
TRANSACTION COSTS IN HEALTH SUPPLY CHAIN AND HEALTH INSURANCE REGULATION

Ronaldo Fiani

Economics Institute – Federal University of Rio de Janeiro
Av. Pasteur 250 , sl. 110 – Instituto de Economia– Urca – RJ – Brazil – 22290-240
E-mail: Fiani@ie.ufrj.br

Abstract

This paper argues that transaction costs are significant in the relation between healthcare supply chain and health insurance. After briefly reviewing the literature on healthcare costs and health insurance, the nature of healthcare supply chain is discussed in order to identify possible sources of transaction costs in the relation with health insurers. The conclusions thus obtained are applied to succinctly evaluate some recent trends of the Brazilian health insurance industry. The objective of the research is to provide more information about transaction costs in healthcare provision and the consequences of such costs to health insurance, which may help to improve health insurance regulation.

Key words: *Transaction costs, healthcare costs, health insurance regulation in Brazil.*

JEL Codes: D23, I11, I13, I18

TRANSACTION COSTS IN HEALTH SUPPLY CHAIN AND HEALTH INSURANCE REGULATION

That risk and uncertainty are, in fact, significant elements in medical care hardly needs argument. I will hold that virtually all the special features of this industry, in fact, stem from the prevalence of uncertainty.

Kenneth J. Arrow

RONALDO FIANI

INTRODUCTION

Notwithstanding the vast literature concerning health insurance regulation, the scope of issues usually discussed is usually confined to the relation between insurer and beneficiary, discussing almost only informational asymmetries like adverse selection and moral hazard problems, as well as rules to mitigate those problems. The most common results of such approach are proposals combining consumer protection and solvency controls over the insurers as an ideal mix of regulatory policy.

However, such approach to health insurance regulation ignores the transaction costs between the supply chain of health services and health insurers. That limitation is particularly harmful to health insurance regulation, for it ignores both the problematic interaction between healthcare providers and health insurers and how regulation may disturb that interaction. The objective of the discussion presented here will be to provide more information about transaction costs in healthcare provision, which may help to devise better regulations. Accordingly, this paper is divided in three sections besides a very brief conclusion. The first section presents what seems to be the most general hypothesis on the cause of healthcare costs increases: Baumol's 'cost disease' hypothesis, discussing its problems for neglecting institutional aspects of healthcare costs.

The second section presents the conventional approach on health insurance regulation, emphasizing the absence of an analysis of the relation between health insurers and healthcare supply chain. The third section deals with transaction costs in healthcare supply chain and its relation to health insurers, also discussing very briefly the consequences of Brazilian regulation for the transaction costs affecting the relations between private health insurers and healthcare providers. Because of extension restrictions, a very brief conclusion ends the paper.

1. The problem of healthcare costs increase: Baumol's "cost disease" thesis

The author who has devised what seems to be the more general economic explanation for the production costs pressures in the healthcare sector (as well as other sectors based on personal services) was William J. Baumol (2012; 1996; 1993; 1990).¹ Perhaps the most outstanding feature of Baumol's explanation is that it simply ignores the presence of transaction costs in healthcare costs, reducing the problem of healthcare cost increases to productivity differentials. According to Baumol, the same kind of cost pressures identified in the healthcare system would also be identified in those sectors with a high content of personal

¹ See also Heilbrun (1997) for the origin of Baumol's ideas.

services, which encompass the educational system, automobile repairs, performing arts, lawyers' services, etc.,² all of them suffering from what Baumol calls a 'cost disease'.

According to Baumol's cost disease thesis, economic activities involving personal services in a high degree are consequently labor-intensive activities, not only in the sense that there is a high proportion of labor in the total sum of factors of production applied in those activities, but also in the sense that labor input in personal services are usually comprised of complex skills and know-hows, making it very difficult to substitute capital for labor in personal services. Those skills and know-hows are necessary because personal services are supplied in a case-by-case basis: physicians prescribe treatments according not only to the characteristics of the disease but also to the specific patient conditions, lawyers devise the best strategies to each case according to its specificities, etc. So the worker in those activities must discern which solution is the best for each case, and that in a large extent also helps precluding standardization and the substitution of capital for labor as ordinarily. Finally, to make productivity gains even more difficult, time reductions in the execution of those activities frequently are not associated to productivity gains but to quality reduction. Abbreviating the time necessary to perform either a surgery or a Mozart's quartet most probably would result in a total failure and not in a productivity gain.

As it is known since Adam Smith, to increase the capital-labor ratio is essential to raise productivity and reduce unitary costs. In the case in question, the technological peculiarity of personal services supply results in such activities presenting gains of productivity in a much lesser degree than industrial sectors, where the substitution of capital for labor is much easier and productivity gains grow exponentially.

So far, Baumol's cost disease thesis amounts only to an explanation of why there is a divergence in productivity gains across the economy; or better, it amounts to an interpretation of why personal services would tend to show gains below the average of the economy. Then comes the second part of Baumol's thesis: in the long run wages must grow at the same rate in all activities in the economy, otherwise the activity in which wages have grown behind the general growth rate will suffer from a shortage of labor supply.³

Given that wages grow at the same rate across sectors but productivity grows much slower in personal services than in industrial sectors (not only because there is no substitution of capital for labor in a significant extent in personal services activities, but also because productivity in industry has risen at astonishing rates since the last century), then it results that costs in personal services rise well above costs in the rest of the economy, or as Baumol (2012) writes, as healthcare becomes more and more expensive computers become cheaper and cheaper. Those fast rising costs affect healthcare, education, car repairs, lawyers' services, art performances etc. According to Baumol, those rising costs are a kind of long run technological fate and there is not much to be done to prevent it. Baumol emphasizes the generality of the cost disease phenomenon by collecting data about costs increases in personal services in different countries, in order to show the pervasiveness of the problem of cost

² Baumol's original concern with the 'cost disease' started much earlier indeed, and was at first only related to performing arts, and not to healthcare. See Heibrun (1997, 91).

