

*Stanford University
Walter H. Shorenstein Asia-Pacific Research Center
Asia Health Policy Program*

*Working paper series
on health and demographic change in the Asia-Pacific*

**Prescribing Institutions:
Explaining the Evolution of Physician Dispensing**

Karen N. Eggleston, Walter H. Shorenstein Asia-Pacific Research Center, Stanford University

Asia Health Policy Program working paper #24

October 18, 2011

<http://asiahealthpolicy.stanford.edu>

For information, contact: Karen N. Eggleston (翁笙和)
Walter H. Shorenstein Asia-Pacific Research Center
Freeman Spogli Institute for International Studies
Stanford University
616 Serra St., Encina Hall E311
Stanford, CA 94305-6055
(650) 723-9072; Fax (650) 723-6530
karene@stanford.edu

Prescribing Institutions: Explaining the Evolution of Physician Dispensing

Karen N. Eggleston

Walter H. Shorenstein Asia-Pacific Research Center, Stanford University

October 18, 2011

A condensed version of this paper is forthcoming in the *Journal of Institutional Economics*.

Abstract: Health systems provide a rich field for testing hypotheses of institutional economics. The incentive structure of current healthcare delivery systems have deep historical and cultural roots, yet must cope with rapid technological change as well as market and government failures. This paper applies the economic approach of comparative and historical institutional analysis (Aoki, 2001; Greif, 2006) to health care systems by conceptualizing physician control over dispensing revenues as a social institution. The theory developed -- emphasizing the interplay between cultural beliefs, interest groups, technological change, insurance expansion and government financing -- offers a plausible explanation of reforms since the 1960s separating prescribing from dispensing in societies such as Japan, South Korea, Taiwan, and China. Technological change and adoption of universal coverage trigger reforms by greatly increasing the social opportunity costs of physician over-prescribing and reshaping the political economy of forces impinging on the doctor-patient relationship.

1 Introduction

A central theme of institutional economics is understanding the evolution of economic institutions. Why are some economic institutions resilient and others fragile? Which institutions create better economic performance? How and why do institutions governing important dimensions of behavior change? Society prospers when organizations have incentive to overcome agency costs and other obstacles to efficiency (Posner 2010) and invest in productive outputs and innovations (North 2005). Yet organizations and the broader incentive structures within which they are embedded often exhibit strong path dependence, with innovations constrained by inherited norms and beliefs.

Health care provides a fascinating laboratory for studying such issues of institutional evolution because of the confluence of three characteristics: (a) cultural norms and beliefs strongly influence health and medical care decisions; yet (b) medicine has experienced rapid technological change that disrupts traditional norms and beliefs; and (c) market failures are rife in health care financing and delivery (Arrow 1963). Health and appropriate medical systems are not only inherently valuable – as a foundation of human capital and social welfare – but also increasingly important for economic performance of an entire economy.¹ The stark human consequences of market and government failures in health care raise the stakes for institutional design and adaptation. The health sector may be one of the few remaining sectors of market-based economies where a lively debate continues about the relative merits of the state and the market in delivering better performance. Moreover, many elements of health systems (including the one studied in this paper) have been adopted or imposed from abroad in some countries, allowing study of institutional transplantation. Finally, health policies provide a useful microcosm for studying resilience of institutions during system transformation during socialist and post-socialist transitions (Kornai and Eggleston 2001). The institution focused on in this paper, for example, survived from pre-modern China through socialist governance during the Mao era through to today’s market-based economy. Understanding why physician dispensing persisted across China’s economic system transformations of 1949 and 1979,

¹Health expenditures tend to grow as an economy develops (with income elasticity greater than 1) so that high GDP countries spend a higher share of GDP on health (Newhouse 1992), in part because the marginal utility of life extension, unlike that of consumption, does not decline as income rises (Hall and Jones 2007). Furthermore, the productivity of health spending can significantly impact macroeconomic performance (as illustrated by the 2009 debate about health care reform being crucial for US economic recovery). Health improvements lengthening life expectancy and improving quality of life constitute a large fraction of overall improvement in living standards over the past century (Deaton 2004; Murphy and Topel 2006).

but then became one of the central targets of China’s 2009 health reform plan, can shed light on how and why institutions change.

International differences in physician dispensing policies have been noted as a puzzle (e.g, “why physician dispensing has remained low in the United States while it become a major source of revenue in Japan is uncertain” Rodwin and Okamoto 2000, p.355). This paper applies the economic approach of comparative and historical institutional analysis (Aoki 2001, Greif 2006) to health policy by developing a theory of physician dispensing and reforms separating prescribing from dispensing, motivated by the experience of East Asia. All models abstract from some aspects of reality to focus on others. In this paper, I focus on prescribing and dispensing, while abstracting from the industrial organization of other stages in the pharmaceutical production and distribution process, such as research, testing, manufacturing or compounding, and wholesale distribution.

According to Greif (2006), an institution is a system of rules, beliefs, norms, and organizations that together generate a regularity of social behavior. I seek to explain physician dispensing as a social institution, or a regularity of physician-patient behavior. I define physician dispensing as physician control over revenues from drug dispensing. Under what conditions were (the rents from) diagnosis and evaluation, on the one hand, and dispensing of medication, on the other, integrated within a single provider rather than separated between a physician and an independent apothecary or pharmacy? With this way of framing the question, it is natural to seek conceptual answers in the literature on the boundaries of the firm. In particular, I modify the theory of asset ownership and relational contracting developed by Baker, Gibbons and Murphy (2002).

Relational contracts are “informal agreements sustained by the value of future relationships” (Baker, Gibbons and Murphy 2002, p. 39). This paper focuses on relational contracts between patients and physicians. Relational contracts provide incentives for noncontractible commitments (diagnosis and prescription efforts by doctors, future business and referrals of patients) if the parties sufficiently value their reputations and the future benefits associated with the physician-patient relationship.

The focus on reputations and relational contracting appears warranted by the unusual confidence many societies place in the self-regulatory ability of physicians, compared to other professions, occupations, and interest groups. Within the health sector and the supply chain for pharmaceuticals, physicians certainly occupy a unique position. In no market-based economy of which I am aware do laws prohibit for-profit corporate ownership of drug manufacturing firms. But several capitalist

societies – including Japan and Taiwan – disallow for-profit corporate ownership of healthcare delivery organizations, while nevertheless allowing individual physicians to own for-profit clinics and hospitals. Such a legal pattern suggests that societies have specific expectations about the medical profession and its ability to self-regulate to restrict fraudulent profiteering in the physician-patient relationship.

