment. This was surprising because it contrasted with the shared vision among ESEs that ARSs did not care about P&P, and that for-profits only cared for their short-term surpluses. But even more contrasting is the fact that ARS not only favoured long-term permanence of enrollees with a given ESE, but also considered that ESEs were doing a poor job in P&P. One officer of an ARS said:

*“...Maybe they incur some investment but not the way it should be. I think [they invest] the minimum, basically to keep the contract with average quality (...) But one doesn’t find that it results in interventions on the population risk, or that they set up intervention programs.”*

An officer at another ARS even reacted to the views of the ESEs by saying:

*“...I have been here for three years, and there are level I hospitals like IE and IJ that have had the same population. They also have to audit the contracts and if they find that an ARS is engaged in that crooked behaviour to selectively substitute [the healthy] they must denounce that with names.”*

Another ARS officer, complaining about the low rates of ESE investment in P&P, said:

*“ They are short of cash and they have to save it to pay for their fixed costs. They cannot invest it in the real social object for which it is intended...”*

One ARS officer said that they would prefer to have all the primary care level provided with their vertically integrated network, but the regulations force them to contract out at least 40% of the premium with level I ESEs. This was not their preferred network design as it did not allow them to have an adequate control of the gatekeeping function. Thus, they tried to control as much as they could with their own network, and they have found that morbidity in the groups under their control is lower than that of the groups capitated to the ESEs.

Regarding turnover rates of the population, the five ARSs unanimously claimed that this problem is due to the migration of beneficiaries within the city. In fact, they reported this as a problem of large urban centers, because heads of households move their place of residence to live close to their temporary jobs in the informal sector. This view was

opposed to the view of some ESEs that most people remain in the area and that it justified investment in P&P. The other reason for turnover is that some people find a job in the formal sector, and it means they have to interrupt their enrollment in the SS, because they have to enroll in the EBS. Temporary interruption means that the worker can retake SS enrollment once the job in the formal sector ends, as long as the job ends within six months of interruption. Regarding adverse selection, ARSs complained that it was a problem caused by the SOH itself, as it preferred to assign new slots in the SS to people with chronic and/or costly conditions, rather than paying their care with their resources, through the FGPP system. Accordingly, a problem that was seen by the ESEs as deliberate adverse selection led by ARSs, was seen by ARSs as people's migration within the city, the revolving door with the EBS, and a deliberate adverse selection exerted by the SOH.

Again, the exception of IK explained many of the issues that have been raised. Given that the population was very stable, this ESE did not face the problem of high turnover. In addition, it was a monopoly in the area, so no matter who was the third party purchaser, the individual would always be the same and stay there. It meant that they had a clear incentive to invest in P&P activities with either short-term or long-term effects, as clearly pointed out by the manager.

#### 7.4. Contract incompleteness regarding relationships with the SOH

Contract incompleteness was assessed within the lines shown in figure 2.2. The interviews ascertained: 1) the components of uncertainty regarding prices and quantities of both input and output markets. And 2) the complexity of contracted products in terms of information asymmetries, difficulties to define the contracted products, and observability/verifiability of outcomes of the contract.

Section 7.4. focuses on uncertainty and complexity of contracts with the SOH, and section 7.5. addresses these topics for ARSs contracts. The aggregate effect of RSI and contract incompleteness as causes of transaction costs, is addressed in sections 7.6. and 7.7. for SOH and ARS respectively. The issue of cooperation or less-cooperation regarding transaction costs was addressed in chapter 6, but will be retaken briefly in the discussion.

#### 7.4.1. Contracts for services to the uninsured

##### *Uncertainty*

Uncertainty due to prices of drugs and consumables was highly reduced through the strategy of a purchasing alliance started in year 2000 as an SOH initiative to take advantage of economies of scale in purchasing. The positive effects of this alliance were quickly evident, as it allowed participating ESEs to achieve large reductions in price variability. Managers did not report major sources of uncertainty in labor wages.

Uncertainty about outputs was related to the number of intermediate activities for each category of FGPP. Given that the ESE received the same payment for a given FGPP category, and each of it involved a list of intermediate activities whose quantities varied from patient to patient, the ESE faced the uncertainty that the actual average value of an FGPP category exceeded the expected value because some patients required a large number of intermediate activities. Technically this is a matter of risk (known probability distribution) rather than uncertainty (unknown distribution); thus, risk should be considered a production cost and not a transaction cost. However, the effects of poor risk management on transaction costs were evident as shown below, so it deserves to be treated as uncertainty.

ESE managers identified some FGPP categories that were always losers, like hospital admission to Internal Medicine, or to the Intensive Care Unit. But other final activities had been identified as profit makers, like those related to gynecology-obstetrics and surgical procedures. The losers were typically FGPP categories with a high variance. The reimbursement rate paid for admission to Intensive Care Unit, for example, was based on an expected value of five days of average stay, but reportedly, most patients stayed more than the 5 days. The FGPP categories reimbursed outliers separately on a fee-for-service basis, to avoid the unpredictable effect on the relationship between expected and actual value. However, one level III manager's comment on this payment mechanism reflected a common misunderstanding among managers:

*“The risk is high, for example in Intensive Care Unit. We get CP\$ 6 million<sup>1</sup> however long the patient stays in the Intensive Care Unit. And the majority stay an average of 25 days, but the SOH only compensates the first five days. From then on, I lose all the money because I cannot bill [the services provided]. In addition, the SOH starts paying again after the 40<sup>th</sup> day on a fee-for-service basis, so if the patient is discharged at day 39, I lose all that money.”*

This misunderstanding led managers to believe that the average of each FGPP category was in fact the cap, and they were careful to avoid delivering any intermediate activities beyond the said average. It was evidenced that they did not view the FGPP categories as a group of patients with different levels of severity that eventually averaged out to the expected value set by the SOH. As a consequence, they had a strong incentive to deny necessary care to keep the cost per patient under the average. The case of therapies (respiratory therapy, rehabilitation therapy, etc) illustrates this, as reported by the manager of IF:

*“...therapies were categorised as intermediate activities, so all [ESEs] started to close down their therapy services.”*

However, no manager considered they were able to pull down the average, because the very sick patients required care that was very difficult to deny. This dynamic also created an incentive for level III ESEs to avoid admissions of uninsured patients within FGPP categories that were losers, and give preference to ARS patients that were paid on a fee-for-service basis.

Another response to the payment mechanism that would be expected is the “quicker and sicker” discharge of patients that has been reported for the DRG system in the United States. Although not paying readmissions within 30 days was the solution applied by the SOH contracts, it was not a strong incentive, as put by an SOH officer:

*“...only a few cases were denied payment for this reason, because it was impossible to link the readmission with a previous admission.”*

---

<sup>1</sup> About USD 2,500.

Another manager complained that one important difficulty in keeping intermediate activities under control was that sometimes the uninsured patients demand activities claiming that the SOH should pay for them. For example, a patient was seen in a visit to the general physician and demanded lab tests that were not paid by the SOH. This situation usually created tensions with doctors that they can quickly solve by ordering the demanded tests, but at the same time adding pressure to the ESE billing cap.

On the side of the SOH, they considered that the FGPP mechanism, being a per-case reimbursement, created incentives for efficiency inasmuch as providers reduced the amount of intermediate activities whose demand was supposed to be more likely to be induced, as pointed out by an SOH officer:

*“...as we paid the final activity, the incentive led [ESEs] to reduce intermediate activities.”*

It was expected that managers and clinicians would devise mechanisms to reduce unnecessary care, like the creation of clinical guidelines and upgrading their information technology capabilities to exert effective concurrent utilisation review. However, on the side of ESEs, except for IIC, it was not obvious that they were adopting these strategies. The exclusion of some intermediate activities from the service portfolio, the dumping of patients to other ESEs and cream skimming of the better paid, and the misunderstanding of the essentials of the per-case reimbursement mechanism, were clear examples that illustrate the predominance of undesirable responses to the payment system.

Key flaws in the estimation of FGPP rates were also pointed out by managers. One of the most commonly highlighted weaknesses was that the SOH estimated one single average per FGPP category for all ESEs, ignoring the differences in fixed labor among ESEs, as shown in chapter 6. This put an additional burden on these high-fixed-cost ESEs, whereas low-fixed-cost ones enjoyed an unwarranted advantage.

It was also recognised by managers that variance hurt more intensely level III ESEs. Case mix and severity was reported to have a larger variance in tertiary care, whereas in level I facilities variance was much lower. The SOH included level of care as one of its adjustment variables, but other variables such as age, gender, case mix and hospital

costs, were not included. An SOH officer recognised this weakness, but the reason for not having gone further in such estimations was basically lack of data.