³ It must be noted that Baumol does not suppose that wages have the same level in all economic activities, for the value of the marginal product of labor is different in each activity and in each job in the same activity. What he says is that the relative level of the wage structure in the economy has to be preserved in order to keep supply and demand of labor in equilibrium in each activity in the long run.

growth *independently of the institutional frame*: “(...) the problem of rising healthcare costs is not overcome either by private or public sector operation” (1993, 19).

Thus, to Baumol the continuous increase in health insurance may be attributed in a large degree to a health costs increase, for there are evidences that labor-saving technological change in insurance, as well as in other financial branches (basically through computing and information systems) have significantly reduced insurance administrative costs (Hecht, 2001). Those reductions in administrative costs, however, would have not compensated for the increase in healthcare costs (Baumol, 1990).

Notwithstanding its simplicity and generality, as well as his efforts to fundament the claim of productivity differentials empirically (Baumol 1993; 2012), Baumol’s ‘cost disease’ thesis has not avoided severe criticism. To mention just a few critics of Baumol’s ‘cost disease’ thesis, Ferris and West (1996) point to the fact that Baumol does not take into consideration the possible effects of organizational changes on the productivity of low-productivity sectors; Cowen (1996) argues that the attribution of a stagnant productivity to performing arts is a mistake for it does not make allowances for the quality improvements that have taken place in those arts; Triplett and Bosworth (2003) found that services productivity was the same as the productivity of the rest of the U. S. economy in the 1977-1995 period (and in fact that it was accelerating), and Gallouj and Savona (2009) point to the usual severe problems in defining and measuring the output of services, which affect not only Baumol’s analysis but have also posed obstacles to a better analysis of innovation in services.

But even if one accepts Baumol’s ‘cost disease’ thesis without criticism, some puzzles of central importance to his thesis still remain: First we have the puzzle of the high level of administrative costs in healthcare, especially in face of the aforementioned general reduction of administrative cost in financial and insurance services produced by the increased use of computers and information systems. In effect, some researchers found nothing less than 31% of administrative costs in the total costs of the U. S. healthcare system (Woolhandler, Campbell, and Himmelstein, 2003). Even those evaluations being subject to methodological issues (Kahn et al., 2005), there is a rather broad consensus that the share of administrative costs in total healthcare costs is significant in the U.S. Comparing data on administrative workers in the U. S. and Canada, Cutler and Ly (2011, 8 emphasis added) have found that:

Perhaps the most troubling difference between the U. S. and Canadian healthcare systems is the differential amount spent on administration (...). For every office-based physician in the United States, there are 2.2 administrative workers. That exceeds the number of nurses, clinical assistants, and technical stuff put together. (...) Canada, by contrast, has only half as many administrative workers per office-based physician.

The situation is no better in hospitals. In the United States there are 1.5 administrative personnel per hospital bed, compared to 1.1 in Canada. (...) *On top of this are the administrative workers in health insurance. Health insurance administration is 12 percent of premiums in the United States and less than a half that in Canada*⁴ (...).

The finding that administrative workers cost amount to an expressive proportion of total premiums in health insurance is a very relevant fact when one considers that: (a) as

⁴ See also Davis, Schoen, and Stremikis (2007).

shown before, there has been significant productive gains in insurance administration thanks to growing incorporation of computing and informational systems; (b) techniques of managed care were first devised and applied in the U. S. in the 1970's, which has accumulated since then a voluminous literature and a large debate on such procedures to contain health costs.

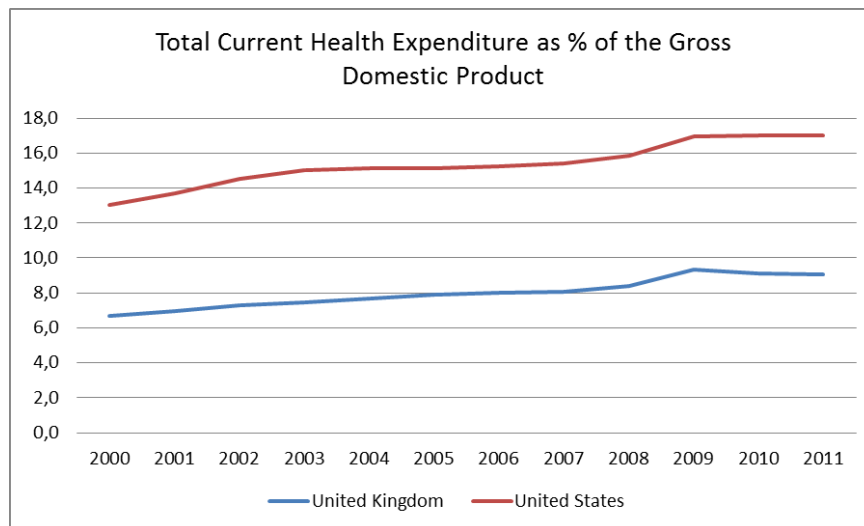
It seems so that not only administrative costs in health insurance have grown despite the aforementioned productive gains, but they have grown notwithstanding the creation and further development of managed care techniques, and it seems to be a general pattern, or general enough to be identified in countries with very different characteristics and regulatory institutions. It will be then argued here that: (1) both the high level of administrative costs and its occurrence in different countries can be explained by transaction costs, which is completely absent from Baumol's analysis; (2) managed care techniques, which themselves help to explain administrative costs increase in health insurance, must be understood as an effort to cope with the high transaction costs in the sector.

