10. DISCUSSION AND POLICY IMPLICATIONS

10.1. Overview of the findings

The separation between provider and purchaser, as proposed by the NPM approach, clearly rests on a TCE-based rationale. Otherwise it would not make sense to search for an alternative governance arrangement for the relationships between these two parties, with the aim of resolving the maladaptation problems of the vertically integrated structure. This same rationale can be applicable to middle income countries like Colombia, since it would be expected that efficiency gains would be obtained through a different governance structure.

However, the types of relationships that were observed in this research, and their explanations, suggest that TCE does not adequately predict all the findings. Relationships between ESEs and the SOH showed dedicated asset RSI, but they were not relevant to shaping the relationship. Their long-term characteristic was more the consequence of an *ex-ante* bilateral monopoly that results in an *ex-post* mutual dependence mediated by a contract, and not the result of a choice between the parties to minimise transaction costs. Within this context, it was found that cooperation reduced transaction costs whereas less cooperation was a cause of transaction costs. It was not clear though, that cooperation aimed at reducing transaction costs or rather aimed at other goals like career concerns.

But uncertainty regarding the number of intermediate activities in each FGPP category (mostly due to poor risk management) or regarding the possibility to reach or not to reach the billing cap, did create grey zones that could not be solved adequately with the contract minute. Complexity of contracted products also led to 1) skimping on intermediate activities, inducement of final activities and cream skimming, based on information asymmetries; 2) possibly higher costs in terms of poorer health outcomes due to lack of concern for technical quality, and 3) cost shifting to other purchasers due to the grey boundaries of the benefit packages. It was not clear though, whether cooperation aimed at reducing transaction costs or rather aimed at other goals like career concerns.

Regarding ARS contracts for level I services, human RSI were found in terms of investment in capitated populations, but it was only relevant for transaction costs among some ESEs that perceived the risk of opportunism on the ARS side. The remaining ESEs invested in this type of RSI as a strategy to guarantee patient loyalty and weed out undesirable ARSs. Uncertainty in capitated contracts was due to poor risk management to deal with primary risk (random peaks in demand or adverse selection) and technical risk, but also due to counterbillings. The effects of the components of complexity of the contracted products were the same as those shown for FGPP, namely skimping on care, higher costs of poor health outcomes and cost shifting. For level II and III, no effect of RSI was found, and all the transaction costs stemmed from contract incompleteness. FFS payments allowed for less uncertainty on the ESE side and difficulties to define the boundaries of benefit packages led to cost shifting. However, a more competitive environment due to greater purchaser's choice of provider prompted both purchasers and providers to engage in long-term relationships to reduce the transaction costs of contract incompleteness, as predicted by theory. Reputation effects played an important role in this case.

A major finding of this research is the fact that contract incompleteness was in and of itself a cause of transaction costs that justified the search for a specialised governance structure. This was clear even in the absence of RSI. The mentioned reactions of providers to the incentives of payment mechanisms made clear that most transaction costs caused welfare losses in terms of: 1) excess growth of expenditures, 2) capture of rents and 3) burden of disease. Regarding excess growth of expenditures, this was the behaviour that was observed during the FFS era of the SOH contracts, which in fact motivated the change towards FGPP. Capture of rents could have happened with both retrospective and prospective payments, because the ESE had a strong incentive to generate surpluses to increase output, and, ultimately, maximise revenues. And the burden of disease could have been affected, at least in terms of not achieving a larger impact with the available resources. Yet lack of concern for health outcomes led the parties to dismiss poor health outcomes as an externality.

The lack of effect of RSI on governance structures has been largely overlooked in the health economics and health services research literature. Except for Allen (2002), Ashton (1998), Palmer (2001) and Guinness (2006), where RSI are explicitly analysed, other studies focused only on contract incompleteness, ignoring the RSI component. For

example, Marini and Street (2006) focus on the administrative costs of contracting by comparing a preexisting payment mechanism with the newly adopted one, but they do not explicitly acknowledge the role of RSI. Preker et al (2000) propose a tool for "make or buy" where the criteria for the decision are measurability and contestability; their omission of uncertainty as a component of contract incompleteness adds to the fact that RSI are not considered at all. Loevinsohn and Harding (2005), although not from an explicit TCE perspective, also focus their analysis of successful contracting out ignoring the potential role of RSI.

It could be said that the fact that RSI do not play any role in public organisations is a reason to ignore them in the analysis. But it is a different thing to say nothing about them than to recognise their lack of effect, as RSI are a key element in TCE. Their limited effect in health care is all the more important because the PPS rests on a TCE rationale. In addition to the role of RSI to shape or not the relationships, the point is the role of contract incompleteness itself, even in the absence of RSI, to justify alternative governance structures to reduce transaction costs. Robinson (1999) has claimed that purchaser-provider relationships in health care lack RSI and this makes vertical integration unjustifiable. However, the effect of information asymmetries was found to be large enough in this research and it could be argued that this effect is probably much larger in health care than in any other industry. This larger effect would justify governance structures closer to the end of vertical integration (yet not necessarily complete vertical integration) than to the end of spot contracting.