The simple theory developed in this paper illustrates institutional evolution and points to several determinants of the social institutions governing the physician-patient relationship. I argue that physician dispensing was a persistent institution because physician interests coalesced around this arrangement and patients came to associate receiving a drug with an effective physician visit. As a result, the greater the number of generations who lived in a society with physician dispensing, the higher the associated costs of physician-induced demand could be before separation became the efficient self-enforcing social institution. Physician dispensing may also provide efficient incentive for investment in drug quality assurance in societies where drug quality is a problem, mitigating the inducement cost. Separation requires payment for enforcement – including overcoming physician interest groups – as well as additional (perceived) travel costs for patients, trading off these payments with the loss of social surplus associated with inducement under physician dispensing. The central proposition is that technological change and adoption of universal coverage both increase losses from inducement and are therefore likely to trigger separation policies. Support for the theory comes from the comparative history of separation reforms in East Asia and in a few other parts of the world (but not pre-modern Europe, where the role of guilds appears to have been paramount).

The paper proceeds as follows. After providing some historical background, I develop a model of physician dispensing that can help to explain the path dependence of health system incentive structures. Subsequent sections discuss extensions of the basic model to encompass regulation and heterogeneity. A section presenting evidence on the history of separation reforms in East Asia compared to Europe, the UK and US, weighing alternative explanation for the patterns of separation reforms, shows that while the theory is not appropriate for explaining separation in medieval Europe, the model's theoretical propositions are largely confirmed by the East Asian experience of separation reforms since the 1960s. The conclusion returns to the question of how the analysis provides an example of institutional evolution.

2 Background

2.1 The legacy of history and culture

Analysis of physician dispensing is related to several strands of literature. In part it is an application to health care of the economic theory of expert-client interactions when information is asymmetric (Darby and Karni 1973; Wolinsky 1993; Taylor 1995; Hubbard 2002; Afendulis and Kessler 2007). In part it represents an inquiry into what Abbott (1988) calls “the ecology of professions” in comparative sociology, exploring the professional boundaries between physicians and pharmacists, and the counterpart exploration in economics of boundaries between professions (Savage 1994) and occupations (Kleiner and Park 2010). Most centrally, however, it is an inquiry applying to health systems the method of comparative and historical institutional analysis (Aoki 2001, Greif 2006).

The cultures of healing and professions of medicine in the east and the west differed in many important ways, and each is itself a heterogeneous collection of many schools of thought. Here I focus on the role of medications in the patient-physician interaction. In the west, prescribing and dispensing of medications has often been ancillary to the primary expected outcomes of a physician visit, diagnosis and explanation.²

In herbal medical traditions as dominated in East Asia, however, the prescription and preparation of medications was central to the entire enterprise, and the dispensing of medicine came to be seen as the central outcome of the physician-patient interaction. In traditional Chinese medicine (TCM) and counterpart herbal traditions developed in Japan and Korea, it was considered improper for doctors to charge patients for the humane service of diagnosis, so that doctors primarily made a living by selling medications (Campbell and Ikegami 1998).

² “Once people began to regard science as a superior and legitimately complex way of explaining and controlling reality, they wanted physicians’ interpretations of experience regardless of whether the doctor had remedies to offer” (Starr 1982, p.19). As Roy Porter quotes in his opus *The Greatest Benefit to Mankind: A Medical History of Humanity from Antiquity to the Present*, distinguished American physician Lewis Thomas, who was a medical student in the 1930s, recalls

we were provided with a thin, pocket-sized book called *Useful Drugs*, one hundred pages or so, and we carried this around in our white coats when we entered the teaching wards and clinics in the third year, but I cannot recall any of our instructors ever referring to this volume. . . . *Our task for the future was to be diagnoses and explanation. Explanation was the real business of medicine.* What the ill patient and his family wanted most was to know the name of the illness, and then, if possible, what had caused it, and finally, most important of all, how it was likely to turn out (Porter 1997, pp.681-2, italics added).

2.2 Defining physician dispensing

Consider two primary institutional arrangements governing clinical reality in a given society at a given time: (1) financial integration of diagnosis and dispensing functions (‘physician-dispensing’); and (2) strict enforcement of patients’ property rights to a prescription, so that the patient may have a prescription filled at any pharmacy and the provider does not receive dispensing revenues (‘separation of prescribing and dispensing’).

Since separation is defined to occur only when the provider does not receive dispensing revenues, physician dispensing includes all organizational and property rights arrangements in which physicians have control over revenues from drug dispensing. Thus, physician dispensing encompasses not only a physician preparing drugs for patients in a back counter of the office, but also any arrangements in which physicians and outpatient clinics “buy” drugs on the open market at one price and “sell” them to patients and their insurers at a higher price (common in Japan, South Korea before 2000, and also true of U.S. oncologists; see Jacobson, Earle, Price, and Newhouse 2010). Physician dispensing also includes a physician hiring a pharmacist for on-site dispensing or owning a free-standing pharmacy (a common arrangement in Taiwan); or hospital-based physicians earning substantial income from hospital outpatient pharmacy revenues, often linked to individual prescribing (a typical situation in China).³

With these definitions, the question of physician dispensing can be distinguished from pure economies of scale and scope. Although for modeling convenience I focus on interactions of a representative physician and patient, classification into physician dispensing versus separation is not based on the scale of the organizations producing the diagnoses and receiving the dispensing revenues. The provider can span the gamut from a solo practice to a large multi-specialty clinic; a pharmacy can achieve considerable scale and still be financially linked or owned by physicians. The question examined here is whether a society retains physician control of dispensing revenues as the technology of pharmaceutical production, distribution, prescribing, and dispensing changes over time.

³On the various forms of physician dispensing in Asia, see discussion in the country-specific chapters of Eggleston 2009.