The second puzzle is the difference in the *levels* of healthcare total expenses in each country. In fact, notwithstanding the fact that healthcare costs increase everywhere (as Baumol has correctly emphasized), there are *significant and persistent differences* in healthcare costs levels across countries, as shown in Graph 1 where the polar cases of U.S. (private and not universal healthcare) and United Kingdom (public and universal healthcare) are contrasted.

In Graph 1 one can see that not only the levels of healthcare expenses as a percentage of gross domestic product in countries with so different institutional frames are very distinct, but also that such difference seems to be increasing in the last ten years. If Baumol was right and healthcare cost pressures were a *pure* technological phenomenon, then different countries (at least developed ones, which are able to adopt the best technology) would show the same level of costs independently of their institutional frame. But Graph 1 presents a quite different picture. One could argue that the higher level of U.S. expenses would be the result of providing healthcare to a greater proportion of U.S. population than in the United Kingdom, but the opposite is true. In fact, while there is a significant number of Americans that have no coverage at all,⁵ since 1948 the United Kingdom has provided universal health services for its population through its National Health Service. Actually, the United States is an exception concerning developing countries in relation to universal healthcare coverage (McKee et al., 2013).

⁵ According to the U. S. Census Bureau, in 2011 a share of 15.7% of all Americans had no coverage at all (http://www.census.gov/hhes/www/hlthins/data/historical/HIB_tables.html). That amounted to more than 48 million people.

Graph 1



Source: OECD Health Data 2013

The conclusion is that there is significant evidence on the impact of the institutional frame on healthcare costs. It seems that hierarchies – specifically public sector management and provision of healthcare – do reduce healthcare costs level. If the institutional frame is relevant and healthcare costs are not exclusively a technological phenomenon as Baumol supposes, then there are good reasons to incorporate transaction costs in the analysis of health economics. But before proceeding to discuss transaction costs in healthcare provision, the conventional theory of health insurance regulation must be considered.

2. The conventional approach on the regulation of health insurance

There is a broad consensus among economists that insurance is subject to severe informational problems, especially in the insurer-insured relationship. That consensus has resulted in a voluminous literature analyzing insurance informational problems through principal-agent lenses, as some combination of moral hazard and adverse selection. But few contributions have extended the informational problem analysis beyond the insurer-insured relationship, notwithstanding the publication of Kenneth Arrow's seminal article (Arrow, 1963) on health economics fifty years ago (from which the epigraph of this paper was borrowed), where Arrow presented health informational problems as rather severe. The answer may be in the fact that there are few studies investigating the sources of transaction costs in the supply chain of healthcare. So, before proceeding to the discussion of transaction costs in healthcare supply chain it is necessary to consider briefly the conventional view on health insurance problems and its regulation.

The conventional view is well illustrated by contributions like Preker et al. (2012), where the authors study the problem of private voluntary health insurance regulation as an appropriate balancing of three kinds of regulation: solvency regulation (for, as they sell insurance, companies in the sector are required to meet financial reserves standards), prudential regulation (aiming to anticipate adverse impacts provoked by financial crises on health insurers), and consumer protection (given the lack of information by consumers on the scope and nature of rights and obligations in health insurance contracts). There is no further analysis on how health insurance interacts with the healthcare supply chain. The problem of health insurance is accordingly described as a problem of excessive demand by the insured.

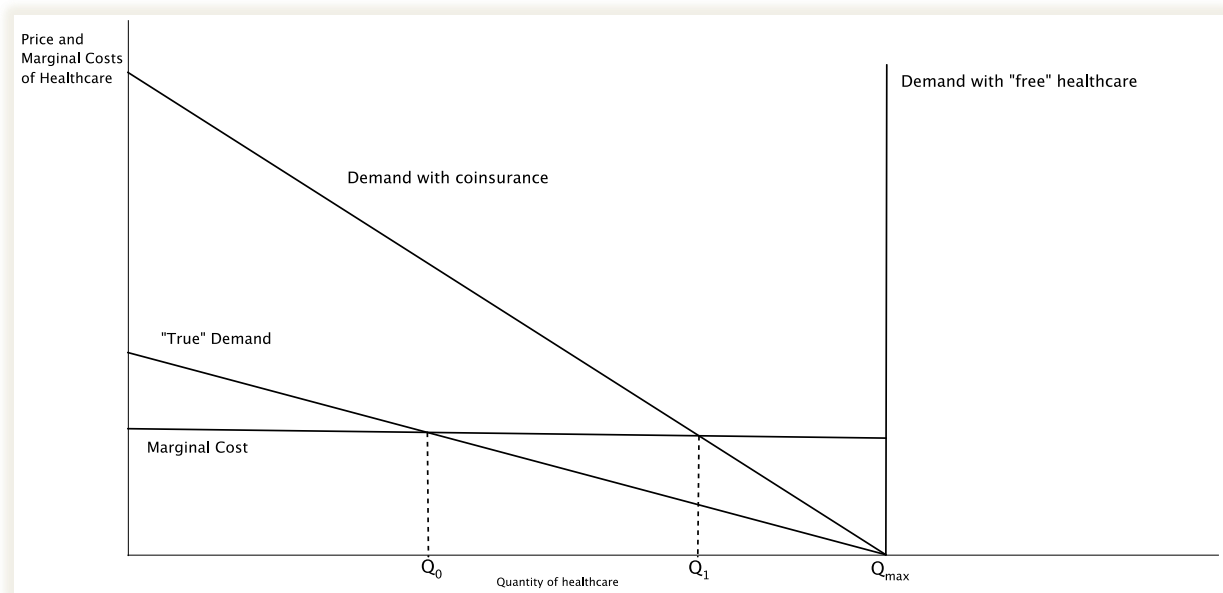
October 07-08th, 2013

Center for Organization Studies (CORS)

USP (University of São Paulo); FGV (Getúlio Vargas Foundation); Insper (Institute of Education and Research); UFBA (Federal University of Bahia); UFRJ (Federal University of Rio de Janeiro) and UFSCar (São Carlos Federal University)

Zweifel, Krey and Tagli (2007) extend that approach to developing countries, considering that coinsurance produces a new healthcare demand curve more vertical and above the “true” demand for healthcare, as in Figure 1.⁶ The demand for healthcare would be perfectly inelastic at its “maximum” quantity if healthcare were “free” (it means, without charge to the demander of healthcare).