Quantities demanded were also a cause of uncertainty at the aggregate ESE level, because some ESEs were unable to reach the billing cap set by the SOH whereas others exceeded it. It was clear though, that wide variations below or above the caps were not reported, because the SOH estimations were carefully done to avoid such wide variations. But small variations were still a source of uncertainty that was readily acknowledged by managers. For the underproducers, uncertainty regarding a fall in demand would cause a painful process of negotiating a bailout with the SOH to pay for fixed labor costs, whereas for the overproducers, uncertainty related to the possibility of writing off overproduction as uncompensated care. In the first case, uncertainty was not so bad, if the manager felt comfortable about receiving a bailout from the SOH. However, bailouts were negatively perceived by managers, who viewed it as a symptom of poor managerial ability to increase output. The challenge among managers who received bailouts was to reduce them progressively, as illustrated by the manager of IIIA and IID, to emulate those who did not receive bailouts, who were in turn considered successful managers.<sup>2</sup> On the side of the overproducers, uncertainty was related to not getting compensation for services provided, versus the possibility to use this overspending as an argument to gain increases in the billing cap for the next fiscal year, as was mentioned by the managers of IIB and IC.

Regarding uncertainty about the output prices, the FGPP categories had a very stable price, which was related more to the resources available at the SOH and the MEE at each ESE, rather than to actual costs of production of each category. However, the prices for some categories that were consistently losers were eventually renegotiated with the SOH.

A final aspect generating uncertainty in the contracts for services to the uninsured was the fact that no matter the contracts were prepared, negotiated and agreed upon between the SOH and the ESEs, the final version of the minutes, specifically the annexes, were usually handed out two or three months after the initiation of the contract. Annexes were key to contract execution, because they described the details of the FGPP

---

<sup>2</sup> A more in-depth analysis of managerial values, which are related to career concerns, will be undertaken in chapter 9.

categories, what was included and what was not included, and other contractual details. A delay in handing out the annexes meant that ESEs started working with the contracts without much certainty about some billing procedures. The problem arose because after having been agreed upon between the parties, the SOH frequently introduced changes to the annexes unilaterally, which ESEs did not realise until they received the final minutes. Once they received them, they realise that they had been making mistakes that caused payment denials. This complaint was generalised among ESEs managers, but the SOH did not pay much attention to their complaints.

### ***Complexity of contracted products***

Regarding contracts for services to the uninsured, complexity of contracted products in terms of difficulties to define the products was not recognised as a major cause of contract incompleteness, large enough to cause contractual difficulties. The products, although showing wide variation and uncertainty, were seen by managers as simple to define, and much simpler for level I ESEs. But even at level III ESEs, the products were not seen as very difficult to define. Instead, all managers agreed that the problem was to define the boundaries between what is covered by the SOH and what is covered by the ARSs. In addition, information asymmetries were not seen by ESE managers as a big problem in these contracts, mostly because the SOH did not care as much about health outcomes as they did about inputs, sometimes process indicators, and always outputs.

However, information advantages on the provider side allowed for the “upcoding” problem, whereby the provider carries out hidden actions to maximise revenues from a given case. This was pointed out by an SOH officer:

*“...they learned to bill and to work the package: if a patient showed up at the A&E service, they kept him for 24 hours to charge the inpatient day.”*

Managers’ views of complexity of contracted products contrasted with their perception of high uncertainty. However, it was clear that managers were aware of the contents of an FGPP category in terms of intermediate activities (i.e., no complexity to define the product), but they acknowledged a lot of uncertainty as to how many times these intermediate activities were expected to be carried out.

Regarding the boundaries with the POS-S, it was evident from all the interviews that these boundaries were very blurred. Although this problem was seen to be worse by ARSs officers, some problems were also mentioned by SOH officers. For example, one level II manager said:

*“When the patient shifts to non-POS-S it is my responsibility to provide prescription drugs for 20 days, provided the patient is enrolled in the Subsidised Scheme with a chronic condition. Then we start with ‘what is a chronic condition?’ (...) the question is what the SOH takes for a chronic condition. That is the problem [when the SOH considers it is not a chronic condition].”*

This problem was raised as a concern for audits, because if the SOH did not consider the patient had a chronic condition, it denied payments to care provided to that patient.

Complexity was reduced by detailed descriptions of FGPP categories in the annexes, but it was also reduced because the SOH did not pay too much attention to outcomes as a source of contract non-compliance. ESE managers recognised they were never penalised for poor quality outcomes, and the SOH focused its concern for quality on input and process indicators. Quality outcomes were assessed through other mechanisms such as epidemiologic surveillance, but these processes had no implications for contract compliance. For example, maternal deaths or nosocomial infections were routinely and carefully analysed to detect system or human errors and to correct them, but they did not generate penalties to ESEs or early termination of contracts. This lack of contractual penalties contrasted with the explicit consideration in contract minutes that one of the obligations of the ESE was to provide services with the optimal quality (timeliness, sufficiency, comprehensiveness, logic-scientific rationality, safety and good manners). It was recognised by all managers that the SOH was not concerned with the strict application of this contract clause, but it rather focused its concerns on inputs (entry requirements), processes and quantities provided (output). A comment from the manager of IIF illustrates this perception:

*“...perhaps some things are measured, like entry requirements to fulfill some minimum standards. (...) we are still far from being able to measure quality, among both public and private providers.”*



Other ESEs were concerned for technical quality, although not necessarily in terms of outcomes but restricted also to input and process measures. For example, the manager of IIID complained that, being a teaching hospital, its staff clinicians devised evidence-based clinical guidelines that implied a much larger use of resources than that allowed by the FGPP categories. It caused the manager great difficulties to control costs associated with intermediate activities. In contrast, the manager of IIC reacted in the opposite direction, i.e., devising the clinical guidelines with the explicit purpose of staying within budget. These guidelines were actively enforced among staff clinicians, as related by the manager:

*“Nobody here, wherever he comes from, is allowed to ignore clinical guidelines. Working at this ESE means adhering to the guidelines. (...) Whoever does not follow them, is brought to a disciplinary process, or his contract is terminated.”*

It was also pointed out by managers that the SOH paid much more attention to attributes of quality related to users’ perceptions, such as patient grievances regarding timeliness of ambulatory visits and full delivery of prescription drugs. In fact, IIIB was imposed a financial penalty in 2002 because it failed several times to deliver prescriptions completely.

It is important to note that the SOH outsources to third parties the audits of the contracts with ESEs. Three major private firms provided the audit services and they followed the guidelines of the SOH to review the bills and apply administrative and clinical denials. Thus, it could be said that these auditors’ lack of concern for best quality health outcomes was the result of the SOH definition of how its outsourcers should audit the contracts.

#### 7.4.2. PAB contracts

Uncertainty with PAB contracts was largely reduced because the contracts were specified in terms of number of activities, and prices reflected input costs. Even with such degree of certainty, some ESEs like IE and IG complained that communities occasionally demand interventions beyond those contracted with the SOH. In these cases, the manager usually denied those services, or when they were provided, the manager usually asked the SOH or the local mayoralty for the additional money. It was

not uncommon, however, that these extra activities remained uncompensated, as reported by the manager of IG, but their amount was not large enough to cause a major financial concern to the ESE.

The PAB products were even more clearly defined as compared to those of contracts for the uninsured. Managers complained that these products were excessively detailed in their description, which left little room for adjustments to local or temporary variations. Concerning quality as a source of contract difficulties, managers also mentioned that the SOH did not care about outcomes, and even less about impact. They restricted their audits to verify that the inputs required by the contract were in fact used, and the number of activities contracted were effectively carried out.

#### 7.4.3. Contracts for ambulances

Uncertainty in these contracts was mostly related to the possibility that the ambulance broke down, because when this happened, the crew was left idle and the SOH did not compensate the ESE for the days it was out of service (except the first 24 hours). When crew members were short-term workers, it was easy to redirect them to perform other tasks within the ESE, but when they were official workers (the most rigid civil service category), the ESE was not allowed to redirect them to other tasks. Thus, the ESE had an incentive to keep the ambulance working, but the probability of breaking down increased with ambulance age. However, the manager of IIA said that preventive maintenance was as a strategy to minimise the chances of break down.

Concern for quality was focused mainly on the fulfillment of entry requirements, which were related to structure indicators, and the CRU exerted control of continued fulfilling of entry requirements.