Indeed, another major finding is that transaction costs, originating mostly in contract incompleteness, although present did not shape the relationships as expected from theory (except for level II-III ESE-ARS relationships). This point suggests that ESEs have different motivations for their performance, and they are not necessarily interested in economising on transaction costs, hence their lack of interest in a specialised governance structure that allows for such economisation. It could be argued though, that SOH and ARSs' search for alternative payment mechanisms was driven by a transaction cost economising objective on the lines of Bech and Pedersen (2005). However, the evidence shown here clearly indicates that changing the payment mechanisms just created other types of transaction costs. In fact, there seemed to be no real gain within a contractual relationship framework that has been frozen by regulations, i.e., the impossibility of vertical integration both on the SOH and ARS sides.

The most plausible explanation for the lack of effect of transaction costs within the observed relationships is the lack of adequate allocation of property rights, because although managers can be declared residual claimants, they cannot privatise profits, either short term surpluses or long term higher value of hospitals (because they are not traded in the market). Strikingly enough, managers were all the time talking about profits, no matter they actually reinvested profits in expanding capacity to maximise output and revenues.

Inadequate allocation of property rights is one key underlying factor for the low-powered incentives of public organisations. In addition, the fact that losses are absorbed by the SOH (due to its commitment to keep the safety net whatever it takes, to show concern for the welfare of citizens), raises a soft budget constraint problem. It is also reflected in the regulations for compulsory contracting for ARSs with level I ESEs, largely a protectionist measure taken by the SOH to prevent the outcome of a negative sum competition and further ESE losses that would have to be bailed out by the SOH itself. On the side of ARSs, their revealed preference for their own vertically integrated primary care network clearly reflects their interest in economising on transaction costs through a hierarchical governance structure. This vertically integrated structure would allow them to economise on transaction costs, i.e., have a better control of the perverse incentives of payment mechanisms and, at the same time, to have access to information on costs, quality and illness, all of it having high strategic value. But the compulsory contracting regulation prevents ARSs from achieving such economies in transaction costs and reduces the need for level I ESEs to search for such economy.

Besides, the lack of a clear and unidimensional objective function and a clear and single minded principal makes it less likely that granting residual claims to managers automatically results in better hospital performance and improved social welfare. As a general rule for experience goods, information asymmetry vis à vis the patient allows for providers (and also purchasers) to focus on objectives other than improving quality at lower cost, without the patient realising the poor quality of the care received. It also allows for providers and purchasers to reduce production costs rather than transaction costs, thus making less pressing the need for a specialised governance structure.

This fact also points towards a weakness in the case for patient choice as a mechanism to improve hospital performance. Patient-driven competition as a mechanism to create market exposure seems at least an oversimplification of the hospital industry dynamics. As commented by Propper et al (2006), the limited effect of patient choice on the performance of health care markets partly stems from the fact that information on providers' technical quality is hard to use for consumer's exercise of choice. This expectation is confirmed by the findings in chapters 6 to 9: the information asymmetry problem weakens incentives for quality improvement, so production costs can be compromised for the sake of output or revenue maximisation, and transaction costs can be ignored.

This lack of predictive power of TCE for the public sector is hardly surprising if it is kept in mind that, from its beginnings, this theory focused its analysis on the private for profit firm (Coase, 1937), and further developers kept the same focus of interest. Only during the late eighties and early nineties the claims for a new role of the state under the rubric of the New Public Management appealed to TCE rationales, although not as explicitly as the literature from the for profit sector. Thus, the arguments of TCE, which are better predictors of the transaction cost economising behaviour of the for-profit firm, show less of a predictive power for the public sector. This causes a gap between what technocrats view as a desirable new role of the state (based on the evidence from the private sector and the privatisation of SOEs) and what actually happens when politicians and bureaucrats implement the technocrats' recommendations.

One plausible link between TCE and the findings of this thesis is that proposed by Frant (1996). He argues that the high powered incentives of reelection make politicians induce allocatively inefficient outcomes when large information asymmetries pervade a given government function. The solution he proposes is to de-politicise the activity by shifting to a separate agent that cannot be controlled by the politician. In addition, he argues that to avoid the effect of incentives it is necessary to appeal to the rule strategy, on the same lines of Tirole (1994), of avoiding regulatory capture and collusion through rigid rules. However, it was shown that autonomy status is not necessarily associated with lower politician influence, because managers still have the decision space to play the politicians' game of patronage. And, contrary to Frant's argument, ESE managers are exposed to the high-powered incentives of the market, i.e., a partial residual claimant status. Thus, unlike his application of TCE to the public sector, specifically the

case of independent authorities (like central banks or regulatory commissions), the evidence of this research suggests that applying his argument to the case of autonomous entities seems to be inadequate.

By contrast with TCE, public choice theory seems to exhibit a better predictive power, based on its ability to reflect reality. On the side of politicians, their objective to assure reelection is intimately linked to their attitudes toward the welfare of their voters. As put by Hsiao (2000): "Good health broadly shared is intrinsically valued in all societies" so politicians will focus their efforts on keeping health systems working as close as possible to the level that is perceived by voters as adequate, no matter it is inefficient or of poor technical quality; hospitals are a very visible part of this construct, hence politicians' interest in shielding them from the risk of closure. The evidence from this research strongly supports this assumption.

In addition, politicians want to control jobs and resources to serve their constituency or for rent-seeking purposes, which was also found in this research. Although it was shown that the three district administrations previous to 2004 successfully fended off political patrons, it was clear that these patrons redirected their focus of attention to the local level, on some ESE managers that were willing to play their game.