Figure 1 – The Effect of Coinsurance on Healthcare Demand According to Zweifel, Krey and Tagli (2007)



Source: Zweifel, Krey and Tagli (2007, 60).

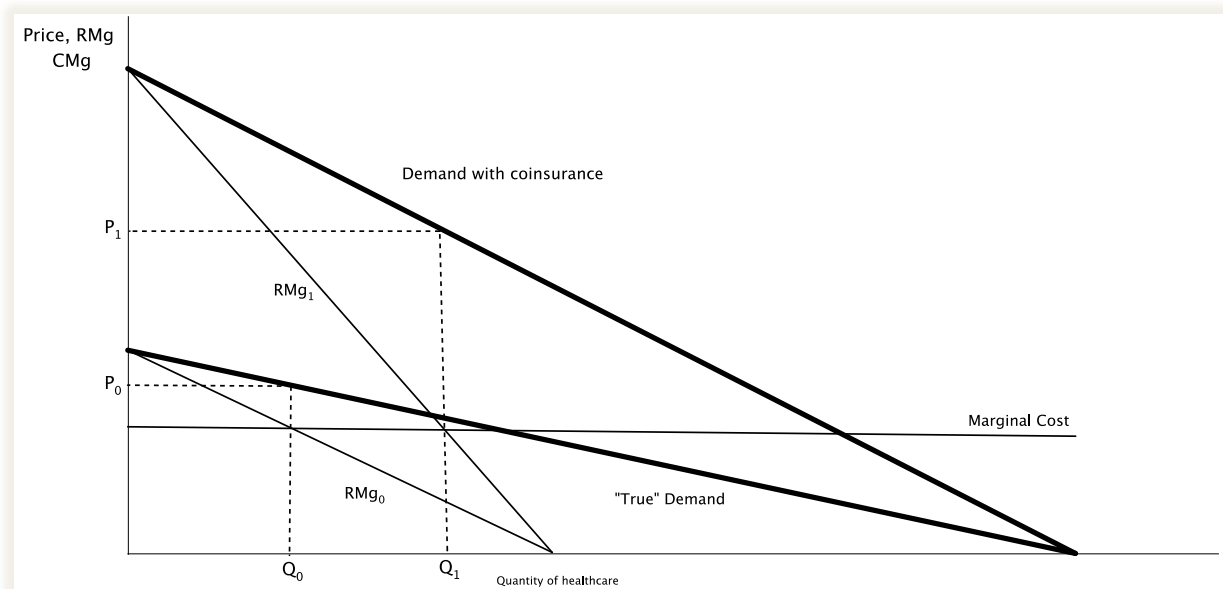
So, even if the healthcare provision were perfectly competitive the effect of health insurance would be to increase the demand of healthcare from Q_0 (its “true” level) to Q_1 (its level with health insurance), both quantities determined as usual by the equality of price (defined by each demand curve) to marginal cost. In the situation that healthcare were “free” to the individual, it means, if healthcare were supplied by a public provider and funded by taxes the level of healthcare demand would be maximum (at Q_{MAX}). Consequently, if healthcare provision were perfectly competitive the problem would be just an “excess” of healthcare demand measured by the difference $Q_1 - Q_0$.

The scenario gets more complicated if healthcare provision is not perfectly competitive. In Figure 2 marginal revenue curves were added to the demand curves of Figure 1, in order to represent the imperfect competition case. Then the “excess” of healthcare demanded is less than in the case of a perfectly competitive healthcare market, but now the price stays above marginal costs. The result is a mix of demand excess and price increase in

⁶ The idea is that insurance (at least to some extent) reduces adverse wealth effects associated to health problems. So, with the usual hypothesis of risk-averse individuals, the more an insured person is covered against health problems, the more his/her consumption of healthcare increases (of course after considering the expenses of the insured with insurance payments).

relation to what would happen if the demand were the “true” one even with imperfect competition ($P_1 - P_0$).

Figure 2 – The Effect of Insurance Coverage on Imperfect Competitive Healthcare Provision According to Zweifel, Krey and Tagli (2007)



Source: Zeifel, Krey and Tagli (2007, 79).

There are two problems with that analysis usually applied to health insurance. The first is that the demand for healthcare is not like ordinary demand, as the demand for shoes. A consumer may autonomously decide how many shoes she/he would buy according to her/his preferences, given personal budget and prices. That usually does not happen with healthcare provision. Normally the kind and extent of the healthcare services the insured may demand are both previously determined and provided by health professionals, who are skilled specialists with years of formal studies and learning by doing in their activities.⁷ Those features make the capabilities and knowledge of health specialists a human specific asset in the sense of Oliver Williamson (1996, 60). The consequence is that no relationship *a priori* can be argued between the expert’s decision and the extension of coverage the insured has, for the simple reason that the specialist decides according to her/his expertise, and not according to how much the insured wants to pay. It then becomes clear that there is no clear relationship between the extension of coverage (measured as the level of coinsurance) and the demand for healthcare, except perhaps for minor health problems and aesthetic corrections. Surely there is no such relationship for complex and expensive healthcare procedures, like complex or risky surgeries, cancer therapy, hemodialysis, etc. The second problem is clearly identified by Hodgson (2009, 104):

⁷ Hodgson (2009, 103) has already made the same point, and before him Hsiao (1995, 132) who wrote: “Unlike in the markets for groceries or clothing, asymmetry of information between buyers (patients) and sellers (physicians) vastly undermines consumer sovereignty”.