#### 7.5. Contract incompleteness regarding ARSs

##### *Uncertainty*

Input prices showed the same uncertainty that was mentioned for the SOH contracts, with respect to both human resources and consumables. With regards to output prices, uncertainty had two clearly different dimensions. On the one hand, fee-for-service

contracts used to be regulated for all public hospitals, so they were required to charge tariffs based on the SOAT fee schedule, which was more expensive than the one typically used by EPSs and ARSs for private providers. However, that restriction was freed during the late nineties, so ESEs were given freedom to compete on prices. This created a lot of uncertainty because ESEs competed fiercely to attract contracts from ARSs, creating a race to the bottom that generated stronger pressures on high-fixed-cost ESEs. On the other hand, capitation contracts were not regulated and they were viewed as a means to avoid the problems observed in the fee-for-service contracts. But fierce competition also forced capitation rates down. This fall of rates was accelerated by the finding of one ARS that its costs of providing primary care services in its own network were much lower than what they were paying to the ESEs, as shown in chapter 6.

The largest cause of uncertainty in capitation contracts, as expected, related to quantities demanded, which varied because of 1) the incidence of disease and patient-driven first-time consultations, and 2) the decisions made by the treating doctor after the patient made the first contact with provider. These two components of risk are called primary risk (or incidence risk) and technical risk (Robinson, 2001c), and are at the root of managed care models to control costs and improve quality. As said for the SOH contracts, risk management, unlike uncertainty, is supposed to be a production cost and not a transaction cost, but it was found that poor risk management caused large transaction costs when ESEs appealed to undesirable strategies to keep expenditures within the capitation budget. The causes of variance in primary risk that were most commonly reported by managers were those related to adverse selection. But as commented above, ARSs and ESEs views of this problem were very different. Another reason was that patients show up at different branches of the same ESE to demand repeat services, medications, and ancillary services for the same episode of illness.<sup>3</sup> This was particularly problematic for ESEs with a large number of branches like IA, IIC and IIE. These ESEs had to invest in strategies, such as call centers and geo-referencing, to reduce this source of uncertainty. A poor use of P&P strategies was also a cause of increased primary risk, which was addressed through various approaches by ESE managers, as shown above.

---

<sup>3</sup> This was a case of moral hazard, when these patients received appropriate care in the first place. In addition, a common practice that has been recognised for a long time is that some patients collect medications at several points of service in order to sell them for personal profit.

In addition, an ideal proactive approach to primary risk should involve both purchaser and provider, because, in a capitated contract, both should be interested in having lower incidence of events. However, one ARS complained that ESEs are not able to engage in such complementarity, either because they do not have the skills, or because they consider it is strategically better for them not to share information on costs and utilisation rates with ARSs.

Regarding technical risk, ESE managers were expected to adopt proactive strategies to reduce unwarranted variations in medical decision-making. Accordingly, the solutions mentioned by managers were similar to those found for the SOH contracts: clinical guidelines and internal utilisation-review, as exemplified by the manager of IB:

*“When you have a capitation contract you have to carry out a very good tracking of it, so that a common cold is not prescribed third-generation antibiotics, which in fact was happening. We know the checkbook is in the hands of whoever cares for the patient, not us [the ESE], so we have to establish controls on prescription patterns and adherence to clinical guidelines. These guidelines allowed us to control [costs] but not reduce quality.”*

Although guidelines were mainly used in capitated contracts, they were also used in fee-for-service contracts to reduce the probability of billing errors and further payment denials. For example, IIC developed clinical guidelines for the most frequent conditions, that were agreed upon with purchasers, but were homogeneous for all of them.

This ESE was an outstanding exception to poorly controlled capitation contracts. It set up a detailed mechanism to keep track of utilisation and the manager introduced a breakeven clause in the contracts with ARSs whereby the ESE made sure it would not lose money. Thus, when actual loss ratios increased beyond expected values, the contract was terminated. This clause had been applied only once, to ARSIV, and the ESE manager mentioned that it was due to deliberate adverse selection, that is, primary risk was generating a high level of expenses that was enough reason for early termination. This no-loss clause was a strategy to minimise risk exposure that had not been applied by any other ESE so far. However, the IIC manager pointed out that the legal framework of ESE allowed them to use the so-called “exorbitant clauses”.

Although exorbitant clauses are a device of the general legal framework for public sector contracting, the private-like legal framework of ESEs allowed for the use of these clauses to protect public autonomous hospitals from losses. Surprisingly, IIC was the only ESE that had used this device as an explicit clause in its contracts with third party purchasers.

Whereas on the side of ESE managers they highlighted their risk management skills, on the side of purchasers they revealed their concerns for undesirable ESE strategies to reduce primary and technical risk. For instance, the SOH argued that:

*“...what one sees with the surveys [of ESE users], is that access barriers are being put to enrollees in the SS, and it is due to the incentive [inherent to] capitation.”*

ARSs argued that ESEs were excessively restrictive in their controls on doctors' clinical decision making. One ARS officer reported:

*“What we feel now is that they capitate and the service is too bad, they use the cheapest antibiotic available, I believe the solution is to shift towards clinical guidelines and open a space for a technical-scientific discussion with them.”*

This same ARS officer complained that signing capitation contracts with ESEs meant a loss of control over risks, as compared to the control they previously had when they paid on a fee-for-service basis, or as compared to the control they were able to exert within their own vertically integrated primary care network. In the context of fee-for-service, this ARS was able to keep track of utilisation rates and to carry out data analyses that gave them a better handle on risks, because ESEs had to report detailed information to get their bills paid. Later on, under capitation contracts, it was very difficult to obtain that information from ESEs, even more when some of them consider it was strategic to hide information to the purchaser.

One concern that was mentioned by several ESE managers was that a rapid growth in enrollees in capitated contracts implied a risk of losing control on loss ratios. The manager of IH mentioned:

*“I believe, it has to be acknowledged, to the extent that one has to bring revenues to the hospital, one does not measure those consequences. One has also to show better results in contracting. Managers take the risks, and in some instances (it is not our case so far), they have had to roll back contracts because capacity is overwhelmed.”*

### ***Complexity of contracted products***

As mentioned before, defining the products was complex in terms of setting clear boundaries of what was included and excluded from the POS-S. Given that this benefit package had a large gap in the middle that excludes intermediate level care (see table 4.1.), ESEs providing care at these levels faced a permanent source of disagreements. One explanation for these blurred boundaries, provided by an ARS officer, was that the original intent of law 100 was that the benefit packages of the SS and the EBS would be equal by year 2001, so it was not necessary to dedicate much effort to define boundaries that were expected to change every year as items covered increased, and even less if such boundaries were to be unnecessary four years after ARSs started operations when both benefit plans would be equivalent.

However, the actual development of financing of the SS prevented the achievement of homogeneous plans, and a temporary situation turned into a permanent one. When it became permanent, the need for a better definition of boundaries was more evident. But as the devil is in the details, a sharp definition of such boundaries proved elusive. This created room for cost-shifting between ARSs and SOH, and the patient had been victim of what is called the bouncing back and forth of patients. This made it very difficult to achieve an adequate design of incentives and provider networks. A level I ESE manager complained that:

*“...to the extent that the three contracts [levels of care] are kept separate and not coordinated, it is not possible to generate improvements [in health outcomes].”*

Almost all interviewees noted that contract minutes were apparently very clear in defining the products by citing the relevant norms. But the blurred boundaries were revealed when the parties built their own interpretations of the ambiguities of the norms. This was reported to be the cause of most of the disagreements between ESEs and ARSs, and it was clearly illustrated with the definition of emergency care. According to

the norms, it is considered any care provided to a patient until stabilisation. Thus, the interpretation of what was to stabilise a patient had given rise to a large number of concepts. One strategy the interviewees mentioned to deal with this source of uncertainty was to hire a lawyer to provide an interpretation of the norm, but it only added more details to the confusion, because both purchaser and provider put their respective lawyers' concepts favouring each party's interpretation of the norm. In addition, a lawyer's concept was not binding to the other party, because it was just a concept.

An alternative strategy to deal with the blurred boundaries was to request the SOH to unify criteria exerting its role as local health authority or appealing to scientific associations. However, the SOH was also unable to define the boundaries. It set up several working roundtables where the parties expressed their points of view, but no strict guidelines emerged from those meetings. Appealing to external academic authorities like scientific associations was criticised by purchasers, who complained that their guidelines were very demanding for the resources available. In the end, it was reported that disagreements were dealt with most of the times on a daily basis by the auditors of each party, when they sat to negotiate the denials of payments either for services provided under fee-for-service contracts for ARSs, or for services shifted to the SOH contracts under capitation contracts with ARSs.