2.3 Physician dispensing as a social institution

Physician dispensing can arise in a society in which no legal or contracting institutions enforce restrictions on the physician-patient interaction or the property rights to dispensing revenues from a given medication. Doctors can freely choose whether to specialize in diagnosis or to integrate diagnosis and dispensing. In its initial form, physician dispensing prevailed in societies in which patients could choose to buy drugs of any form in any place where they were marketed. There often was no distinction between prescription and over-the-counter drugs, and few if any licensing and accreditation standards for providers. The developer of an effective new medical formulation had no enforceable property rights. Others could expropriate the innovation unless the originator kept it secret by, for example, dispensing the medication directly to the patient. (Indeed, a special term in Chinese, *mifang*, refers to ‘secret formulations’ in TCM.)⁴

In contrast, some societies have adopted restrictive laws that enforce a division of labor between physicians and pharmacists, including legal restrictions over physician or hospital ownership of dispensing revenues. Such a distinction may be reinforced by the organization of separate guilds or professional associations for physicians and pharmacists (Sonnedecker 1986; Savage 1994; Cochran 2006). Other aspects of separation that usually follow from proscribing physician income from dispensing include enforcement of property rights over drug discovery and a legal distinction between prescription-only and over-the-counter medications. Kickbacks from pharmaceutical firms to doctors are made illegal (although enforcement may be difficult). Only when a society has invested in such laws and enforcement institutions can genuine separation prevail. In South Korea, for example, separation reforms in 2000 forced hospitals to shut down and divest from outpatient pharmacies. Although intermediate institutional arrangements may exist, such as restrictions on ownership that are not enforced or even flagrantly violated, I focus here on physician dispensing and separation to illustrate the two dominant extremes.

I denote all contracting costs associated with separation as Ω . These contracting costs may be closely linked to the political economy of professional organizations in a society. In Europe, for example, contracting costs of separation in the modern era were relatively low because enforcement did not require overcoming the political and economic interests of strong physician-dispensing inter-

⁴Historical accounts provide some evidence that inability to enforce property rights over secret formulations played a role in shaping the institutional features of healthcare delivery. For example, both doctors and pharmacies hired their own relatives to help prepare herbal prescriptions so that the formulations could be kept secret (see Cochran 2006 and Eggleston and Yang 2009).

est groups. By contrast, contracting costs of separation are high in societies in which the medical profession organized to protect dispensing rents. In such a case, development of interest groups or lobbying organizations representing the interests of dispensing physicians reinforces physician dispensing as an institution. I therefore index contracting costs with time, Ω_t , to denote that political interests of providers will coalesce into more powerful interest groups as a function of the length of time a society has experienced physician-dispensing.

An arrangement prevails as a social institution, for the purpose of this analysis, when both parties – the upstream provider and the downstream purchaser – find the value of future interactions under the current institutional arrangement outweighs the temptation to renege on the relationship (and resort to the spot market or, eventually, change the social institution).

The model seeks to explain when parties will find physician dispensing to be the preferable self-enforcing structure governing the physician-patient relationship. To move from physician dispensing to separation requires paying $\Omega_t > 0$: overcoming the political, contracting, and enforcement costs of removing dispensing revenues from physicians.

Separation also requires that patients separately visit both the physician and an independent pharmacy, making two (or more) trips instead of one for each treatment episode.⁵ Let patients' perception of the differential travel cost associated with separation, at time t , be T_t . (If travel costs under physician dispensing are T^L and under separation are T^H , then $T = T^H - T^L$). Consumers' perception of the costs associated with the need to visit doctors and pharmacies on separate trips may differ according to how long they have experienced integration. The perceived cost is likely to be largest, for example, if consumers have lived with integration for many generations and therefore consider it unnatural and wasteful to make two visits rather than obtain the medication as part of the physician visit.

Claim 1 $\Omega_t + T_t$ increases with the length of time under physician dispensing.

Inertia in consumer expectations and the political economy of supply, subsumed in these two factors ($\Omega_t + T_t$), suggest that societies will continue physician dispensing until change of some

⁵Darby and Karni (1973) also emphasize the transaction costs of avoiding integrated 'diagnosis and repair' across many economic fields, using the specific example of dispensing: "If there were no additional costs involved in buying diagnosis and the actual repair service from different sources, then the consumer would generally do so, thereby avoiding the fraud at no additional cost. [Footnote] An example is the separation of prescription and preparation of drugs. The costs of separation of diagnosis and repair, in general, involve not only possibly increased monetary costs, but also may involve additional time or other transactions costs, which must be compared with any expected reduction of fraud" (p.69).

other factor, alternatives for which are discussed next.

2.4 Why separate prescribing from dispensing? Alternative hypotheses

Several economies of East Asia – including Japan, South Korea, China and Taiwan – have recently implemented policies to separate prescribing from dispensing, often closely upon the heels of increased public financing and health insurance expansion. Other Asian economies such as Singapore have not separated prescribing from dispensing. This section explores several alternative hypotheses about what can explain such a pattern of separation reforms. (I return to a more detailed historical account and discussion of these hypotheses after developing the model.)

One hypothesis might be that physician dispensing is determined by culture. Both theoretical and empirical evidence shows how cultural traits are transmitted intergenerationally through multiple mechanisms (Cavalli Sforza and Feldman 1981; Bisin and Verdier 2010) and how culture can shape incentives for specific forms of innovation (therefore playing a role in long-run economic growth; Gorodnichenko and Roland 2010). Perhaps culture determines the physician-patient relationship, and separation only materializes when there has been a change in culture. Yet separation reforms have not coincided with large changes in the use of herbal medications or other specific clinical manifestations of culture in East Asia, and Singapore shares a similar cultural background yet has not separated prescribing from dispensing. Thus it would appear that to explain the pattern of institutional change that we observe, culture must interact with other factors.

Another explanation is rent-seeking among closely affiliated professions, such as pharmacists and physicians. The hypothesis that separation is determined by a power struggle among guilds would appear to have considerable appeal for explaining separation in pre-modern Europe (as discussed further in the final section of this paper). However, it has less success in explaining the fact that nowhere in East Asia have pharmacist associations or pharmaceutical firms played a decisive role in separation reforms, whereas physician groups remain quite powerful. Moreover, the rent-seeking hypothesis does not present a compelling explanation for the timing of separation reforms, such as why separation followed closely in the footsteps of expanded insurance in Korea, Taiwan, and China. Of course, rent seeking logically does shape the contested borders of professions (and as we argue below, rent seeking is closely tied to physician-induced demand, since inducement creates and expands rents); but sans interaction with other forces, the rent-seeking theory seems incomplete.