“Furthermore, even when affected by a similar injury or infection, the nature and severity of the outcome can vary from individual to individual. Health care needs are idiosyncratic, reflecting substantial physiological and neurological variations between individuals. Differences in health problems emanate from differences in past environment and genetic endowment. The peculiarities often vary significantly from person to person; each patient requires an individual diagnosis and remedy”.

Hodgson is simply pointing to the ordinary fact that there is no simple univocal relation between the health problems the patient has and the therapy she/he needs. The variance among patients and the fact that the health expert defines the kind of health services the patient need are sufficient to invalidate any presumption of a typical demand curve of the insured patient for health services. The inevitable conclusion is that the traditional supply-and-demand apparatus is definitely inadequate to analyze healthcare economics. One needs to turn to an alternative approach, and that approach must consider transaction costs in healthcare.⁸ This will be the next subject.

3. Transaction costs in the provision of healthcare services

It should be clarified first that Oliver Williamson’s (1985) conditions for the significance of transaction costs are all present in healthcare supply. Since Arrow (1963) it has been recognized that the provision of healthcare is clearly subject to bounded rationality, complexity and uncertainty, which together provide the occasion for opportunistic representations according to Williamson (1985). All those features plus the already discussed specificity of each patient’s therapy suggest that transaction costs are a significant share of total costs in healthcare provision. In sequence the sources of transaction costs in healthcare supply are specified, with particular emphasis on the patient-physician relationship.

3.1 Trust in Patient-Physician Relationship

The new research agenda initiated by Arrow (1963) produced a vast research on market imperfections produced by moral hazard and adverse selection in the insurance carrier-insured relation, and the resulting deviation from the competitive ideal.⁹ However, much less attention has been paid to some very important (although brief) insights from that seminal work. For example, discussing what he calls ‘third-party control’ over payments, which means the control insurance companies must have over the relation between the health services supplier (physicians, hospitals etc.) and the patient (who is insured), Arrow (1963, 962) observes that:

The moral hazard in physician’s control (...) shows itself in those insurance schemes where the physician has the greatest control, namely, major medical insurance. Here there has been a marked rise in expenditures over time. In prepayment plans, where the insurance and medical service are supplied by the same group, the incentive to keep medical costs to a minimum is strongest. In plans of the Blue Cross group, there

⁸ It will *not* be argued here that transaction costs paradigm is the only valid approach to understand all healthcare economic aspects. The hypothesis here is just that transaction costs are significant for the study of healthcare economic problems, and it invalidates the usual supply-and-demand approach.

⁹ See, for example, Hurley (2000).

has developed a conflict of interest between the insurance carrier and the medical-service supplier, in this case particularly the hospital.

Arrow is pointing here to conflicts in the relations between health services providers and insurers, something that has not attracted much attention since then. Those conflicts are exacerbated when the health insurance is of the prepayment kind, when the insured pays a fixed amount regularly (usually on a monthly basis) in advance of any healthcare expenses. In this case insured health risks are all borne by the insurer, who must face the variance in the insured's health expenses. If the insurer is risk averse – a natural consequence of the fact that he must comply with the financial requirements made by the insurance regulator – there is a strong reason for the insurance carrier to be quite interested in the costs of the healthcare supplier. When one admits informational problems between the insurer and the healthcare provider and so the possibility of opportunism from both sides, it becomes clear that the relationship between them is not an easy one, which was illustrated by Arrow in the quoted passage above with the reference to Blue Cross' problems with hospitals.

In the sequence Arrow observes that:

The need for a third-party control is reinforced by another aspect of the moral hazard. Insurance removes the incentive on part of individuals, patients, and physicians to shop around for better practices for hospitalization and surgical care. *The market forces tend to be replaced by direct institutional control* (Arrow 1963, 962 emphasis added).

Here Arrow – of course unintentionally – seems to be anticipating Oliver Williamson's governance structure analysis. When there is a third-party paying healthcare costs with severe informational problems it is not enough to talk about contract clauses to correct price distortions and to realign incentives: one must have institutional controls, it means one must have a governance structure – in Williamson's terms – to contain healthcare costs.

However, the problem actually is even worse than Arrow supposes, for there is more than a problem of moral hazard induced by the insurer's promise to pay the healthcare costs whatever they might be. Trust is deeply involved in the relation between insured and health provider. It is clearly unrealistic to imagine one quitting over her/his trusted surgeon when there is the need for a surgery because she/he found a cheaper professional, no matter how much one loves to 'shop around' other goods and services. In fact, the remark on the lack of motivation of patients to 'shop around' for cheaper healthcare is surprising when one considers the importance Arrow himself attaches to trust in patient-physician relations.

Moreover, discussing the expected behavior of the physician Arrow acknowledges in another passage the important role trust plays:

It is clear from everyday observation that the behavior expected of sellers of medical care is different from that of business men in general. These expectations are relevant because medical care belongs to the category of commodities for which the product and the activity of production are identical. In all such cases, the customer cannot test the product before consuming it, and there is an element of trust in the relation. (1963, 949)

Afterwards, discussing how the physician-patient relation affects the quality of medical care, Arrow (1963, 951) comments that: "A pure cash nexus would be inadequate; if nothing else, the patient expects that the same physician will normally treat him on successive

occasions”. Finally, in the postscript Arrow associates the patient-physician relation to that kind of “personal and especially family relationships”, which “though declining”, are still economic relevant “in the most advanced economies” (1963, 967). Not only trust based on personal relations is actually important to healthcare perceived quality, but also in contradiction to Arrow’s argument on the decadence of trust in patient-physician relationships, Hall (2003) presents evidences that trust continues to be strong in those relations.