Beyond what is included or excluded in the POS-S, the other problem was the definition of medical necessity. This was a problem of fee-for-service contracts. Again, this problem was illustrated by ESE managers with care at the A&E service. Given that most of the capitated contracts did not include A&E visits, these were paid on a fee-for-service basis. Thus, when the ESE billed the ARS for A&E visits, the ARS denied payment for many visits they considered were not emergencies. Extreme versions of abuse can be seen on both sides of the problem. The manager of IIB complained:

*"...there is no sharp regulation that clarifies what is initial emergency care. You and I doctors know what initial emergency care is. But the ARS [demands that] a patient with a cardiac tamponade be managed only with IV fluids. We know that a cardiac tamponade has to be opened and decompressed."*

Stories to the other extreme were also told by ARS officers:

*“I cannot show up at the ER with a four-day pain for an ingrowing toe-nail. But [an ESE] that is very unique in Bogota, charges that visit [as an ER visit] with the argument that if the patient is turned away he can die and the manager can face criminal charges.”*

Non-comprehensiveness of the POS-S and the consequent incentives for cost-shifting, were exacerbated by compartmentalised contracting by the ARS. It meant that a single patient did not receive all the necessary care with the same provider, and the transfer from one provider to another created problems of coordination across stages of care. This was mentioned by the manager of IID, an ESE that also provides level I services. The ESE ran a capitated contract for level I services for a population of a given ARS. If a patient from that contract needed level II services, he had to be referred to another ESE for a service the first ESE already had in place.

The difficulties related to boundaries of the POS-S and medical necessity created much tensions among ESEs and ARSs. On the side of ESEs, the problem was seen as the result of ARSs trying to shift costs to the SOH or to avoid payments as much as possible. But on the side of ARS the problem was seen as lack of ESEs’ ability to adequately manage risk and a prevailing short-term view of cash flows.

These contradictory views also yielded opposite views of quality concerns. ESEs considered most ARS were not as concerned for quality as they were, but on the ARS side the perception was exactly the opposite. Obviously, the restriction of lack of observability and verifiability of quality outcomes was also found among ARS contracts, and quality measurements focused mainly on structure and process indicators. Structure indicators were those established as minimum entry requirements, which the ARS had the obligation to verify before signing a contract with any provider. Process indicators were the same as those shown for the SOH, namely patient satisfaction, waiting times and complete delivery of prescription drugs. Patient satisfaction was a major concern for all ARSs and they were ready to respond to patient grievances.

Two of the analysed ARS were considered by most ESEs as really concerned for technical quality, although on the side of these two ARS it was found that they did not



have the same views about the ESEs. An officer of one of these ARSs, showing interest for best health outcomes, recognised the difficulties in achieving this:

*“...if you introduce quality indicators in the contract, no ESE manager signs it. (...). Last year, for example, I tried [to introduce an indicator] to keep an average length of stay on X, but they didn’t sign it. They resist the introduction of such precise indicators in the contract because they know themselves and they know they can fail. For the last two negotiations we have hardly managed to introduce indicators of timeliness, and other topics they do not find thorny. From my viewpoint, the contract should have all the quality indicators (...). But they do not sign it.”*

It is curious, however, to note that one of these two ARSs has a clause in the contract that allowed the ARS to deny payments when clinical errors were detected, but it had never been applied despite the occurrence of errors that were recognised by both parties. But even in the presence of blatant errors the ARS complains that all it would be able to do was to file a complaint with the SOH, because terminating a contract for quality reasons would mean falling below the 40% compulsory contracting with the local level I ESE.

An officer of one ARS that had its exclusive private provider for tertiary care manifested his concern for quality in two different situations. On the one hand nosocomial infections were a big issue, because the patients of this ARS typically were transferred from one facility to another before getting complex enough to be taken to its private tertiary care hospital. When the patient was referred to the private hospital, a wide variety of nosocomial bacteria had been collected through all the previous venues, and these bacteria increased the risk and complexity of nosocomial infections at the last receptor, the private hospital. The other concern for quality expressed by this ARS officer was that after the initiation of capitated contracts, complications at tertiary care soared:

*“...we noticed a big change, we then started tracking back the cases and we found that [the ESEs] were discharging the cholecystectomies the following day [of an open surgery]. We noticed that savings from capitation were causing an impact on high cost care.”*

In a further comment, this same ARS officer said:

*“...all these [contract] clauses are prepared for the time the relationship is broken up, but at the end of the day [all the ESE managers care about] is the terms of cash flows and capitation amounts.*

#### 7.6. Transaction costs regarding SOH contracts

As expected from theory, the presence of RSI and contract incompleteness causes a series of risks and prompts strategic behaviours by the trading parties that generate transaction costs. The costs of the contracting process itself include the design, negotiation, monitoring, enforcement and renegotiation of contracts. Regarding the costs of design and negotiation, it was found that all ESEs pointed out that the SOH leaves little room for negotiation of contracts, because of the dependence on SOH revenue for paying fixed costs. Marginal changes to prices, billing caps and quantities contracted can be negotiated but the room for bargaining was reported to be very limited. The manager of IIID stated:

*“...In legal terms it is called a Contract of Adhesion: Either you sign it or you sign it. So the ESE’s bargaining capacity is minimal”*

The ESEs also recognised that negotiations were more difficult at the beginning of the FGPP system, when the minutes had to be first designed to reflect the new payment system. But after the first round of negotiations, a learning process made contract renewal easier.

As shown in chapter 6, this situation of mutual dependence was key for ESE managers to cooperate or not, depending on the chances they had to force the SOH to satisfy their demands or to tolerate their poor performance.

In addition, the SOH had to incur extra costs during the first year of the FGPP because some ESEs, led by IIIC, were challenging the new payment mechanism on the grounds that it was causing a decrease in quality. The SOH had to contract a study on the impact of FGPP on quality, which showed no negative or positive effect (see in chapter 4 the study by Marulanda et al (2002)). All ESEs considered the costs of design and

negotiation of contracts with the SOH were reasonable. However, the costs of monitoring the contracts through outside audit companies were considered high by most ESE managers, as illustrated by this quote from the manager of IIC:

*“...I am not sure about the figures, but I believe that those auditing firms (...) were costing about CP 5,000 millions or a bit more (...) [compared to the total amount] among PAB and uninsured of about CP 250,000 million, I think it is an exaggerated cost, from any point of view ”<sup>4</sup>*

In addition, ESEs had to incur other costs of monitoring because when the auditing company goes to the ESE, it has to mobilise human resources and clinical records at a cost that is not reimbursed by any party. In fact, the manager of IJ said:

*“...If the contract were more simple and, accordingly, its verification also more simple, it would not be necessary to incur those costs and [the ESE] would be able to make a surplus. A large share of the money one cannot take as surpluses is spent in supporting administrative costs to justify contracting.”*

In terms of administrative costs of billing procedures, these were considered high by the parties. For example, IIC kept a large group of administrative personnel on each floor of the hospital, to make sure every bill was issued correctly and minimise the possibility of denials. These costs were seen by the manager as an investment because they allowed the ESE to increase revenues due to fewer denials. On the side of the SOH, they considered that shifting from fee-for-service to FGPP system simplified the billing processes, because the ESE did not have to issue a detailed itemised bill but just a global bill with the number of final activities provided.

The follow-up of contracts also was seen by ESEs and the SOH to generate costs to the parties, not only regarding the outsourced auditing companies, but also the bi-partite committees that were set up to keep track of contract execution. These committees were seen by most ESEs as positive, because they were used as a means of communication with the SOH, and to be informed about its changes in policies and procedures. Others like IIF considered that these committees did not add value to patient care.

---

<sup>4</sup> This perception proved overestimated. The actual amounts paid to these auditing firms were CP 1,971 millions (data officially provided to the researcher by the SOH's directorate of health insurance).

Enforcement costs were not reported. The parties usually resolved their disputes in a friendly manner, and if not friendly, no considerable extra costs of enforcement were reported. The only exception that was found was that of IID, which was demanding from the SOH the payment of a large amount of ER consultations during 1999 and 2000. These services were provided without a contract, and the SOH claimed it was not liable to pay services it had not contracted for, whereas the ESE claimed that those services were provided to uninsured people, who were the responsibility of the SOH. This dispute was taken to the Courts and by year 2004 it had not been settled. It was also found that some of the ESEs that exceeded the billing cap demanded the payments from the SOH, but these disputes had not been taken to the courts. Usually they were settled by the SOH by just saying there was no money, and considering these overprovisions as social surpluses. Although they could be a source of deficits, they typically took place with self-sustainable ESEs like IIC, IB, IIB, IC, and IIF and did not generate deficits to these ESEs.