Regarding the SOH bureaucrats, their behaviour is also determined by the politics of Mayoral elections. As Mayors are publicly elected bureaucrats, they are expected to show the same behaviour of politicians expecting to be reelected or at least to transfer power to one of their party's allies. This creates a strong incentive for incumbent Mayors to appeal to the wishes of their constituents, and keeping hospitals open is a means to the end of staying in power. Therefore, bureaucrats will follow the Mayor's policies, reflecting the strategic implications of the Mayor's decision-making. This prediction is supported by the findings, when the SOH officers acknowledged that the hidden agenda was that hospitals ought not to be closed down. Nonetheless, the three administrations between 1995 and 2003 showed an unusual willingness to make policy decisions with long-term effects (i.e., beyond the political cycle of a given Mayor), but without trespassing a threshold of tolerance for poor performance of ESEs.

On the side of bureaucrats' selfish behaviour, it could be argued that ESE managers' career concerns were not necessarily aligned with the need to economise on transaction

costs, hence their lack of concern for specialised governance structures. Additionally, the bilateral monopoly with the SOH and the protectionist regulations with ARSs for level I services made the need for specialised governance structures even less pressing, no matter that large transaction costs pervaded these relationships. It could be argued that ESE managers' search for alternative payment mechanisms as governance structures was aimed at maximising revenues rather than at reducing transaction costs. In fact, they took the necessary steps to avoid denials and counterbillings, and at the same time overreacted to the perverse incentives of the payment mechanisms, taking advantage of information asymmetries vis à vis the patient.

Another factor that helps to create this specific context is the fact that autonomy has not evolved as it was supposed to from the beginning of the 1993 reform. Insufficiency of the economy to generate the necessary resources turned a transitional situation (towards universal coverage and equal benefit packages between the EBS and the SS) into a permanent one. Autonomisation also fell in between, resulting in a halfway progress that has been interrupted indefinitely. The role of the SOH as insurer of last resort, and the ESEs as providers of last resort, make the parties to be locked-in indefinitely.

In addition, autonomy has opened room for managers to engage in undesirable responses because they do not face the tight controls of the private sector stemming from adequately assigned property rights and accountability devices including the market for corporate control. Although major failures have also been seen in the private sector (e.g., the Enron case) as a result of the imperfect ability of corporate governance to control corporations (Hermalin and Weisbach, 2003), it can be expected that these accountability devices will be more effective when the markets in which they operate are more competitive. In their absence, the SOH has to keep a set of control knobs that substitute for both the direct control of a hierarchical structure and the discipline of input, capital and output markets. This could be argued to be a complementary explanation for the halfway autonomy that was found. Incomplete transfer of decision rights was seen as a strategy to keep control over certain key decisions to reduce the risk of ESE misbehaviour. In addition to the contract itself, the participation in the board of directors, the use of Performance Agreements, the tutelage function and the more basic regulatory function as local health authority provide strong control knobs to make sure ESEs behave correctly. Informal control knobs are also relevant, such as peer pressure,

personal relationships and the politics of reelection of managers or their promotion to better jobs.

The creation of a market for insurers in the Colombian health system design was an incentive for ESEs to reduce the oversight of the SOH by increasing their share of revenues from ARSs. If universal coverage and equal benefit packages were achieved, the SOH would be able to control ESEs only indirectly, through the boards of directors, the tutelage function, and the regulatory function. The extent to which this hypothetical situation would lead to better or worse outcomes in terms of autonomy (due to less direct control of the SOH), is a purely speculative exercise. However, it could be speculated that due to its public nature, ESEs are unlikely to respond to market forces in the same way as private for profit firms. This would justify that the SOH keeps its stewardship role to avoid the undesired responses that were pointed out in chapter 8, and to allow for a cautious introduction of market forces.

10.2. Further general reflections

Several general reflections can be added from the findings reported, beyond the ones that have been thoroughly addressed in the previous chapters.

Differences between contracting in and contracting out.

Although the PPS can be said to be the result of a search for a governance arrangement to reduce transaction costs, the NPM approach addresses two different situations as if they were the same. As noted in chapter 2, contracting *out* is considered different from contracting *in*, at least within the context of this thesis. In contrast, the arguments proposed by Walsh (1995) simply compare vertical integration to contracting, overlooking the different implications of contracting *in* (i.e., the PPS) and contracting *out*. When contracting out, the government calls for external firms or organisations to place bids for contracts, and the selection process and final granting of the contract to a given contractor can be competitive or contestable. By contrast, contracting in, although not in principle, in practice implies little or no competition or contestability. The dynamics of the relationship are thus determined by the historic background of the root organisation both parties were part of, so the mutual dependence that is created makes linger some traditions of the past, as has been shown by Goddard and Mannion (1998)

for the NHS. The struggle of the SOH to get rid of those traditions is clearly shown by the findings, and the prospects of reducing them to zero are rather limited, given the ESE status of provider of last resort, its public characteristic and the politics around it.

These lock-in dynamics were also found in the country experiences discussed in chapter 3, which clearly point out the differences between contracting in and out. For the cases of Sub-Saharan African countries show that the extent of transfer of decision rights is severely limited, and contracts are rather ineffective. In other experiences like those of China, Eastern Europe/Former Socialist Economies, Pakistan and Peru, contracts were not even reported or did not exist at all. It could be said that, except for the case of China, in all the reviewed experiences the government is still highly tied to its hospitals and the effectiveness of contracting (when it exists) is rather limited.