The relevance of trust in patient-physician relationship is in the core of the analysis of transaction costs in healthcare supply in this paper, as that relationship is usually the starting point for all the health services, which are provided in sequence to the patient. Unfortunately, until the moment the trust in that particular relationship has not attracted attention from researchers that employ transaction costs theory. That lack of interest is however understandable when one considers the fact that Oliver Williamson, the most important researcher in transaction costs and governance structures has been rather skeptical about trust, which he has called an ‘elusive notion’ (Williamson 1996, 250), but one cannot understand the patient-physician relationship without incorporating the concept of trust into the transaction costs analysis. Not surprisingly, Williamson’s lack of interest in non-calculative trust has attracted some tough criticism from researchers studying the healthcare system, where trust has an important role in its functioning (Gilson 2003). Yet incorporating trust into the transaction costs theory is not an easy task, as trust and opportunism (which is one of fundamental conditions for significant transaction costs) seem to be on opposite sides in human behavior for opportunism presupposes self-interested calculation, and calculation is obviously absent when there is trust. Nevertheless some clarifications may help to somehow incorporate trust in the analysis, trust in the sense of “the emotional bonds and obligations generated through repeated interaction, empathy and identification with the other’s desires or intentions, or the desire to treat the other as I would wish to be treated myself” (Gilson 2003, 1456).

First, trust in the patient-physician relation does not exclude opportunism, for trust is asymmetric in that relation: the patient trusts the physician more than the physician trusts the patient, as the physician concentrates the knowledge about how to handle the patient’s health condition. Second, while trust plays an important role in the patient-physician relationship, trust does not seem to play any significant role in the physician-insurer relationship, which seems to be governed only by rational calculus on both sides, and that opens way to opportunism for both parties. Third, and perhaps more important to transaction costs analysis, trust in patient-physician relationship generates switching costs for the insurance carrier as the patient will usually resist to abandon his trusted physician, even when the physician recommends expensive therapies.¹⁰ In fact, *trust makes the patient-physician relationship a typical “small numbers relation”* (Williamson, 1985), which is the final and essential condition for transaction costs to be relevant. *Trust increases as the physician accumulates knowledge and particular capabilities suited to the patient’s health idiosyncrasies*, which is a specific asset.

But the small numbers nature of the relationship between patient and physician has consequences that surpass their immediate relation. As the physician usually is the gatekeeper

¹⁰ The point that trust and liking can be sources of transaction costs through switching costs has been already made by Nicholson, Compeau, and Sethi (2001, 12).

to the supply chain of healthcare services and products the patient must access to implement the therapy, the small numbers character of the patient-physician relation extends to the other parts of the healthcare supply chain the physician recommends to the patient (hospitals, laboratories, other specialists for more complex and specialized levels of care,¹¹ etc.).

Such effect of extending the small numbers character of the patient-physician relation to other parts of healthcare supply chain put the whole pack of goods and services of a therapy outside the scope of price and cost control, which is the objective of the health insurer. Actually, notwithstanding the voluminous literature concerning empirical tests on the existent physician-induced demand,¹² the issue is not if the physician will or will not induce demand to increase her/his revenues. The issue is that any independent professional included in the healthcare supply chain — i.e., independent from a health insurer — does not have the same objectives as the insurer's. While the insurer wants to contain costs, the health professional decides in conformity with her/his knowledge, her/his experience, and her/his personal evaluation: even when the health professional is not acting opportunistically to increase profits or revenues, she/he is unconcerned about cost containment, and that ensues consequences for the management of the health supply chain, as will be discussed in the next subsection.

3.2 The Management of the Health Supply Chain

Unfortunately Arrow did not develop the insight of the need for a third-party institutional control on the health provision costs, but managed care¹³ (originally developed in the United States in the seventies) introduced a form of institutional third-party control in healthcare services very similar to Arrow's anticipation, and managed care has assumed a central role in the working of healthcare systems not only in the U.S., but also in countries with a large share of private agents in healthcare like Brazil. So, that insight is of great importance to the comprehension of how the healthcare supply works, and why it is affected by significant transaction costs in its operation.

One of most common managed care techniques – either in public health systems like the National Health System of the United Kingdom or in private health insurance companies – is the *gatekeeper*. The function of the gatekeeper – usually a general practitioner – is to control patient access to secondary care, which is not only more specialized but also serve as the starting point to even more specialized and expensive levels of healthcare, like tertiary

¹¹ Usually healthcare is divided in four levels: first there is primary care (the first level of consultation), then secondary care provided by more specialized professionals like cardiologists, psychiatrists, urologists, etc. If necessary, the professional who provided primary care directs the patient to secondary care specialists, who by their turn manage the network of the more specialized healthcare services the patient should receive at that level. The same process of directing the patient and managing the network of healthcare at the proper level happens from secondary to tertiary healthcare (provided usually for inpatients in need of even more specialized services like cardiac surgery, neonatal complex treatments, cancer therapy, etc.), and from tertiary to quaternary care (highly specialized and restricted services, like experimental therapy). The important point here is that at each level of care there is always a professional or group of professionals managing the network of healthcare services at that level or guiding the patient to the professionals who will act as gatekeepers of the next level.

¹² Which to date seems rather inconclusive: see Rice (2006).