Disputes arose most of the times because of the gray zones left by the regulations, as mentioned before. These gray zones were not solved by the contract clauses because they referred to the current regulations, which left lots of gray zones. As mentioned above, the parties adopted a series of strategies to settle disputes, but all of them added to the costs of contracting. The manager of IJ pointed to the fact that disputes were not directly with the SOH but with the outsourced auditors:

*“...very often our conflict is with the auditor firm and with the way they interpret [the contracts], and the SOH does not take part in this. The SOH simply takes what the auditor firm says because it is its agent vis à vis the contract.”*

The manager of IIF argued that differing perspectives were the source of costs:

*“...I think the issue is that the two parties must clearly understand what providing health care implies. So, when the local authority achieves a better knowledge of what survival means for the provider, the situation will be more clear.”*

With regards to PAB contracts, a similar dynamic was reported. The initial contracts were more costly to design, but a learning process made renewals less costly. These

contracts were also contracts of adhesion, so there was little room for ESEs to bargain. However, all managers of the ESEs that provided PAB services complained that the liaison for each program within the PAB spent a lot of time in meetings at the SOH, both during preparation of new contracts and during their execution. Nonetheless, the manager of IIC considered it was better for the contract execution:

*“...[the PAB contract] is a bit more demanding [but] it is based on a better consensus. (...) it involves more activities, they are more rigorous in the technical aspects of each activity, it requires a much longer negotiation process (...) It takes more time, they end up being much more specific, but they are based on a better consensus as compared to [the contract for the uninsured].”*

Billings for these PAB contracts were seen as quite simple by managers, but the audits were seen as more stringent. The same issue of dispute resolution was found with these contracts, i.e., the audit firm adhered to its interpretation of contract clauses and the norms, so the costs of dispute resolution were incurred with the audit firm rather than with the SOH. However, disputes were less protracted than those of contracts for the uninsured, because there was a stronger consensus prior to the signing of the contract. No costs of enforcement were reported for PAB contracts.

With regards to contracts for ambulance services, the costs associated with contracts were reported to be the lowest because this contract was very simple. The only complaint was that sometimes the CRU decided to transport an ARS patient without the previous authorisation of the ARS; when the ESE billed the ARS for that service the ARS denied the payment on the basis that it had not authorised the service.

The overall perception of completeness was inquired about with a direct question: “How complete do you think these contracts are?” 13 of the 22 ESEs considered the contract for services to the uninsured was incomplete, whereas 3 out of 15 considered the PAB contracts were incomplete. All ESE managers considered the contracts for ambulances were complete.

#### 7.7.Transaction costs regarding ARS contracts

Given that the types of relationships with ARSs varied from cooperative to less cooperative, as shown in chapter 6, the costs of contracting also varied accordingly. Learning effects were reported to have an effect on decreasing the costs of contracting. The evolution of minute design can be illustrated with the case of ARSI:

*“...For year 2000, when we started with capitation contracts, I sat down with the legal department of ARSI and we came up together with a model minute, but experience has shown us a lot of good and bad things of contracts. For example, two years ago I went to the Ministry of Health and we evaluated the minutes. We have been adjusting them not only to the new regulations but also to what we have learned from experience and to ESEs’ remarks. It is clear for us that a minute is not a straightjacket that I impose on the provider, but we have always negotiated on the basis of what we worked out two years ago and that we have successively modified.*

This same ARS illustrated how the negotiation took place:

*“...[the cost of contracting] is not exorbitant. It takes time because although the minute is long, I think it is good, it is clear. ESEs take a long time, one or two months, to make their remarks before legalising the contract. In the meantime, we work with a short pro-forma contract we sign initially, and then the long minute takes a bit longer to come in.”*

However, ESE managers also considered they had little room for negotiating contracts with ARSs. This was confirmed on the side of ARSs, in two ways: 1) they usually sent their minutes to the ESE and minor modifications took place. And 2) ESEs had a different minute for each ARS, whereas ARSs had basically the same minute for each type of contract (capitation, fee-for-service, packages) and the same minute was sent to the ESEs.

The negotiation process typically focused on two big issues: prices and counterbillings. For the fee-for-service contracts, prices referred to the fee schedule to be adopted and the percentage discounts. For the capitation contracts, prices referred to the percentage of the premium that would be paid for a given package of benefits. Both contracts also involved price discounts for early payment. Counterbillings referred only to capitation contracts. Beyond prices and counterbillings, the rest of the contract clauses were

usually taken for granted and were paid little attention. Some ESEs like IIF, IID and IB had standardised negotiating processes to reduce their costs, to such an extent that managers delegated it to their marketing area and they only participated when disagreements arose or key decisions had to be made. In the case of IB it was found that the contracting process had evolved so smoothly that the minutes were three or four years old, and they were simply renewed. Although the contracts were for one year, their almost automatic renewal made the relationships long-term in practice.

The costs associated with the execution of ARS contracts were related, as shown above, to the issue of gray zones in the boundaries of the POS-S and the gray zones of clinical decision making, namely medical necessity. However, these were costs typical of a fee-for-service contract, whereby it would be expected that the risk bearer (i.e., the ARS) had a strong incentive to engage in cost-shifting (either to the SOH, the ESE or the patient) interpreting the gray zones to its own advantage. In fact, denials of payments in fee-for-service contracts was pointed at by ESE managers as the largest source of transaction costs because they had to invest time and effort to avoid the denial, or to get it paid once it took place.

The costs of preventive approaches to denials were evident with the case of IIC, with its large group of administrative personnel that helped to prevent billing mistakes (to keep clear the gray zones of coverage), and its active enforcement of clinical guidelines (to keep clear the gray zones of medical necessity). Some ESEs decided to outsource the billing process to specialised firms, in the expectation that these firms would reduce the amount of denials because of their better knowledge of regulations and of ARSs standards for auditing bills. These outsourcers operated within the ESE, preparing the bills and making all the necessary checks to minimise the possibility that the bill was denied payment. However, the manager of IIC was clearly skeptic of this strategy to handle billings:

*“...ESEs experiences with outsourcing the billing process have been a disaster. [These outsourcings] do not work because billing depends on many things that are completely independent of the billing system itself, because a bill starts at the doctor. (...) you cannot put an auditor behind each doctor and nurse and lab technician [to make sure they do not make billing mistakes]. Billing is not only about consolidating numbers.(...) to me, outsourcing the billing process does not work because when doctors know it is*

*the outsourcer's responsibility, they say 'the outsourcer is getting paid for making bills, it is not my problem' and they do not engage in the process. Instead, when it is the hospital [which makes the bills], the doctor knows that if he does not [make it correctly] it is the hospital who loses and it is his business and he knows that if he fails, I will crucify him. Billing is not rocket science, the problem is the professional, to make him sensitive to the hospital, and to raise his interest in reporting [activities done]."*

A sort of specialised governance structure that was reported by ESE managers to emerge in ESE-ARS relationships, was that of the follow-up committee. This structure was made up of representatives of the two parties and tried to reduce transaction costs with a preventive approach. These committees were very helpful to clarify disagreements before they caused a denial, and to consolidate these clarifications into a common understanding of gray zones so as to clearly define for the future, who pays what under what circumstances. In fact, they were considered a very good complement to contracts and regulations, as these two latter were recognised to be unable to fill the gaps associated with contract incompleteness. However, follow-up committees had their limitations, because most of the times the participants were not decision makers, so they had little power to resolve disputes outside the contracts.

Complaints were also raised about the learning process. The manager of IIA complained that:

*"...There are not clearly defined and fixed rules of the game, but they are subjective, according to who is there (...) if my experience tells me that I am being denied payments for this and that, then I change to try to reduce denials. But the next month I find new causes of denials."*

Another complaint from managers refers to the wide diversity of contracts that a single ESE had to manage. Given that contracts differed in terms of interventions covered, payment mechanisms, referral procedures, etc., holding several contracts increased the likelihood of making mistakes that resulted in payment denials. In addition, the specific processes of a given contract were so detailed that it was very likely that a mistake was made in preparing a bill. However, this was a complaint among level I ESEs like IF and IJ, whereas a level III ESE like IIIC, as shown above, had implemented a sophisticated apparatus precisely to minimise the probability of mistakes.