Technological change is another obvious candidate for explaining institutional change in medicine. In the modern era, specialization is increasingly common across the supply chain of a range of production activities; medications are no exception. Once pharmaceutical firms became the primary producers of research and development as well as manufacturers of medications, and the processes of verifying the chemical content and purity of medications became too complex for a community doctor or pharmacist, then the integration of prescribing and dispensing increasingly became a relic of a by-gone era. In exploring the history of the pharmacy profession in the U.S., Deborah Savage (1994) provides a succinct statement of the technological-change-in-production hypothesis: “There was a fundamental change in the technology of drug production and distribution, especially as it affected pharmacists and physicians as producers and distributors. These well-recognized and fundamental changes were more subtle than ‘simple’ cases of rapid technological change because of their cumulative effects on professional institutions, but no less revolutionary” (p.154).

However, the technological change hypothesis in its pure form also has several weaknesses. Technology may change, but what changes the laws about who owns and profits from technology? Technological change in pharmaceuticals alone cannot explain why a country like Singapore – as affected by modern biomedicine and forces of globalization as any country – has never separated prescribing from dispensing. It hardly seems satisfying to posit that the separation reforms since the 1960s in Asia are simply a lagged policy response to globalization, implying that Europe was precocious and East Asia backward. Furthermore, since separation does not merely rearrange organizational boundaries but also re-defines ownership of dispensing revenues, technological change by itself cannot fully explain separation reforms. For example in China, medication inventories and dispensing functions may have moved from individual physician offices to the hospital pharmacy, but hospitals were legally allowed to charge a dispensing mark-up to cross-subsidize other medical services, and the average government-owned hospital in China by the mid-2000s earned close to half its revenues from dispensing medications; physicians, as employees of hospitals, often received bonuses linked to hospital profitability (Sun et al. 2008). Thus the integration of diagnosis and dispensing persists as a legally legitimized form of physician compensation despite technological change in the scale of medication manufacturing, prescribing, and dispensing.

In this paper, I explore the logical implications of an alternative, yet in some ways complementary, hypothesis: that separation reforms are directly linked to physician command over expanding

medical care resources and the incentive for wasteful over-use associated with physician dispensing. Continuing to allow physician dispensing avoids the contracting and interest group costs (Ω_t) and the transaction costs (T_t) of enforcing separation, but it also gives incentive for inefficient physician behavior, such as over-prescribing and dispense medications according to profitability. I denote the potential social welfare loss from ‘physician-induced demand’ using notation P .

The definition of physician-induced demand given in the *Handbook of Health Economics* is as follows: “Physician-induced demand exists when the physician influences a patient’s demand for care against the physician’s interpretation of the best interest of the patient” (McGuire 2000, p.504). Similarly, Eisenberg defines inducement as “prescription of services that a well-informed consumer would not want to use” (quoted in McGuire 2000, p.505). That physician-induced demand for pharmaceuticals may significantly impact drug use and spending patterns is clear from several recent studies (discussed in more depth in the final section). For example, Iizuka (2007) finds that physician dispensing in Japan leads to 15 percent higher expenditures for hypertension drugs than would be the case if the physician mark-up were eliminated. An audit study by Currie, Lin and Zhang (2010) finds that when facing patients who not only report symptoms that do not warrant antibiotics but also state reluctance to take antibiotics, Chinese dispensing physicians nevertheless prescribe antibiotics to more than a third of such patients. Chen, Gertler, and Yang (2011) find that Taiwanese physicians, when no longer paid a profit margin for dispensing, reduce prescriptions (measured by expenditures) by almost 30 percent.

Arguing that physician-induced demand may play a large role in decisions to prohibit physician dispensing does not deny the power and influence of technological change in medicine. To the contrary, physician-induced demand is elevated to a large social issue primarily thanks to advances in technology which add a wider array of medical resources to the physician’s armamentarium. Contemporary physicians decide not only if a patient should take a medication for a given acute episode; they can put patients on multiple pills a day for the rest of their lives, and physician recommendations for diagnostic procedures and therapies can encompass a vast array of additional expensive resources. In the classic economics treatise on modeling physician behavior, Pauly (1980) notes that physicians only absorb about 18 percent of total health spending – the 2010 figure for the US is still just 20.9 percent⁶ – yet “the relatively small amount that goes to physicians provides the

⁶ “Physician and clinical services” represented \$535.8 billion out of \$2,569.6 billion total health expenditures in the US in 2010; prescription drugs accounted for another \$260.1 billion, or 10.1% of total health expenditures (Truffer et al. 2010, Exhibit 1).

financial incentive which determines the bulk of resource use, output quantities and characteristics, and total costs in the health sector” (p.1). Physician authority⁷ over use of such resources leads to larger rents associated with integration and larger social costs from prescribing according to profitability rather than clinical indication.

Separation reforms, under this hypothesis, flow directly from a change in how health systems control physician-induced demand. Prior to the modern era, physician-induced demand arguably was a minor factor in overall health spending, given the limitations of medical technology. Even when technology expands and the expenditure implications of wasteful over-use increases, patient ability to pay constrains health spending in most societies at low levels of per capita income (and, in a different format, in contemporary Singapore). However, when the constraint on induced demand from patient demand-side cost sharing is removed – through expansion of insurance and increased public financing for health care – the social costs associated with wasteful overuse increase and strain government budgets; it is these forces, in the model, that spur separation reforms.

As we will see, encouraging an unproductive action such as excessive prescribing reduces total surplus but may also reduce the reneging temptation in the physician-patient relationship enough to make it worthwhile (for a similar argument, see Baker, Gibbons and Murphy 2002, p.67). Moreover, committing to reward physician prescribing, even with prices reflecting market power, can serve as a relational incentive for quality effort, such as assurance of pharmaceutical quality. Thus, it is not clear a priori that eliminating incentive for physician-induced demand is the socially efficient institutional structure under all parameter values or historical contexts.