Ownership role as a complement to contracts

The incomplete granting of property rights could be argued to be a strategy to deal with *ex-post* maladaptations of ESEs to contract terms. Whereas in the private sector the only formal link between client and provider is a contract (plus informal links like reputation effects or the threat of exit), it was shown that in the SOH-ESE relationships the SOH acts both as a client and as an owner. If the only lever the SOH had to control the ESE were the contract, it is quite likely that the undesirable responses to autonomy would have been very difficult to control. However, the findings show that keeping its ownership role allowed the SOH to limit the scope of undesirable responses to autonomy, specifically in those areas more likely to lead to a negative sum outcome.

Capitation contracting: a step backwards to block grants?

The evolution of payment mechanisms in the SS toward capitation, partly fuelled by the compulsory contracting regulations, seems to be a step backwards towards block grants. This is paradoxical because the idea with the contract-based relationship is that the contract serves as an accountability device to make sure that the best value for money is bought on behalf of patients. However, the consequences in terms of reduced flows of information and perverse incentives towards skimping on care certainly reduces the ability of contracts to achieve a better coordination of care and the maximisation of health value to patients. On the side of the SOH contracts the same problem is taking

place, to the extent that the FGPP system reduces the flow of information towards the purchaser. As has been repeatedly argued in the previous chapters, ESEs behaviour as revenue maximisers leads them to take money in by any means, including taking advantage of monopoly power. Accordingly, the shift from retrospective to prospective payments weakens incentives for ESEs to be responsive to purchasers' needs in terms of coordination of care and best health outcomes.

Across-the-table governance structures

It has been pointed out by Ashton (1998) that the heterogeneity of health services, because of wide differences in terms of frequency, uncertainty, asset specificity, and measurability, imply different transaction costs, which imply different governance structures. In the case of Bogota, the SOH-ESE contracts for curative services showed no concern for these differences, as the contract is the same for all activities. It is based on the FGPP categories, and the per-case payment applies to any type of service. According to Ashton, it would have been expected to find different contracts for different types of services, or at least for some gross categories of them, depending on the specific sources of transaction costs shown by each category.

Perhaps the SOH is still in a learning process whereby sooner or later it will realise that this variability of transaction costs justify different contracting mechanisms and their corresponding specialised governance structures. However, given the finding that the parties are not interested in reducing transaction costs, it is unclear if they would find it attractive to explore alternative governance structures.

Individual profile of managers as predictor of success

The overall outcome of hospital autonomy can be said to be highly influenced by the individual profile of each manager in place. This profile is determined by individual skills and leadership. Career objectives and the decision space created by autonomy interact with the individual profile of manager to lead to a good or bad outcome. This same effect has been reported by Jakab and Preker (2002b) for Eastern Europe countries, Eid (2001) for Lebanon, and Gouveia et al (2006) for Portugal. The observed outcome in the Bogota case is that ESEs apparently increased their output and sustainability was improved, and at the same time the SOH decreased the rate of growth

of budget. On the side of the ESEs that showed the largest increases in output and improvements, it could be said that *ceteris paribus*, differences in leadership itself explain part of the differences in outcomes. The question if autonomy was relevant or not, remains to be answered because there is no way to build a counterfactual within the 22 ESEs, as they were all granted autonomy at the same time.

10.3. Methodological issues and limitations

The limitations of this study have to be clearly stated. On the one hand, the generic limitations of qualitative methods do not allow for direct generalisation of the findings to other settings. However, it is also clear that the case-study method is not aimed at statistical generalisation but at analytic generalisation, i.e., the testing of theory (Yin, 1994). Therefore, lacking a sampling logic, it is not necessary to think of how the findings apply in other settings. Nonetheless, the reader might feel tempted to extrapolate findings to other settings, which warrants some caveats before attempting such generalisations. The most important limitation in this sense is the fact that Bogota is a unique case for hospital reform in Colombia. This uniqueness stems from the following facts:

- Being a District, Bogota acts both as a municipality and as a department. It implies a big advantage when compared to other sub-national administrative units of the government, because it keeps the purchasing function for services for the uninsured at all levels of complexity. This common funding of all levels aligns incentives and makes coordination easier. However, the typical situation in the rest of the country is that the municipality is responsible for level I services and the department is responsible for level II and III. This creates an additional fragmentation and consequently an additional fence to shift costs from one side to the other, because each administrative unit receives its monies separately from the other, and these resources are not pooled to align incentives. This has proven a big challenge in the rest of the country.

- Being the capital of the country, Bogota is by far the richest area and has the highest formal economy. This means that resources to purchase the services for the uninsured and to enrol people in the SS are less scarce than in the rest of the municipalities of the country. It also implies that it tends to attract more skilled

professionals and the availability of resources to improve human capital is much higher. It could also be argued that the supply of qualified potential managers is better than in the rest of the country.

- The three district administrations between 1995 and 2003 were recognised as less prone to political patronage, which paved the way for ESE managers and SOH officers to give priority to technical aspects of management and policy implementation. The meritocratic selection and appointment of managers was pointed out by interviewees as a remarkable feature of the previous administrations; this feature had been relaxed during the new administration which started in 2004.

These factors and the interactions among them, clearly determine a better performance of the SOH, in terms of stewardship of the local hospital network, than is likely to have been the case elsewhere in Colombia.