A reaction to the transaction costs associated with fee-for-service contracting was to shift towards capitation contracts. This was what both ARSs and ESEs expected. However, this type of payment also had its own costs. On the one hand, the issue of counterbillings that has been mentioned before, created the need for preventive mechanisms such as call centers to direct patients to their geo-referenced provider, transportation costs to pick up patients at out-of-network providers, and the *ex-post* review of services provided out-of-network to agree if a counterbilling was justified or not. All these preventive mechanisms added costs to the capitation contracts. In addition, the strategy to pick up patients to reduce counterbillings requires cooperation among ESEs, which was not always the case. In fact, it was reported by the manager of IE that:

*“...we have someone 24 hours a day on the radio who is permanently in contact and is well informed about everything. The problem is that level II and III ESEs are doing things that [do not correspond to their levels of care]. And we argue with our medical auditor because they did not notify us, and we were willing to take care [of the patient].”*

On the side of the ARS, the shift towards capitation was also expected to reduce the costs it incurred to control the fee-for-service incentives towards demand inducement. However, they quickly realised that there were other types of transaction costs, as reported by an ARS officer:

*“...we propose weekly follow-up meetings, we have an agenda like a school schedule for those meetings. But when one proposes [the meetings], some [ESEs] say ‘for what? Why do you want to know that, if I’m in the driver’s seat and it is my money.’ But we want to know what inconveniences our enrollees have.”*

Another ARS officer said:

*“...[we said] let’s check how it works and if it really decreases transaction costs and [allows us to have] a smoother operation, because the issue of denials and conciliations [involved in fee-for-service contracts] was becoming a headache. So we started to capitate with ESEs we had a good relationship with. (...) What we thought would solve*

*transaction costs ended up causing greater [patient] illness and more administrative processes (...) The [costs associated with] denials in the fee-for-service payment became the [costs associated with] counterbillings of capitation. I think the relationships were more alienated with capitation.”*

Another ARS officer had a more general view of transaction costs in all types of contracts:

*“...The problem is that there is no trust between the parties. So I am always thinking the provider is inducing more interventions than those the patient needed, and the provider is thinking that I will make more denials than they expect. Then they induce 100.000 more because that is what they expect I will deny. That has caused instability due to distrust.”*

So, apparently capitation and fee-for-service contracts had each their own types of transaction costs, making it difficult to determine which costs were higher. Accordingly, an alternative governance structure was vertical integration of primary care networks, as suggested by one ARS officer:

*“... Our stance on capitation is to reduce it as much as we can and administer it directly.”*

However, as said before, Law 715 compelled ARS to contract at least 50% of the premium with public hospitals and the SOH compelled ARSs within Bogota to contract at least 40% with the local ESE, so the prospects for increasing control over the gatekeeping function via vertical integration seemed rather limited.

This created a less cooperative relationship among level I ESEs, as shown in chapter 6. Conflict burdened relationships were explicitly reported by IG, IJ and IF.

All in all, transaction costs were not reported to be high enough to appeal to more powerful enforcement mechanisms such as the abuse of monopoly power or litigation. In addition, several regulations were issued by the government to reduce the problem of delayed payments, another pervasive problem reported by ESE managers in their relationships with ARSs. The latest regulation was the decree 050 of 2003. Before then,

ESEs reported that they had a hard time collecting their receivables, but now they all recognise that the decree has had a very positive effect.

The use of monopoly power (due to the compulsory 50% contracting stated by Law 715 of 2001 and complemented by SOH regulations to contract 40% with the local level I ESE) was not seen by managers as an alternative to force ARS to pay on time. The manager of IB openly said it was not a good idea. And litigation was not considered because the risks and costs were also high. In this sense, the manager of IIC said:

*“...we had the option [to sue the ARS] but it would be a serious mistake from our viewpoint. A judicial process takes, in the best case scenario, three years, if not five or six. (...) and the commercial relationship would have been broken up (...)*

Accordingly, alternative strategies to force ARSs to pay on time were used by ESE managers. The manager of IIC, for example, considered that:

*[it is better] to use the SDS pressure, close services [to the ARS], etc. We had a strategy to put ARSs with overdue payments on the notice board of all points of service, and that exerted pressure because everybody saw them.*

The manager of IIIA had a different strategy. Once the 30-day term to pay bills expired, she sent memos to the overdue ARSs, with copies to the controlling authorities, and these sent memos to the ARSs. It caused reactions on the side of ARS, as reported by the manager:

*“...IIIA was considered the enemy of ARSs (...) I had my debt turnover at 96 days but they threatened that they will not contract with me. Then, I had to invite them to a working breakfast, and I told them: ‘I will send you the memo with no copy to the SOH, but if you don’t pay, the next memo goes with copy to the SOH.’ And debt turnover has improved.”*

Other preventive strategies also worked at ESEs. The manager of IB and IIC, for example, pointed out that the key was to have a very careful negotiation, so that both parties had complete clarity about the issues arising during the execution of the contracts. They emphasised that it was preferable to spend a lot of time during

negotiation, so that misunderstandings were reduced to its minimum in a later phase. In fact, these were the ESEs that had more stable relationships with ARSs and could afford to renew contracts smoothly. The other strategy, applied by IIF was to carefully select the ARSs they would contract with, so as to minimise the chances of opportunistic behaviour, as shown in chapter 6.

#### 7.8. Summary of findings

To summarise the findings of this chapter on RSI and contract incompleteness, the propositions set out at the beginning are tested against the reported findings:

- **Proposition 1:** There are relevant RSI in the relationships between ESEs and purchasers (SOH and ARSs).

This proposition is confirmed: the SOH contracts for the uninsured exhibited a high degree of dedicated capacity because a large share of ESE revenues depends on the utilisation of its capacity for services to the uninsured. The PAB contracts required some physical RSI but its impact was rather limited. Ambulance contracts showed no RSI, except for the fact that the SOH leads the utilisation of these assets, but a competitive market outside the SOH contracts made specificity irrelevant.

With regards to ARS contracts, capitation contracts, which involved a transfer of risk to the provider, created incentives for the ESEs to invest in P&P to reduce risk, a human-type of RSI.

- **Proposition 2:** The presence of RSI increases the transaction costs of relationships observed between ESEs and purchasers (SOH and ARSs).

This proposition is partially confirmed: Although present, RSI (dedicated capacity) did not create the risk that the SOH held ESEs up. Clearly, it would not make sense for the SOH to opportunistically extract a quasi-rent from the ESEs if it had to bail them out later, given that the SOH had a commitment to keep ESEs open. However, regarding ARS contracts and the presence of a human-type RSI (investment in capitated population), two responses were found: some ESEs perceived the risk of extraction of quasi-rents and preferred to underinvest in P&P. But other ESEs thought it was

strategically better to incur the investment to create patient loyalty and secure contracts with the well reputed ARSs. But in general, RSI were not found to be the major source of transaction costs as compared to contract incompleteness.

- **Proposition 3:** Contract incompleteness pervades the relationships between ESEs and purchasers (SOH and ARSs).

This proposition is confirmed. Contract incompleteness was caused mainly by uncertainty about the number of intermediate activities in the FGPP and the interpretation that ESE managers had on the average versus the cap for each FGPP category. The level of overall billings to the SOH was also a source of uncertainty both for those who exceeded the billing cap and those who didn't. However, lack of concern for best health quality outcomes reduced the problems of asymmetric information, difficulty to define the product, and observability/verifiability of the product; thus, quality compliance focused on observable/verifiable aspects of inputs and processes, which are easy to comply with. The difficulties in definition of the product were mainly found regarding the boundaries of the POS-S, which created opportunities for cost shifting both on the provider and the purchaser side; these blurred lines were not adequately dealt with in the contract minutes.

Regarding PAB contracts, uncertainty was largely absent and there were low information asymmetries due to the same reason explained above: lack of concern for outcomes and impacts, and a focus on observable/verifiable features of the product, basically inputs and sometimes process indicators. Regarding ambulance contracts, uncertainty was found to be low except for the normal operational risk associated with ambulance break down. Information asymmetries had no effect because the product was largely observable and verifiable, although concern for technical quality (safe transportation of the patient) was not evident and the same focus on input and process indicators reduced the need for bridging information gaps related to health outcomes.

On the side of ARS contracts, they showed a very different pattern. Contract incompleteness stemmed from three basic causes: 1) uncertainty regarding quantities demanded, 2) difficulty to define the product in terms of a) the boundaries of the POS-S, and b) considerations on medical necessity. Causes 1 and 2.a. were related to capitation contracts, whereas causes 2.a. and 2.b. were related to fee-for-service

contracts. In the case of fee-for-service contracts, endless discussions on the boundaries of the POS-S and medical necessity induced the parties to shift towards capitation contracts in the expectation of reducing these contracting costs. However, in the case of capitation, the incentives for skimping on care increased the risk of poor quality outcomes and consequently created other types of transaction costs associated with the control of inadequate responses to incentives.

- **Proposition 4:** Contract incompleteness is by itself a source of transaction costs of relationships observed between ESEs and purchasers (SOH and ARSs).