Baker, Gibbons and Murphy note that “the key question in our analysis is whether choosing appropriate asset ownership (integration or nonintegration) can make a given promise self-enforcing” (p.41). Here, the key question is whether physician or patient ownership of the prescription “asset” (rights to where to fill the prescription and the associated revenue) can make two promises self-enforcing: (a) the purchaser promise not to expropriate the *mifang* of a physician-inventor or ‘defect’ from the physician-patient relationship to self-diagnose (perhaps with pharmacists ‘quasi-prescribing’ thereafter); and (b) the physician promise of not (excessively) inducing demand. The

⁷Savage (1994) defines professional authority as “command capabilities not available to non-professional economic agents,” derived from expert knowledge but stopping short of ability to force clients into actions; “examples include the attorney’s command over the legal system’s resources and a physician’s command over the health care system’s resources” (p.140). Starr (1984) recounts how the medical profession in the US rose to a position of cultural authority as well as social authority, entailing not only the command of medical resources for therapy but also legitimate claim to defining what constitutes illness, insanity, disability, and death.

“temptation to renege” is important because while physician dispensing may be an institutional equilibrium in the overall society, the meeting of any one patient with any one physician may be infrequent. The physician’s ability to induce demand will vary across encounters. As the model developed in the next section shows, when many physician-patient interactions lead to temptation to renege on the relational contract upholding physician dispensing, then that institution governing the physician-patient relationship begins to break down and is no longer self-enforcing. Separation reforms ensue.

3 A simple theory of physician dispensing and separation

The upstream party, the physician, produces services which, combined with the filled prescription, treat the downstream patient’s medical condition. If the physician owns the prescription asset, then physician dispensing prevails; Baker, Gibbons and Murphy call this ‘relational outsourcing.’ If the downstream patient owns the prescription asset, then separation prevails; Baker, Gibbons and Murphy call this ‘relational employment.’ I modify the model to fit institutional evolution of physician dispensing and separation, including induced demand, demand-side liquidity constraints, insurance coverage, incentives for drug discovery, and regulation.

Each period, the physician chooses two actions, quality-related effort and effort to induce demand, with an associated effort cost. No one except the physician herself can observe the actions. In this ‘unproductive multitasking’ setting, the value of the physician-patient relationship would be highest if the physician invested only in quality-improvement effort and never induced demand. The provider’s actual effort choices depend on both who owns the prescription “asset” and whether the patient-provider interaction is either (1) relational – with provider commitments to high effort and limited physician-induced demand rewarded by loyal patient consumers willing to pay a premium for quality – or (2) a “spot market” exchange without hope or commitment to future interactions. The focus is on comparing relational physician dispensing and relational separation. As Baker, Gibbons and Murphy highlight, asset ownership shapes reneging temptations in relational contracts. Whether or not the patient has property rights over where to fill a prescription will affect the provider’s temptation to overprescribe and thus whether physician dispensing will be self-enforcing.

3.1 Relational governance of physician dispensing

First consider physician dispensing: the patient ‘outsources’ diagnosis and prescription services so that revenues from both services flow to a single provider. In an ongoing relationship, the patient agrees to pay the physician in three components: a fixed payment for the contractible minimum quality; a bonus for high quality; and another payment that depends on physician effort to induce demand. Clearly, the patient will not wish to agree to pay for physician-induced demand, but may be unable to distinguish required services for quality outcomes from excessive and unnecessary services, and may be willing to pay the physician a premium or monopoly price ($\beta_H > 0$) to assure quality. The bonus β_j essentially represents the premium that patients are willing to pay when receiving medication as part of the physician-patient encounter (i.e., under physician dispensing).

The patient may come to realize that the physician is charging a monopoly price, or prescribing excessive drugs to increase net revenue, and will be tempted to renege unless the benefits of a continued relationship outweigh the costs of paying for the excessive prescriptions. The physician may be tempted to induce demand as far as possible, but has to weigh the short-term benefit against the loss of reputation and future patronage if the patient goes elsewhere. These temptations to renege on the relational contract depend on what happens when the relationship breaks down. Assume that a one-time physician-patient interaction under physician dispensing (‘spot outsourcing’) involves haggling over the price, with the final spot market price determined by Nash bargaining and trigger strategy equilibria: the parties are assumed to honor the relational contract as long as the other party does, and to revert to spot market transactions forever after if the other party reneges (as detailed in the appendix).

For a given discount rate, the physician will honor the relational contract when the current payment plus the present value of the future relationship exceeds the spot market payment in this period and forever after. The patient will honor the relational contract when the current net benefit and the present value of the future relationship exceed the alternative, the net benefit from paying the spot market price and forever after having spot market physician-patient interactions. For physician dispensing to be an equilibrium, the maximum ability to induce demand must not be too large, so that the patient does not renege and the patient out-of-pocket spending does not exceed the patient’s ability to pay (see appendix).

When patients lack insurance coverage, ability to pay is likely to be the binding constraint on physician-induced demand. An increase in coverage generosity, such as adopting social health

insurance, enables patients to buy more health care, relaxing the ability to pay constraint on both quality (classic moral hazard) *and* physician-induced demand (supply-side moral hazard).⁸ Societies that provide risk pooling will confront higher social costs of retaining the physician dispensing institution.

In addition, it seems plausible that doctors' ability to induce demand depends on available technology. Therefore we assume that as the capabilities of medicine grow, the social opportunity costs from physician-induced demand also grow. As a result, technological change will play a role in determining when physician dispensing and separation are self-enforcing institutions.

3.2 Separation of prescribing and dispensing

Under separation, the patient is relatively "empowered" vis-a-vis the provider.⁹ The physician has no incentive to induce demand because all dispensing revenues flow to independent pharmacies. The physician also has limited ability to charge a monopoly price, since the patient and perhaps pharmacist can see the prescription and how it is matched to symptoms and "defect" from the physician-patient relationship to engage in self-diagnosis or pharmacist 'quasi-prescribing' in the future. The physician may also have little incentive to invest in quality-improvement effort. In addition, compared to physician dispensing, the patient under separation must incur higher travel costs to visit both a physician and a separate pharmacy. When will it be socially efficient to require patients to pay these extra travel costs and to invest in the contracting institutions to uphold such a separation equilibrium? In other words, when will separation facilitate physician-patient relationships that generate social surplus sufficient to justify abolishing physician dispensing?