A second limitation of this study is that most of the dynamics described are endogenous to a particular style of governing of the Mayor of Bogota. These dynamics can change and create a completely different dynamic in major respects. An interesting question would be how these dynamics changed after a new administration came to power with a different view of political patronage and a different ideological stance with regards to health care. Given that this research focused on a cross-sectional analysis of the situation in 2002, which reflected the effects of the current local political situation, a longitudinal analysis should be undertaken to isolate the effects of changes in government orientation. It would also be possible to compare hospital reforms in countries with very different settings to contrast the findings on a cross-sectional basis.

A third limitation that has to be taken into account is the fact that interviewees did not necessarily have to reveal their true motivations or negative perceptions in their answers. In fact, they might have answered strategically, emphasising some things to their convenience and minimising or hiding others, as well as giving answers more likely to please the interviewer. This respondent bias was partially reduced by triangulation, i.e., getting the same information from different sources. For example, ESE managers' negative views on ARS openly contrasted with ARS officers' views on them, and the complementary view of the SOH partially solved the stalemate of

interpretation. Unfortunately, an in-depth analysis of thorny subjects like corruption, incompetence or political patronage goes beyond the scope of this research and could not be easily clarified with the research methods used. This, obviously, leaves unanswered questions regarding the influence of these thorny subjects on the findings reported.

A fourth limitation has to do with the fact that some transaction costs arise from behaviours that cannot be easily contrasted to determine their magnitude. For example, what is the optimal level of investment in a capitated population against which the actual investment can be compared? Or what is the optimal quantity of interventions against which the actual quantity provided can be said to be demand inducement or skimping on care? The fact that these behaviours are rooted on opportunism and bounded rationality, and contract incompleteness cannot adequately deal with them *ex-post,* clearly justifies their classification as transaction costs. However, the challenge to measure them to carry out a more robust test of the theory points to some structural limitations of the currently available empirical methods and the consequent difficulty to thoroughly test the theory. These limitations are obviously reflected in this research.

To assess the relationships between purchasers and providers, specifically in the public sector, it would be necessary to analyse a wide range of health systems in order to capture variations in types of relationships and the transaction cost variables. In the case of Colombia, it is not possible to achieve such variation, because all the hospitals are autonomous, so there is not a hierarchical organisation to use as comparator. The scope of this research proposal did not allow for an inter-country comparative analysis, and a case-study of the Bogota hospitals was adequate to improve understanding of the selected research topic.

Regarding analytic generalisation, it was clear that the case-study method allowed some gaps in theory to be detected that are relevant for further research, as shown below. The research design described in chapter 5 seems to have worked adequately for this purpose. The initial insights of theory that guided the interviews were helpful to evidence the gaps that were later fleshed out more accurately with the propositions set out for chapters 6 to 9. However, further issues were detected after the interviews when it was difficult to go back to interviewees to elaborate them with more precision. Issues like objective functions of the agents and degree of cooperation were analysed only *ex-*

post through the existing transcripts, but they would have been certainly better ascertained in a second round of interviews. Yin (1994) points out that a feature of case-study methods is their flexibility to re-shape the enquiry without losing the focus on the main research question. Unfortunately, the emergence of these issues after finishing field work did not allow for such detailed enquiry.

The exercise of testing theory through the case-study method deserves a deeper reflection of what it means to test a theory. As shown in chapter 5, such testing is not aimed at ruling in or out a whole theory, because it is unlikely that a single theory explains all the findings. Rather, a set of theories should be used to build the best explanation of the findings, which in turn are also used to detect the gaps in the theories used, so as to streamline them and increase their explanatory power. This study allowed for such type of exercise, and as shown below, there were evident gaps that are largely overlooked in the literature. The exercise also allowed reasonable arguments to be built for further research questions and for the strengthening of existing theories.

10.4. Implications for policy

A general implication that flows from the local to the international context is the fact that information asymmetries, apparently much larger in health care than in many other industries (Dranove and Satterthwaite, 2000), cannot be underestimated when making predictions based on evidence from other sectors or industries. In addition, the fact that most of health care is either financed, provided, or both, by the government, gives more importance to the predictions of public choice theory. The implications of information asymmetry demand a larger effort on the government to bridge them and make health care market approaches work better. It does not mean that consumers must become experts, but at least some information on outcomes should be made available for the public to make decisions and make patient-driven competition meaningful.

Propper et al (2006) cite the experience of the New York coronary artery by-pass graft report card system as an example of how such information can lead to better outcomes. Although a conspicuous exception in the general field of information on outcomes for patient choice, it gives the idea that such information exerts positive effects that remain to be better understood. Perhaps a small reduction in the information gap would force purchasers and providers to adopt behaviours that improve quality and reduce costs, and

discourage them from ignoring the effects of transaction costs just by reducing production costs at the expense of quality.

Clearly, variations in institutional settings influence the extent of over-response of providers to high-powered incentives in the presence of information asymmetries and their negative consequences. Hsiao (2000) points out that weaker ethical constraints in developing countries are likely to leave room for stronger responses. Liu and Mills (2005) report such over-responses in China, and Mills et al (2000) report that private providers in Thailand were more likely to skimp on care when paid on a capitation basis. Skimping strategies were similar to those found among ESEs in Bogota: restrictions to medical decision-making, restrictions to access to inputs, and lower intensity of service use.