This proposition is confirmed: the effects of the abovementioned features of SOH-uninsured contracts on the costs of contracting were clear, mainly related to poor quality outcomes that remained unchecked, i.e., ESEs spent less than the optimal in terms of health outcomes because skimping on care was not a major concern for the SOH, and only the most outstanding cases (maternal mortality, nosocomial infections) were analysed. This is expected to cause a welfare loss, and it could be argued that this was a societal transaction cost. In addition, the costs of contracting were also relevant, not only regarding negotiations and follow-up, but also the direct costs of outsourced audits.

Regarding PAB contracts, the costs of these contracts were mostly related to design and negotiation but this reduced the costs of execution because a big effort was invested in clarifying *ex-ante* the largest possible number of contingencies. These contracts also involved the costs of audits carried out by outsourced auditors.

Regarding ARS contracts, a concern for quality outcomes was clearly reflected in the costs to be borne by ARS for poor outcomes that generated high-complexity events. Thus, ARSs could not ignore the importance of good quality outcomes, but these were difficult to define, observe and verify, and information asymmetries would make it very difficult to enforce compliance by the ESE, because most of the times it was difficult to attribute poor outcomes to specific errors of clinical management. Consequently, the relationships were burdened with transaction costs, not only related to sub-optimal investment in RSI but also to skimping on care, and distrust.

Other costs of distrust were reflected in ESEs' unwillingness to share information with ARSs that might be useful for the parties to improve risk management and increase the likelihood of best health outcomes.

- **Proposition 5:** The transaction costs of these relationships, stemming both from RSI and contract incompleteness, result into long-term relationships

This proposition is partially rejected. Regarding SOH contracts for the uninsured, follow-up committees could be said to be governance structures that reduce transaction costs associated with contract incompleteness, although their attributes were not strong enough to achieve such reductions. In fact, the long-term *ex-post* relationships that were observed were not the result of choices made by the parties to reduce transaction costs, but rather the result of an *ex-ante* bilateral monopoly. It can be argued that cooperative relationships were motivated by reasons other than reducing transaction costs. Regarding PAB and ambulance contracts, transaction costs had no effect in shaping the observed type of relationships.

With regards to ARS contracts, the transaction costs associated with contracts for level I services, arising from contract incompleteness and RSI, were more likely to be higher due to the lock-in as a result of compulsory contracting regulations. For level II and III ESEs, a more competitive environment led them to build long-term relationships to reduce transaction costs.

## **Discussion**

As was shown in chapter 2, one of the rationales for the PPS is to avoid the transaction costs of vertically integrated hierarchies. This rationale rests on the unproven assumption that the transaction costs of the contract-based relationship are lower than those of vertical integration. This point begs the question: is it true that transaction costs are lower in the PPS than in the vertically integrated structure? Although the research literature on TCE shows that quantitative methodologies would help to answer this question partially, it cannot be tested in this particular research because all the ESEs have the same contractual arrangement with the SOH. Nonetheless, an in-depth qualitative assessment of the case of Bogota helps to detect the sources of transaction

costs and their potential impact on the relationships between the trading parties, in a way that cannot be ascertained through quantitative methods.

Hence, the exercise of analytic generalisation requires consideration of the expected dynamics of transaction costs, based on theory, and comparison of expected dynamics with empirical findings to see if they confirm or reject the theory. It will shed light on the presence of RSI and contract incompleteness, their degree of intensity and how they influence the relationships to have the observed characteristics that have been described in chapter 6. This discussion is organised in two parts: first, a comparison between expected and observed types of relationships will be made. And second, the explanations for the effects of RSI and contract incompleteness on the observed types of relationships, or their lack thereof, will be analysed.

#### 7.9.Expected versus observed types of relationships

Chapter 6 described and analysed the observed types of relationships between ESEs and purchasers, in terms of the short-term/long-term continuum. It was found that relationships with the SOH were determined by the *ex-ante* bilateral monopoly and the fact that they are long-term was not the result of a choice of the parties. Thus, the relevant question is how cooperative these relationships were. With regards to ARSs, it was found that relationships between level I ESEs and ARSs tended to be less cooperative, because of the bilateral monopoly created by the compulsory 40% contracting with the local level I ESE. But relationships between ARSs and level II and III ESEs showed a pattern more similar to that predicted by TCE. Table 7.1. shows a summary of the findings of chapter 7, and a contrast between the expected type of relationship according to theory and the type that was actually observed.



**Table 7.1.**

**Summary of findings of sources of transaction costs and the expected vs. observed types of relationships**

		Type of RSI & cost	Components of contract incompleteness				Type of relationship	
			Uncertainty	Information asymmetry	Unobservability unverifiability	Difficulty to define the product	Expected	Observed
Uninsured		Dedicated capacity (High cost)	High: number of intermediate activities; hitting the billing cap	Reduced because of no concern for best health outcomes	Low because focus on outputs	Boundaries of POS-S with SOH	Long term <i>ex-post</i>	<i>Ex-ante</i> bilateral monopoly, full cooperation to less cooperation depending on other factors
PAB		Physical assets (small cost)	Low	Reduced because of no concern for best health outcomes and impact	Low because focus on outputs	Low because focus on outputs	Short term <i>ex-post</i>	<i>Ex-ante</i> bilateral monopoly, full cooperation to less cooperation depending on other factors
Ambulances		No	Low	No	Low	No	Short-term <i>ex-post</i>	Long-term
SS	Capitation	Human (P&P) (medium cost)	High: quantity demanded	High because ARS concerned for high-cost outcomes due to skimping on care	High	High because of boundaries of POS-S, medical necessity	Long-term <i>ex-post</i>	Level II and III: trust-based, win-win, make long-term <i>ex-post</i> more likely. Level I: long-term <i>ex-post</i> because of compulsory contracting regulations; less-cooperative
	FFS	No	No	Induced demand	Low	High because of boundaries of POS-S, medical necessity	Long-term <i>ex-post</i>	Level II and III: trust-based, win-win, make long-term <i>ex-post</i> more likely

## 7.10. Explanations for the effects of RSI and contract incompleteness or its lack thereof

As shown in chapter 2, Hart, Shleifer and Vishny (1991) suggest that when quality is noncontractible, private firms are more likely to skimp on quality. Eggleston and Zeckhauser (2001) point to the same expected results of contracting out for health services, while Chalkley and Malcomson (1998) obtain the same outcome but they classify contractors between a purely selfish and a purely benevolent continuum. Accordingly, what can be expected from ESEs is that, being public organisations, their behaviour would be benevolent because they are not interested in making profits. However, it is not completely obvious from the findings that this is their actual behaviour.

Going back to the issue of theoretical models of hospital behaviour, it is possible that ESEs behave like a profit-maximising firm, or a not-for profit firm that acts as a for-profit, as proposed by Feldstein (1993, p. 237). If they behave like purely selfish contractors, the room opens for the two most powerful groups (administrators and doctors) to use the hospital to satisfy their private interest. The outcome will depend on the relative power and the financial motivation of each group (Muurinen, 1986). The less selfish the most powerful group, the more likely the ESE will behave like a benevolent contractor. But the findings of this research warn that it cannot be assumed that the fact that a hospital is public, automatically makes it benevolent. This reinforces the experiences of Sub Saharan Africa countries, summarised by Hanson et al (2001). Context differences between the UK case and those of developing countries make it necessary to carefully consider the effects of incentives and the necessary role of regulation.

These considerations are at the root of the explanation of the effects of RSI and contract incompleteness on transaction costs and the observed types of relationships. In the for profit sector, it is obvious that there are incentives for the parties to minimise transaction costs, and this creates the specialised governance structures that are observed between spot contracting and vertical integration. However, in the public sector this is not so obvious. On the one hand, public entities lack a single-minded objective function that aligns purchaser and provider, and on the other hand, the outcome of the exchange, good quality health, is difficult to observe and verify, and is

plagued by information asymmetry, not only between purchaser and provider but mainly vis à vis enrollees.

It is worth emphasising that the research questions did not considered objective functions initially, but they emerged during data analysis. Exploring interview data for references to objective functions, it was shown that the SOH may have a diffuse objective function: maximise social welfare, subject to the budget constraint. However, it could also be said that its objective function is to keep ESEs open, to avoid political turmoil, as pointed out by an SOH officer. Or it could also be simply to reduce spending, as will be discussed in chapter 9. On the side of ESEs, it could be argued that their objective function is to maximise revenues. The fact that profits could not be privatised points toward revenue maximisation to generate surpluses to make capital investments or expand inputs, which further increase output and, hence, revenues. These two functions imply clearly identifiable outcomes, yet not necessarily best health outcomes, just patient satisfaction. ESEs can also have the objective function to maximise social welfare if the agenda of the most powerful group (administrators or doctors) is aligned with this objective. A final objective function could be the manager's individual career concerns, which are not necessarily aligned with best health outcomes, but in the case they are, it is likely that social welfare is maximised. It could thus be inferred from these considerations that the ESE objective function is endogenous, i.e., it is chosen by the key decision-maker depending on their personal objective function.