Under separation, the patient 'employs' the physician as his or her agent to diagnose the medical condition and prescribe a remedy. Under 'relational employment,' the upstream provider's temptation is not inducing demand but rather stinting on effort. Indeed, when separation is enforced, the fall-back spot market transaction is that the physician provides only minimum contractible quality. By design, separation removes the opportunity for excessive prescriptions motivated by prescriber-dispenser profits (P^H). But the purchasers (patients and their insurers) must pay for the

⁸Some evidence in support of this assumption comes from Iizuka (2007), who finds that Japanese dispensing physicians will forego one dollar in mark-up in exchange for a 28 cent reduction in patient out-of-pocket cost. He hypothesizes that patient ability to discipline the provider, by referrals or control of repeat visits (and thus future demand and revenue for the doctor) constrains demand inducement. Increasing generosity of coverage weakens this demand-side constraint. Also see Lundin 2000.

⁹See Hart (1995) on how asset ownership under incomplete contracting conveys power.

contracting costs of legislating and enforcing separation (each period), Ω_t , as well as bear the additional travel costs, T_t . These costs will be greater, the longer the time period over which provider interest groups have coalesced around the rent stream associated with physician dispensing and the more generations of patients have become accustomed to receiving a medication as the ‘payoff’ from visiting a physician.

The patient honors the relational contract if the maximum promised bonus for quality is less than the incremental benefit of relational contracting compared to spot market contracting. The physician chooses whether to honor the relational contract or refuse the bonus as too low for the requisite effort, instead reverting to one-time interactions with patients for spot market prices and minimum contractible effort. The physician will honor the relational contract if the incremental value of continuing the relationship exceeds the minimum expected “pay for performance” bonus. Ability to induce demand plays no role in the reneging constraint for separation (see Appendix).

Separation must also fulfill the ability to pay constraint, but separation by definition removes physician-induced demand from the patient out-of-pocket burden and thus may relax the constraint. In effect separation substitutes a payment to enforcement institutions ($\Omega_t + T_t$) for the extra payment to the provider (P), allowing patient payments at time of service use to be better targeted on increasing the value of treatment.

4 Physician-dispensing and separation reforms

What does this theory predict are the determinants of separation reforms (i.e., switching from physician dispensing to separation)? The theory predicts that physician dispensing dominates separation if and only if

$$\frac{(P^H)^2}{2} < \Omega_t + T_t. \quad (1)$$

Proposition 2 *The higher the contracting, political economy, and perceived travel costs of separation, the wider the range of parameters over which physician dispensing is self-enforcing.*

Since Ω_t and T_t increase in the number of periods a society has been under physician dispensing, the institution is self-reinforcing (Greif 2006). Each individual does not foresee this reinforcing of physician dispensing, however. If players are unaware of the reinforcement mechanism, then at

time τ they imagine that the costs of switching to separation will remain the same into the future, although in fact it will increase if physician dispensing continues. Thus $\Omega_t + T_t$ is not a fixed parameter, nor an endogenous variable; it is a ‘quasi-parameter’ (Greif 2006).

Proposition 3 *The longer a society remains under physician dispensing, the higher the associated costs of physician-induced demand (P^H) can be before implementing separation becomes the efficient self-enforcing social institution.*

The potential for physician-induced demand in any given period – P^H – is also itself a function of insurance and technology, and is thus another quasi-parameter.

The ‘efficient organizational form’ – physician dispensing (‘relational outsourcing’) or separation (‘relational employment’) – varies with the discount rate r and the magnitude of P^H . For high discount rates, neither patient nor physician cares much about the future, and therefore reputation-based relational contracts are not self-enforcing; spot market interactions prevail (see Figure II of Baker, Gibbons and Murphy 2002). At very low discount rates, both institutional arrangements can achieve socially efficient outcomes. Probably most interesting and policy-relevant is the intermediate range of low to moderate discount rates, where the purchaser and provider agree to cooperate in an ongoing relationship, with physician dispensing dominating for low P^H and separation dominating for high P^H .

Proposition 4 *Separation policies – switching from physician dispensing to separation – can be an efficient response to increases in P^H , the potential loss in social surplus from maximum physician-induced demand.¹⁰*

Figure 1 illustrates how the quasi-parameters T_t , Ω_t , and P^H impact the range over which physician dispensing and separation are self-enforcing institutions, assuming an intermediate range of discount rates that can support relational contracting. The horizontal axis graphs the potential loss in social surplus from maximum physician-induced demand, P^H . The vertical axis shows the costs associated with enforcing separation, $\Omega_t + T_t$. The upper left-hand side, with high $\Omega_t + T_t$ and low P^H , represents a range of over which physician dispensing is the efficient equilibrium institution. The lower right-hand side, with low $\Omega_t + T_t$ and high P^H , illustrates the range of over which separation is the efficient self-enforcing institution. The dividing curve illustrates when

¹⁰ Compare BGM’s “first result” – that switching from relational outsourcing to relational employment is an efficient response to widely varying supply prices (BGM, p.64).

the social surplus associated with physician-dispensing exceeds that of separation, as given by the $\frac{(P^H)^2}{2} < \Omega_t + T_t$.

What determines the magnitude of P^H ? As discussed above, technological change continuously pushes up the capabilities of medicine, and with it, the potential welfare losses associated with physician-induced demand. Thus, all else equal, separation is more likely to prevail as physician recommendations influence utilization of a wider range of expensive health care technologies.

A second critical determinant of P^H is the generosity of health insurance or healthcare coverage, such as through National Health Insurance (NHI) or a National Health Service as in the UK. Adoption of universal coverage is plausibly associated with a significant one-time increase in P^H : reducing the patient burden of out-of-pocket spending will relax the ability-to-pay constraint on physician-induced demand.

A political mechanism may also link technological change to reductions in patient out-of-pocket burden when technological advances imply that access to healthcare becomes more unequal and has more significant impacts on life chances. Consider a certain level of social solidarity, manifest as a political threshold of acceptance for disparities in access to healthcare. Technological change will typically increase health spending and strain the poor's ability to pay, increasing disparities in medical care. Once those disparities breach the political threshold of acceptability, they will trigger political reform to adopt or extend risk pooling for medical expenses (captured in the model by a decrease in the parameter θ).