Another institutional variable is the ethos of public service that shows differing degrees of intensity in different countries. Le Grand (2003) suggests that a tradition of public service that rewards "knightly" behaviours can substitute for the imperfection of markets for social services, as had been argued for the NHS in the first post-war decades. Although this variable was not ascertained in this research, it could be said that some managers' explicitness about social service objectives for their ESEs is a symptom of such public service ethos. What the findings did not provide evidence for, is how pervasive that ethos was among ESE managers and SOH officers.

10.4.1. Implications for policy at the local level

It has been shown that the FGPP generates transaction costs that are just different from those of retrospective payments. Therefore, apparently payment mechanisms are not the solution in the Colombian setting to reduce transaction costs as proposed by Bech and Pedersen (2005). But the transaction costs of contract-based relationships also cast doubts on the attributes of this governance structure. The inadequate allocation of property rights seems to be at the root of the disfunctionalities of the public sector, so the solution to the transaction costs between purchasers and providers cannot be searched for in alternative governance arrangements.

Nevertheless, the problem of controlling budget growth could be addressed in a very cynical way: assuming hospital care is inefficient and ineffective no matter the payment

system or the governance arrangement, it is unrealistic to expect that hospital policy would contribute to improve hospital performance. In fact, it was shown that contract-based relationships also show *ex-post* maladaptations and their costs cannot be said to be lower or greater than those of vertical integration. The same can be said of the payment mechanisms. Accordingly, the problem for hospital policy comes down to how the SOH keeps hospital expenditures at bay or even ratchets them down. It is here where the Purchaser Provider Split provides its most pragmatic consequence: if hospital survival is left to depend on contracts, it will be easier for the financing authority (the SOH) to restrict budgets by turning back the responsibility on the managers. However, as pointed out in chapters 6 to 9, the mutual dependence relationship can only be changed on a piecemeal basis. In fact, it could be argued that the evolution from vertical integration to autonomy and from FFS to FGPP are steps towards a future situation where ESEs will be more directly responsible for their survival. This process was progressing in the same direction during the three district administrations before 2004, but a major shift in policy and ideology in 2004 makes it unlikely that the same direction is followed.

The compulsory contract regulation for ARSs (40% of premium with level I and 10% with level II and III) is obviously a protectionist measure for level I ESEs, inspired by the need to avoid double payments on the SOH side, because if ESEs do not get the contracts with ARSs, their idle capacity has to be paid by the SOH anyway. However, the complaints of ARS officers and the SOH itself regarding the lack of control over patient care and coordination of care, is a clear transaction cost that has to be solved by searching for specialised governance structures to reduce them. The negative consequences of ESEs' revenue maximising behaviour seem to be hardly solved by changing the payment mechanism, so changing the governance structure seems more likely to solve the transaction cost problem. The space needed for this search requires more flexibility on both purchaser and provider side to try alternative governance arrangements.

Given that purchaser and provider are locked-in by the compulsory contracting regulations, the ESE has no incentive to search for a governance structure that reduces transaction costs. This is clearer when the situation with level I ESEs is compared to that of level II and III ESEs, where more choice on the ARS side forces ESEs to be more responsive to the demands of purchasers. The weaknesses of contracts in the ARS-

ESE relationships in the level I context, and the contrast with relationships at the level II and III context, suggests that alternative governance structures would be likely to reduce the transaction costs stemming from contract incompleteness. As shown in level II and III contexts, a concern for win-win situations prompted some ESEs to build trust with ARSs they considered desirable trading parties. It could be argued that this trust substitutes for other governance structures, as proposed by Holmstrom and Roberts (1998), but the fact is that the relationship is shaped in a transaction-cost economising way. Such building of trust is unlikely to be found in a relationship where one of the parties has no exit option.

10.4.2. Implications for policy at the national level

The issue of information asymmetry mentioned above can be addressed at the national level by a policy to create and disseminate information on health outcomes. Although it is a big challenge, some information is better than no information. In this sense, the Ministry of Social Protection released in November 2006 a quality ranking of hospitals based on indicators of structure (Ministerio de Proteccion Social, 2006). Although it is still full of weaknesses, it could be argued that this ranking will spark reactions among providers and the public that will help improve its validity and in the long run will prompt purchasers and providers to invest effort to improve their performance in terms of outcomes. Although Propper et al (2006) are sceptical of the effects of information on patient choice and, ultimately, on competition among providers, they acknowledge the positive effect of the New York State experience.

On the side of insurance markets, the problem of segmented benefit packages (i.e., the gap in the medium complexity interventions) creates an evident tension due to the incentives for cost-shifting. This misalignment of incentives has proven very difficult to overcome. Given that resources to fill this gap and to enrol the uninsured are unlikely to be mobilised, it would be necessary to think of alternative policies. One possibility would be to give people with chronic conditions a comprehensive package, because these are the ones most badly hurt by the problems of coordination. In addition, these people should be enrolled with a risk-adjusted premium, given their inherently higher primary risk.

10.4.3. Implications for policy at the international level

The greatest implication for international policy from the evidence of this research is that it suggests that policies aimed at exposing public hospitals to market-like incentives are likely to result in other types of inefficiencies. This research suggests that public hospitals seem to respond inadequately to market incentives, due to information asymmetries and inadequate allocation of property rights. Their behaviour as revenue maximisers in the Colombia case suggests that their responses to incentives can be disproportionate. This over-response can have intermediate positive outcomes, such as higher output (like in Pakistan (Balal, 2006)) and higher revenues (like in China (Eggleston and Yip (2004)) , but these can come at a cost in terms of overall allocative inefficiency and health outcomes, as shown by Eggleston et al (2006). This lack of adequate response on the hospital side casts a serious doubt on the issue of residual claimant status, which has been shown by Harding and Preker (2003) as a key to generate incentives to improve performance.