¹³ Managed care is a set of techniques designed to control health costs. It usually involves: negotiating price schemes like capitation payments and rebates that transfer some of the risks to healthcare providers, previous authorization for complex medical procedures, utilization review, copayments, gatekeepers etc. See Glied (2000).

healthcare (as surgeries) or even quaternary healthcare (experimental and uncommon surgeries).¹⁴

The first point about the role of the gatekeeper in the working of any health system is that not only the general practitioner formally appointed as gatekeeper, but also actually *any* organization providing healthcare (like a hospital, a clinic, a medical cooperative etc.) acts *as* a gatekeeper to some level of the healthcare supply chain, in the sense that these organizations control the access of the patient either to the first level of the system or to more complex healthcare supply chain levels, and so they regulate in some extent the nature and volume of the provision of healthcare services.¹⁵ Consequently, besides healthcare “demand” being ‘irregular and unpredictable’ as Arrow (1963, 948) describes it, the usual hypothesis of independence of demand from supply is not acceptable as a tool for healthcare economic analysis.

The second point about the gatekeeper is the fact that *the specialist or organization that acts as a gatekeeper also manages the healthcare supply chain necessary to provide the service*. In other words, the physician, clinic, hospital, etc. who gives the patient the access to some level of the healthcare system usually selects which health services will be provided, who will provide that services and when. The management of the healthcare supply chain by the gatekeeper may be indirect — through a network of associated providers — or direct — through a vertical integrated healthcare organization, as a hospital that has its own laboratories, medical equipment for long therapies, etc. In the case of an indirect management, there will be transaction costs in each node of the providers’ network. In the case of direct management, transaction costs will be present in the relation between the gatekeeper and the insurance carrier, if the patient is insured.

The third and more important point about the gatekeeper has direct connection with what has been said above about the way a health professional decides: as the gatekeeper in practice acts as the manager of the healthcare supply chain (at least until the next level of care, where the patient will face another specialist who will act as the gatekeeper of that level), and given the fact that the gatekeeper is a health professional or a group of health professionals, it follows that the management of the healthcare supply chain has no relation to cost containment (even if there is no opportunism by the healthcare provider). It has important consequences to the regulation of health insurance, as will be seen in the next subsection.

¹⁴ See footnote 10 above.

¹⁵ The fact that physicians (or healthcare organizations) act in practice as gatekeepers to the following level of healthcare, combined with the significance of trust in patient-physician relationships paradoxically may render the obligation of previously visiting a general practitioner as gatekeeper to access the public health system quite ineffective as an instrument of cost savings. In fact, discussing the adoption of the gatekeeper in OECD countries and the small cost savings that it generated, Reibling and Wendt (2012, p. 502) point out that:

Our literature review indicates that gatekeeping has the potential to increase efficiency and reduce costs, but savings are probably smaller than political expectations. One reason for low cost savings is that many people in free access countries have already voluntarily followed a gatekeeping patient pathway, as indicated by the high number of persons with a family doctor who also usually forms the first point of contact in free access systems (...). As a result, most people in free access systems see no problems in the development towards gatekeeping.

See also Dourgnon and Naiditch (2010).

3.3. Brazilian regulatory rules and transaction costs in private health provision

When facing the problem of a management of healthcare supply which is at best alien to cost containment, if not contrary to cost containment (because of opportunism), health insurers have generally adopted four possible reactions, individually or in some combination: (1) to increase the price or reduce the coverage of health insurance; (2) to adopt more rigorous managed care controls; (3) to exercise its monopoly power when negotiating with healthcare providers, and (4) to vertically integrate reducing the importance of independent healthcare providers along the healthcare supply chain. Let's consider very briefly (because of extension restrictions) each of those four reactions in relation to the Brazilian regulatory frame.

Price and coverage adjustments are very limited due to Brazilian regulations. Health insurance coverage is determined by the Brazilian regulatory agency, ANS (*Agência Nacional de Saúde Suplementar* – The National Agency for Private Health Insurance and Plans).¹⁶ The price of insurance is free for new contracts (though subject to ANS's supervision), but the following price increases are determined by ANS, which again limits the possibility of adjusting price to cope with the cost management difficulties in healthcare supply chain. To adopt more severe managed care controls – like denials of expensive health treatments – usually results in three kinds of problems for health insurers in Brazil: bad reputation, punishments from the regulatory agency and costly litigations (case law shows that Brazilian courts frequently adjudicate in favor of the insured). Besides those problems peculiar to the Brazilian case, there are also the additional administrative costs of adopting managed care controls. Those additional administrative costs of managed care controls help to explain the high level of administrative costs of health insurers, notwithstanding the mounting use of computation and information systems, one of the puzzles Baumol's 'cost disease' thesis cannot explain.

The option of exercising monopoly power against health providers has been largely used by health insurers, but in the last years it has produced growing discontent with two large physicians lockouts. Vertical integration to eliminate independent healthcare providers has been the only option remaining, which has been increasingly used especially by all kinds of non-financial firms providing some kind of health insurance coverage (like Brazilian health maintenance organizations – *medicina de grupo* – and medical cooperatives), with the only exception of financial health insurers, which are prevented by Brazilian law from operating outside the financial sector.

Conclusions

There is an intrinsic incompatibility between health insurer cost containment objectives and the way the healthcare supply chain is managed by its gatekeepers. This fact is obscured either by a purely technological approach of healthcare costs like Baumol's cost disease theory, or by the traditional supply and demand approach. An alternative approach based on transaction costs is necessary in order to understand movements like vertical integration waves as the one happening now in Brazil.

¹⁶ The word *suplementar* in Portuguese (supplementary, in English) refers to the role private health insurance was supposed to perform in Brazilian health system: it was originally supposed that private health insurance would only complement the public health system supply. In reality the Brazilian private system duplicates the public health system.