Proposition 5 *Technological change directly increases P^H (because technology tends to increase spending) and indirectly increases P^H through an induced increase in coverage. Therefore separation policies are more likely when there has been (i) rapid technological change in healthcare, and/or (ii) a large increase in generosity of insurance coverage, which may itself be a political result of technological change increasing disparities of access beyond the acceptable social threshold.*

To the extent that per capita income is correlated with the level of healthcare technology perceived as available in a society, the international pattern of higher public spending at higher per capita income levels is consistent with this hypothesis. Further empirical evidence consistent with this hypothesis comes from adoption of universal coverage at ever lower levels of per capita income during times of rapid technological change in health care, such as the last several decades. Of course there also is evidence that insurance coverage itself can spur technological change (e.g.

Finkelstein 2007).

Societies could be arrayed in Figure 1 in different ways, with many European countries presumably toward the southwest (early separation) and the US and UK towards the middle (later separation). The theory predicts that for societies with high contracting costs and low potential for physician-induced demand historically, physician dispensing was a self-enforcing and self-reinforcing institution. Over a long period of time, these societies moved from southwest to northeast on Figure 1: physician dispensing was self-reinforcing as patients and doctors came to believe that dispensing was the natural outcome of any effective physician-patient encounter and the political authority of providers grew to be able to block reforms. Several economies of East Asia appears to fit this description, as discussed in more detail below. On net, arguably $\Omega_t + T_t$ increased in these societies (until recently). Technological change raised the social opportunity costs of wasteful overuse (P^H) ever higher. Together these forces pushed East Asian societies upward and left in Figure 1, perhaps close to but only recently crossing over the boundary between physician dispensing and separation as self-enforcing institutions.

4.1 Regulation

The contracting costs Ω embody numerous costs of regulating and enforcing strict separation of prescribing from dispensing. Yet a society may also use regulation within the physician dispensing institutional framework to mitigate physician-induced demand. One advantage of such an approach is avoiding the costs of switching to separation, especially the patient perceived travel cost T_t . Moreover, physician dispensing may help to assure quality by linking physician reputation to drug quality. Counterfeit medications constitute a world-wide problem, especially in developing countries. Hospitals and doctors have incentive to invest in costly quality assurance when they earn substantial revenues from drug dispensing.

Proposition 6 *When physician dispensing offers a quality assurance advantage over independent pharmacies (by, for example, linking physician and hospital reputations to the quality of the drugs they dispense), physician dispensing dominates separation up to a higher threshold of physician-induced demand.*

Consider a regulatory structure that with probability π can detect if the physician induced demand and impose a fine of $X > 0$. Following Glaeser and Shleifer (2003), X denotes the maximum

fine that can be enforced. Mechanisms might include claims review to detect over-prescribing and investigation of doctors' financial relationships with pharmaceutical firms and pharmacies to penalize kickbacks, with associated enforcement costs of $\Omega(\pi X)$. Regulation of prescribing practices can reduce doctors' expected benefits from induced demand to $\hat{\beta}_H = \beta_H - \pi X$, lowering inducement and making physician dispensing sustainable over a broader range of parameters.

A second important avenue of regulation involves reimbursement or relational payments, especially when the downstream purchaser involves an organization (such as public insurance agency or a managed care organization). Moving toward "pay for performance" (with relational bonuses linked to evidence of patient benefit) or toward bundled prepayment for all medical care including drugs (such as through capitation for "all medically necessary services") also removes the incentive for physician-induced demand without the need to enforce separation. Thailand's "30 baht" public insurance program is a good example of such a case. Of course, the cost of this form of payment is the lack of incentive for quality effort and the possibility of underprovision, as in spot market transactions.

4.2 Heterogeneity

Clearly, the benefits and costs of physician dispensing can differ across population sub-groups. For example, the theory predicts that physician dispensing will be the more efficient institutional arrangement when the perceived costs associated with traveling to two different locations are sufficiently high. This arguably explains why hospitals dispense medications to inpatients, even when they are prohibited from dispensing to outpatients (as in Korea). Another, somewhat related case arises when the physician diagnosis and treatment services are inextricably linked to the dispensing of medication, such as for injections ostensibly requiring physician oversight.

Probably the most common exception to separation arises when separation is prohibitively costly because of insufficient human resources relative to high travel costs, such as in rural areas. Occupational distinctions between physicians and pharmacists have been blurred when one or the other of the professions was scarce relative to effective demand. This allowance for integrated prescribing and dispensing either took the form of pharmacists prescribing medications when physicians were not available (such as a royal ordinance in France in 1724 [Sonnedeker 1976, p.71], rural UK today, and many developing countries), or physicians being allowed to dispense directly to patients when pharmacies appear inadequate. In Switzerland, for example, some can-

tons (counties/districts) allow physician dispensing when pharmacy availability is below a certain range. However, physicians are still required to “inform patients about their right to obtain their prescription and buy the drugs in a pharmacy of their choice” (Rischatsch and Trottmann 2009, p.1), illustrating that modern governments even in Europe (with a very long history of separation) regulate patient ‘ownership’ of prescriptions, apparently as a device to preserve patient autonomy and thus constrain physician-induced demand.

A final important dimension of heterogeneity involves the distinction between acute care episodes and management of chronic disease. When a patient has a chronic disease, the travel costs associated with frequent adjustment of prescriptions may also weigh in favor of integrating some diagnosis and dispensing functions.

5 Evidence on historical episodes of separation

In this section, I first provide an overview of policy reforms dismantling physician dispensing in East Asia and then weigh competing explanations of the pattern of reforms in Asia and beyond.

5.1 Separation reforms in East Asia

Both Korea and Taiwan adopted separation policies soon after extending health insurance coverage to their entire populations. Japan had tried to impose western medicine and pharmacy on Korea during Japan’s colonial rule of the peninsula in the early 20th century, but largely failed (Eggleston and Yang 2009). It was not until after South Korea achieved universal coverage (in 1989) that it implemented a strict – and very politically contentious – separation of prescribing from dispensing in 2000. Physicians went on strike nationwide three times to protest separation reforms. The South Korean government granted concessions that included raising physician fees and indefinitely postponing other important policies for sustainable healthcare financing such as provider payment reform (Kwon and Reich 2005; Kim and Prah Ruger 2008; Eggleston and Yang 2009).