Public hospitals' status as provider of last resort, and health authorities as insurer of last resort, create market conditions very different from those of the rest of insurers and providers serving the less poor in society. This fact strongly restricts the effects of policies based on evidence from the for profit sector or from previous waves of privatisation of SOEs and public utilities. Consequently, the safety-net status of public hospitals makes it unlikely a commitment can be realised to cost-effective allocation of resources away from tertiary care and towards primary care, outpatient and public health interventions. In addition, given the political visibility of hospitals, politicians are likely to take advantage of the public's perception of the role of hospitals in society to keep inefficiency at a level that is tolerated by their constituencies.

Given these points regarding the limits of markets to yield improvements in efficiency and quality, the wave of marketising reforms in health care systems raises a question about its adequacy and feasibility, at least for developing countries. In developed countries such as the United States, with market-based systems, some hospitals carry out a social function as safety net providers, and part of this function consists of providing uncompensated care whose costs are shifted to other purchasers (Eldenburg and Krishnan, 2003). Once competitive pressures increase, this implicit subsidy and the social function of hospitals start to disappear (Propper, 2006). This problem is even bigger for developing countries, because their economies are unable to support

comprehensive benefit packages and universal insurance. It means that the poor and the high-risk individuals are not adequately protected. It is unlikely that pure market incentives can solve this problem, and the high-risk and poor individuals will still generate uncompensated care for providers. There is where public hospitals still play a key role.

The problems arising from market failure and safety-net status make it unlikely that external incentives and regulations will lead to meaningful changes in hospital behaviour ?when exposed to market forces. In addition, it seems that a key determining factor is the individual profile of hospital managers and their career concerns. If these two are aligned with better hospital performance and improved social welfare, it is more likely that the hospital improves. Otherwise it is not necessarily the case. Thus, the policy challenge is how to select the individuals with those characteristics, or to create a cadre of public managers, and to attract them to the public hospitals. This obviously entails competing with the private sector where better salaries are paid and better long term career opportunities can be pursued, which restricts in one more way the scope of hospital reforms.

Additionally, it could also be the case that poorly skilled or corrupt managers are selected for ESEs. In these case, autonomy does not adequately safeguard the hospitals from the expropriation of rents. The boards of directors can be part of corrupt processes that trickle down from corrupt local authorities and the fact that a hospital is autonomous does not prevent this from happening any more than it does in a vertically integrated structure. Although Bogota, at least during the period of study, was notable for the absence of corrupt practices, it cannot be said this is the rule in developing countries. It cannot either be said that an ethos of public service prevailed in Bogota or that its level was high enough to fill the information gaps of market approaches.

The evidence shown by this research implies a strong critique of the hospital autonomy policy, which suggests that marketising reforms do not necessarily cause the effects observed in other sectors or countries. Key trade-offs have been identified between the low-powered incentives of hierarchical relationships and the high-powered incentives of market-based ones. In addition, the inherent disfunctionalities of the public sector that create a different set of low-powered incentives open room for stronger effects of the market-like incentives. Thus, it seems that there is no such thing as a straightforward

solution to the problem of incentives in public hospitals that strikes an adequate balance between efficiency, quality and equity. The contribution of this research to a better understanding of the dynamics of purchaser-provider relationships could be said to be one step more towards a robust theory of public hospital behaviour that allows for a better design of policy.

10.5. Further research questions

The process of analytic inference used in this research to test theory (see chapter 5) identifies some gaps in TCE and other NIE theories. The most basic question that remains unanswered is the actual objective function of public hospitals in developing countries. This question underlies the development of a more specific and sophisticated modelling of hospital behaviour that would allow for a better guide to, and interpretation of, research. Although this research suggests that revenue maximisation was the objective function of ESEs, it can be the result of the prevailing incentives and regulations. Following McPake and Archard (2002), this objective function could be endogenous to the system design, which means that a more fundamental analysis has to be undertaken to ascertain the very basics of the behaviour of public hospitals in developing countries.

Along these same lines, a better understanding of the effects of allocation of property rights to hospital managers, and in general to public bureaucrats, deserves more theoretical development. Hart (1995) underscores the separation of residual control rights and residual income rights in public organisations. It could be hypothesised that this separation makes it more likely that bureaucrats enjoying residual claims invest time and effort to increase residual revenues instead of investing to increase residual controllers' payoffs, i.e., the welfare of the community.

One key further research question for the streamlining of TCE theory is the role of contract incompleteness as a cause of transaction costs in the absence of RSI. The wide information asymmetries between patients and providers, and between purchasers and providers, create wide room for opportunistic behaviours that generate the risk of expropriation of rents as hardly seen in other sectors of the economy. This fact, in and of itself, justifies the search for specialised governance structures including outright vertical integration between purchaser and provider. Although this argument has been

dismissed in the US (Robinson, 1999), it is still not very clear whether different stages of evolution of local health care markets justify governance structures other than arms-length contractual relationships.