REFERENCES

- Arrow, Kenneth J. 1963 "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review* 53(5): 941-973.
- Baumol, William J. 2012. *The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't*. New Haven: Yale University Press.
- _____. 1996. "Children of Performing Arts, the Economic Dilemma: The Climbing Costs of Healthcare and Education." *Journal of Cultural Economics* 20(3): 183-206.
- _____. 1993. "Health Care, Education and the Cost Disease: A Looming Crisis for Public Choice." In *The Next Twenty-Five Years of Public Choice*, edited by Rowley, Charles K., Friedrich Schneider, and Robert D. Tollison, 17-28. Dordrecht: Springer Netherlands.
- Cowen, Tyler. 1996. "Why I Do Not Believe in the Cost-Disease." *Journal of Cultural Economics* 20(3): 207-214.
- Cutler, David M., and Dan P. Ly. 2011. "The (Paper) Work of Medicine: Understanding International Medical Costs." *The Journal of Economic Perspectives* 25(2): 3-25.
- Davis, Karen, Cathy Schoen, and Kristof Stremikis. 2010. *Mirror, Mirror On The Wall: How the Performance of the US Healthcare System Compares Internationally: 2010 Update*. New York: Commonwealth Fund.
- Dourgnon, Paul, and Michel Naiditch. 2010. The Preferred Doctor Scheme: A Political Reading of a French Experiment Of Gate-Keeping. *Health Policy* 94(2): 129-134.
- Ferris, J. Stephen, and Edwin G. West. 1996. "The Cost Disease and Government Growth: Qualifications to Baumol." *Public Choice* 89(1-2): 35-52.
- Gallouj, Faïz, and Maria Savona. 2009. "Innovation in Services: A Review of the Debate and a Research Agenda." *Journal of Evolutionary Economics* 19(2): 149-172.
- Gilson, Lucy. 2003. "Trust and the Development of Healthcare as a Social Institution." *Social Science & Medicine* 56(7): 1453-1468.
- Glied, Sherry. "Managed Care". 2000. In *Handbook of Health Economics*, edited by Anthony J. Culyer and Joseph P. Newhouse, v. 1, 707-753. Amsterdam: North-Holland
- Hall, Mark A. "Arrow on Trust". 2003. *Uncertain Times: Kenneth Arrow and the Changing Economics of Health Care*. Durham: Duke University Press.
- Hecht, Jason. 2001. "Classical Labour Displacing Technological Change: The Case of the US Insurance Industry." *Cambridge Journal of Economics* 25(4): 517-537.
- Heilbrun, James. 1997. "Baumol's Cost Disease." In *Baumol's Cost Disease: The Arts and Other Victims*, edited by Ruth Towse, 91-101. Aldershoot: Edward Elgar Publishing Ltd.
- Hodgson, Geoffrey M. 2009. "Towards an Alternative Economics of Health Care". *Health Economics, Policy and Law* 4(1): 99-114.
- Hsiao, William C. "Abnormal Economics in the Health Sector". 1995. *Health Policy* 32(1): 125-139.
- Hurley, Jeremiah. 2000. "An Overview of the Normative Economics in the Health Sector". In *Handbook of Health Economics*, edited by Anthony J. Culyer, and Joseph P. Newhouse, 55-118. Amsterdam: North-Holland.

-
- Kahn, James G., Richard Kronick, Mary Kreger, and David N. Gans. 2005. "The Cost of Health Insurance Administration in California: Estimates for Insurers, Physicians, and Hospitals." *Health Affairs* 24(6): 1629-1639.
- McKee, Martin, Dina Balabanova, Sanjay Basu, Walter Ricciardi, and David Stuckler. 2013. "Universal Health Coverage: A Quest for All Countries But Under Threat in Some." *Value in Health* 16(1): S39-S45.
- Nicholson, Carolyn Y., Larry D. Compeau, and Rajesh Sethi. 2001. "The Role of Interpersonal Liking in Building Trust in Long-Term Channel Relationships." *Journal of the Academy of Marketing Science* 29(1): 3-15.
- OECD Health Data 2013, accessed 15 August 2013, <http://www.oecd.org/els/health-systems/OECD-Health-Data-2013-Frequently-Requested-Data.xls>
- Preker, Alexander S., Birgit Hansl, Vijayasekar Kalavakonda, Somil Nagpal, and Nicole Tapay. 2012. *Private Voluntary Health Insurance: Consumer Protection and Prudential Regulation*. Washington, DC: World Bank Publications.
- Reibling, Nadine, and Claus Wendt. 2012. Gatekeeping and Provider Choice in OECD Healthcare Systems. *Current Sociology* 60(4): 489-505.
- Rice, Thomas. "The Physician as the Patient's Agent". 2006. In *The Elgar Companion to Health Economics*, edited by Andrew M. Jones, 261-268. Cheltenham: Edward Elgar Publishing Ltd.
- Triplett, Jack E., and Barry P. Bosworth. "'Baumol's Disease' Has Been Cured: IT and Multifactor Productivity in US Services Industries." 2006. In *The New Economy and Beyond: Past, Present, and Future*, edited by Dennis W. Jansen, 34-71. Cheltenham: Edward Elgar.
- Williamson, Oliver E. 1996. *The Mechanisms of Governance*. New York: Oxford University Press.
- _____. 1985. *The Economic Institutions of Capitalism*. New York: The Free Press.
- Woolhandler, Steffie, Terry Campbell, and David U. Himmelstein. 2003. "Costs of healthcare administration in the United States and Canada." *New England Journal of Medicine* 349(8): 768-775.
- Zweifel, Peter, Boris B. Krey, and Maurizio Tagli. 2007. "Supply of Private Voluntary Health Insurance in Low-Income Countries". In *Private Voluntary Health Insurance in Development: Friend or Foe?*, edited by Alexander S.Preker, Richard M. Scheffler, and Mark C. Basset, 55-113. Washington, DC: World Bank.