The key point in these reflections is that maximising social welfare through best health outcomes is not necessarily the objective function of ESEs and the SOH. Accordingly, it is not necessarily true that, even in the presence of RSI and contract incompleteness, the choice of a given type of relationship is determined by the objective of minimising transaction costs. For example, unobservability and unverifiability of health outcomes make it possible to skimp on care, which makes minimisation of transaction costs less necessary because overall costs can be reduced through skimping. When ESEs cut intermediate activities within FGPP categories, or when they put access barriers to capitated enrollees of the SS, they end up reducing overall costs at a much larger rate than if they spent time and effort trying to reduce transaction costs. Obviously, it requires monopoly power so that the purchaser cannot search for alternative providers. In fact this is the case, as has been shown, for SOH-ESE, and ARS-level I ESEs relationships. However, in the case of an ESE that shows a welfare maximising

objective function (either because of personal values, or for career concerns), its manager will be concerned to minimise transaction costs, and at the same time minimise production costs without affecting quality.

It is useful to contrast this behaviour with that of fee-for-service contracts, where it is clear that ESEs invest a lot of effort to reduce denials. Given that denials partly have their origins in purchasers' (SOH and ARSs) efforts to shift costs to each other, and in second-guessing clinical decisions, ESEs have adopted proactive strategies like accumulating experience to solve issues of POS-S's boundaries and developing clinical guidelines agreed upon with purchasers to reduce the issue of medical necessity. Some of them also set up sophisticated billing processes while others outsourced these processes to billing experts (although with limited success). Perhaps the costs of contracting related to denials are very relevant to the extent that they reduce revenues to the ESE, so if the ESE manager wishes to show higher revenues it is wise to devise mechanisms to assure them.

Another point that shows ESEs' concern for reducing the costs of contracting is the shift from fee-for-service towards capitation, in the expectation that prospective payments would reduce such costs. Regarding SOH contracts, the same could be said considering that the costs of contracting, particularly regarding billings, are lower in the prospective FGPP mechanism than in the fee-for-service payment. This line of reasoning follows Bech and Pedersen's (2005) point that payment mechanisms are alternative ways to reduce transaction costs within a given type of relationship. However, the perverse incentives of prospective payments (skimping on care), the problem of counterbillings, and ESE unwillingness to share information with purchasers, generate other types of transaction costs that are not adequately dealt with by the parties just by shifting from one payment mechanism to another. Hence the rationale to consider poor risk management skills as a source of transaction costs, even though they might be more correctly considered a production cost.

As the findings show, the parties lack concern for best quality health outcomes. As suggested by McPake et al (2003), quality decreases as opportunistic behaviour of providers lead them to concentrate effort on observable and verifiable features of products, but not on the really important issues of technical quality. The additional finding of this research is that purchasers also lack concern for technical quality. It

could be argued that this lack of concern makes it less likely that the parties invest effort in reducing transaction costs, because they just cut production costs. But the search for a payment mechanism that reduces transaction costs reveals that the parties are actually concerned about reducing such costs.

One point that could explain the difference in approaches to transaction costs is where financial risk is located. In the case of SOH contracts, risk at the ESE is reduced by the expectation of bailouts from the SOH and by the aggressive cutting of intermediate activities. More fundamental risk-mitigating factors are the bilateral monopoly and no-exit option mentioned in the chapter 6. But the SOH faces a budget constraint that is determined by the District Authorities. This budget constraint is not 100% hard, and how much below 100% it is, depends on local authorities' objective functions. But in the end, whatever the hardness of the constraint, the SOH is the risk bearer. Depending on the hardness of the budget constraint, it could be said that the SOH is more or less tolerant to higher or lower transaction costs.

However, as shown in chapter 6, ESE managers still have the option to cooperate or not, and this choice might be explained by the transaction costs associated with contract incompleteness. But it was evident in that chapter that managers who decided to cooperate had interests other than reducing transaction costs, like, for example, increasing the chances of being reelected for the next office term. It could also be argued that career concerns or even personal relationships between a given manager and the Secretary of Health were more influential on the decision to opt for cooperation. But in the case of ESEs whose objective function was aligned with maximising social welfare, they were more likely to cooperate, as illustrated by the case of IB. Such cooperation has an effect (purposively or not) to reduce transaction costs.

On the side of ARSs, it is more obvious that they face a hard budget constraint (even the nonprofits), although their objective function is likely to differ between nonprofits and for-profits. This makes ARSs the single risk bearers of the costs associated with the POS-S, which makes them more interested in reducing not only production costs (medical loss ratios and insurance administration) but also transaction costs. Their continuous search for modifications in payment mechanisms, follow-up strategies, and the consideration for vertical integration as a governance structure of last resort, reveals their concern for transaction costs. In their effort to reduce transaction costs, ARSs

obtained two different responses from ESEs: among those facing competition (level II and III), some ESEs exploited their competitive advantage to build trust and generate win-win situations, which were more likely to result in long-term relationships, as was shown in chapter 6. Consequently, trust-based, win-win, long-term relationships between ESEs and ARSs are more likely to reduce transaction costs. But those protected from competition (level I ESEs), were less responsive because they knew they would not be substituted. However, within this last group, some ESEs were responsive, either because the managers' objective function was aligned with social welfare, or their career concerns entailed incentives for a more cooperative relationship with ARSs.

ESEs willing to cooperate, because of market exposure and/or alignment of objective functions, were more likely to engage in long-term relationships but as an *ex-post* choice of both parties, to reduce transaction costs. How negative would this effect be on competition? It could be argued that the effect is not too bad because the parties usually look for each other based on the reputation that each has built. So, it is expected that the small numbers game does not create room for opportunism because the parties have been selective regarding with whom to establish a long-term relationship. Those unwilling to cooperate will, instead of reducing transaction costs, increase them. Hence the double-way effect of cooperation as a choice to reduce transaction costs or non-cooperation as a source of these costs.

Transaction costs in both SOH and ARSs relationships stemmed basically from contract incompleteness. RSI were not relevant for SOH contracts, and they were relevant only for capitation contracts with ARSs, but not all the ESEs reacted to the possibility of the extraction of a quasi-rent as predicted by theory. But the single presence of contract incompleteness was a reason strong enough to be considered as a source of transaction costs and to search for specialised governance structures to reduce them. This is consistent with what has been reported in the literature on health care systems in the US (Robinson, 1999). In this country, during the early years of managed care, HMOs chose vertical integration as a measure of last resort when they faced large difficulties to control spending within a very limited budget constraint.

However, this is a major departure from the mainstream literature on TCE, as TCE gives paramount importance to RSI, its role in transforming an *ex-ante* competitive market into a bilateral dependency, and the risks of quasi-rent extractions. In

mainstream TCE, contract incompleteness seems a complementary issue that increases the risk of *ex-post* hazards and maladaptations. However, in the health care literature, it seems that contract incompleteness in and of itself justifies the search for a specialised governance structure that reduces the costs associated with contracting. Information asymmetries play a key role as has been shown here for ARSs, when the purchaser is really concerned with health outcomes (at least those that reduce its profit-maximising motive), because the incentives of prospective payments increase the likelihood of skimping on care.

## **Conclusions**

The findings shown in this chapter demonstrate that there are RSI and contract incompleteness in the relationships between ESEs and SOH/ARSs, but their effects on the observed types of relationships are more obvious for the relationships with ARSs. Although RSI are present in the relationships with the SOH, their effect as a source of transaction costs is nonexistent, because the SOH has no reason to extract quasi-rents that will backfire on itself. Contract incompleteness did not create transaction costs in terms of unobservability or unverifiability of clinical outcomes, because of lack of concern for health outcomes. It made it easy for ESEs to comply with input and process indicators, because they are observable and verifiable, and products are easy to define. However, contract incompleteness did create transaction costs in terms of the inability to adequately control the responses to perverse incentives of payment mechanisms, namely skimping on care (in FGPP and capitation) and demand inducement (in FFS contracts).

The predictions of theory are more evident in the relationships with ARSs. However, investment in P&P for capitated populations, which is a human type of RSI, faced two different approaches: one conservative, as predicted by theory, and one non-conservative. Strategic choices and an objective function to maximise social welfare explain the non-conservative approach to investment in P&P. But the role of contract incompleteness explains by itself the need for better control of the provider, which is restricted by the 40% compulsory contracting with the local level I ESE.