A more general concern for future research on TCE is the adequate application of this theory to the public sector. It was shown that the impossibility of adequately allocating property rights weakens one key assumption of TCE theory: that the parties align the costs of transactions with the attributes of governance structures, in a transaction-cost economising way. The other key assumption that is violated in the public sector, particularly in the case of the PPS, is that contracting starts as a voluntary decision of two parties that want to engage in a win-win exchange; in contrast the parties in the PPS are *ex-ante* locked-in (Goddard and Mannion, 1998).

The application of TCE to the public sector that was proposed by Frant (1996) does not adequately solve the challenges posed by hospital autonomy, because he focuses his analysis on the de-politicisation of public functions by creating independent authorities so that the high-powered incentives of re-election do not apply. This rationale has been shown to work well for central banks and for regulatory commissions, where de-politicisation of decision-making allows long-term commitments to be made that would not be possible within the short-term political cycle of re-election. Unlike these cases, hospital autonomisation has been shown in this research to be rather irrelevant in terms of de-politicising decision making and reducing rent-seeking, because what autonomy induces is hospitals' exposure to market-based high-powered incentives. This area of research remains largely unexplored.

It is important for future research to distinguish between contracting *in* (i.e., the PPS) and contracting *out*. In fact, a comparison of the agency problems and transaction costs between the two settings would be very helpful to clarify the possibility that the government can consider the complete contracting out of hospital services just as happens for Medicare or Medicaid in the US. This would make it necessary to consider outright privatisation of public hospitals, although not politically feasible in the current state of affairs, as one of the policy options besides corporatisation or autonomisation.

On the side of statistical inference, further research questions for generalisable findings are also derived from this research. Regarding the observed lack of effect of RSI and

contract incompleteness in shaping relationships, and the role of contract incompleteness as a source of transaction costs in the absence of RSI, it is necessary to do research aiming at more generalisable results in term of larger cross-sectional studies that span different autonomy and country settings, as well as longitudinal studies that control for the effect of changes of government leaders.

Another field of interest is that of the effects of autonomisation on efficiency and quality. Although this is the typical question that many researchers have tried to answer, it could be the case that the research questions are not adequately addressed due to a limited understanding of the process of autonomy itself. For example, the fact that autonomy is not a dichotomous state but a continuum of a large number of variables, makes it difficult for researchers to classify a hospital as autonomous or not.

11. CONCLUSIONS

The main aim of this research was to describe and analyse the role of transaction costs in shaping the relationships between purchasers and public hospitals in Bogota. It has been shown that transaction costs were present in purchaser-provider relationships. In the case of SOH-ESE relationships, transaction costs stem from contract incompleteness but not from RSI. In some cases of ARS-SOH relationships for level I services, transactions costs stem from the fact that some ESEs do not invest enough in promotion and prevention in the expectation of opportunistic behaviour on the ARS side. But for other level I ESEs this expectation does not deter them from investing because they use it as a strategy to strengthen patient loyalty and weed out undesirable ARSs. In the case of ARS-ESE relationships for level II and III services, transaction costs stem only from contract incompleteness and no RSI were found to be sufficiently relevant.

However, the presence of transaction costs did not shape the relationships as expected from theory, except for ARS-ESE relationships for level II and III services. SOH-ESE relationships were long term but as the result of a pre-existing bilateral monopoly that resulted in an *ex-post* mutual dependence, given that ESEs are providers of last resort to the poor, and the SOH is the insurer of last resort. This lock-in or mutual dependence situation opens room for ESEs to decide if they want to cooperate or not, because under certain circumstances non-cooperation was seen as a strategy to obtain quasi-rents from the SOH. Although openly non-cooperative behaviour was not found but rather less-cooperative behaviour, it caused transaction costs in the relationships with the SOH. On the other hand, cooperation reduced transaction costs. This same logic applies to ARS-ESE relationships for level I services, because the compulsory contracting regulations gave monopoly power to level I ESEs and their willingness to cooperate was less than that of level II or III ESEs. For these latter two levels of care, the fact that they were not protected by compulsory contracting regulations made them more competitive and their behaviour was more similar to that predicted by theory, i.e., they invested time and effort in creating win-win situations and building trust as the basis for long-term relationships to reduce transaction costs.

Regarding the relationship with the SOH, the mutual dependence is also reflected in the fact that hospital autonomy has not evolved into a total autonomy status. This is partly a consequence of the impossibility of achieving universal coverage with a comprehensive

benefit package, which causes the SOH to keep its role as purchaser at the same time that it is owner of ESEs. However, it is also a consequence of the risks that granting unfettered autonomy to a public institution would entail. Given that ESEs lack the accountability devices of input, output and capital markets, the SOH has to keep control over certain decision rights as a mechanism to safeguard ESEs from the expropriation of rents or from managers' outright poor judgment.

According to the findings summarised above, TCE has little predictive power except for level II and III ESE relationships with ARSs. The lack of adequately allocated property rights provides a strong explanation for the lack of effect of transaction costs on the observed types of relationships. This implies that granting residual claimant status to an ESE is not enough to create the incentives to economise on transaction or production costs.

Among the New Institutional Economics theories, the one that better explains the findings is Public Choice. It was found that the SOH concern was the sustainability of ESEs, and not necessarily their best efficiency or the highest quality. In addition, the impossibility of carrying out large-scale reforms as opposed to rather piecemeal improvements, and the lack of risk exposure (i.e., poor decision-making did not lead to hospital closures), could be argued to be a reflection of politicians' desire to keep the hospital network as a proof of their concern for the welfare of citizens.