Taiwan implemented National Health Insurance in 1995, and began only two years later with separation reforms. In 2002 Taiwan implemented a further reduction in payment for drugs dispensed in physician clinics. Although Taiwan attempted to legislate full separation, the achieved level of separation is probably lower than that of Japan under a more incremental approach. Taiwanese hospitals could retain their outpatient pharmacies; the majority of physician clinics hired on-site

pharmacists; and “gateway pharmacies” (with ownership links to physicians) fill about 80 percent of prescriptions dispensed by Taiwan’s pharmacies (Chou et al. 2003, Hsieh 2009). Physician dispensing incentives also shape prescribing of generics compared to brand-name drugs (Liu, Yang, and Hsieh 2009).

Careful empirical analysis by Chen, Gertler, and Yang (2011) reveals that drug expenditures are probably inflated (at least) 30 percent by physician dispensing in Taiwan, because that is how far expenditures fell after enforcement of separation. (This figure may be an underestimate because it is derived from the 55 percent of physicians who did not find it worthwhile to use the policy exemption to continue dispensing, and therefore those who were presumably making the least profit from dispensing.) Moreover, the higher expenditures under physician dispensing did not provide measurable health benefits for patients, because when drug expenditures declined after separation, “patients treated at clinics without an onsite pharmacist did not have greater observable adverse health events than patients treated at integrated clinics” (Chen, Gertler, and Yang 2011, p.1).

Japan’s incremental separation policies probably have achieved greater removal of dispensing revenues from physician control. Separation, as measured by the share of prescriptions filled at pharmacies rather than dispensed by physicians, increased from only 20 percent in 1996 (Rodwin and Okamoto 2000, p.353) to 60 percent in 2007 (Tomita 2009). Japanese policy well illustrates the regulatory approach and political acceptance of heterogeneity, with policy allowing physicians and patients to self-sort between integration and separation. Incremental separation reforms began in earnest after Japan attained universal coverage (1961) and healthcare technologies and total spending grew rapidly (1970s and 1980s). Earlier separation attempts – during the Meiji reforms in 1874 and under the Allied occupation in 1951 – both failed (Tomita 2009). Those earlier reform failures are consistent with the theoretical proposition that physician dispensing was a self-reinforcing institution prior to the confluence of universal coverage with rapid technological change and population aging in the latter decades of the twentieth century. Regulations of the 1980s and 1990s included reduction in the allowed mark-up for drug dispensing; higher prescribing fees if the physician did not dispense; regulatory encouragement of independent pharmacies; and penalties for self-referral to “gateway pharmacies,” regulated as a form of double billing (Fujii and Reich 1988; Rodwin and Okamoto 2000). The dynamic pricing formula introduced in 1992 imposes larger reductions in retail prices for drugs that earn doctors higher margins between wholesale and retail prices (Iizuka 2007 and 2009). Adjustment of the medical and pharmaceutical fee schedules within a global budget

(Campbell and Ikegami 1998) also moderates the incentive and opportunity for physician-induced demand in Japan. Combined, these policies have moved Japan incrementally from a physician dispensing equilibrium toward separation as the dominant social institution.

In China, prescribing and dispensing have long been integrated. Soon after the communist revolution brought Mao Zedong to power, PRC policymakers chose to allow physicians and hospitals to sell medications directly to patients at a mark-up of 15 percent above acquisition cost. The choice to make the profit margin on dispensing drugs positive seemed logical and appropriate, the theory would suggest, because it reflected an institutional element defining clinical reality for centuries in China, and the level of technology and per capita ability to pay suggested that there was little scope for overuse. The allowed margin of 15 percent for drug dispensing had little effect on behavior prior to the 1980s because physicians were salaried public servants and hospitals were largely government-budget financed. Since then, however, physicians and hospitals have come to rely significantly on drug dispensing revenues from charging patients mark-ups over acquisition costs. Social dissatisfaction with perceived over-charging and over-treatment has increased. “Supporting medical services through drug sales” (*yi yao yang yi*) has been widely criticized amongst mounting evidence that such financial incentives distort prescribing and contribute to rising expenditures. In one excellent recent study, Currie, Lin, and Zhang (2010) audit the antibiotic prescribing behavior of hospital-based physicians in two cities and one rural area using student “simulated patients” during the 2008 and 2009 flu seasons. They find that Chinese physicians prescribe antibiotics for a startlingly high proportion of patients (averaging 62 percent), even when patients report symptoms that do not warrant antibiotics; and 39 percent of physicians still prescribed antibiotics when the simulated patients signaled to doctors that they knew that taking antibiotics would be inappropriate. These results provide strong evidence of physician-induced demand in China, with adverse consequences not only for medical spending but also for patient health and development of antibiotic resistance.

Alongside expanded insurance and improved public health, separation is one of the stated objectives of China’s 2009 national health reform plan (Standing Conference of the State Council of China 2009), for which the government committed 850 billion RMB (about 125 billion US dollars). Specifically, government-owned primary care (“grassroots”) providers must sell drugs on the essential medications list without a mark-up, and pilot reforms of public hospitals are encouraged

to include separation.¹¹ For example, separation – removing hospital reliance on drug dispensing revenues – is one of the three major pillars of reform for government-owned hospitals in Beijing.¹² The theory developed in this paper suggests that this newfound interest in separation has been inspired in no small measure by China’s expanding insurance coverage, itself spurred by rapid technological change and income growth. Government-subsidized health insurance for rural residents and urban non-employees, added to pre-existing insurance for urban employees, has dramatically increased coverage: the number of uninsured in China fell from 77.9 percent in 2003 to only 12.9 percent in 2008, according to the National Health Services Surveys.¹³ The propositions above predict that China will adopt more rigorous separation policies as it (gradually) replaces demand-side constraints with supply-side constraints on health spending. The large increase in available medical technologies in China over the last several decades, combined with the current dramatic expansion of insurance, both push toward separation. Introducing insurance lowers the costs of separation eventually, because provider payment need no longer be defined by patient willingness to pay linked to expectation of obtaining a drug during a physician’s visit. Insurance not only makes separation ‘easier’ in this sense, it also makes separation increasingly necessary. Without a strong demand-side constraint, physician-induced demand can threaten the financial sustainability of social health insurance. Similar concerns appear to have been the primary antecedents of separation reforms in Korea and Taiwan, as the theory would predict.

However, the theory also suggests several reasons why some integration of prescribing and dispensing may be a self-enforcing institution in China for a longer period than in Japan or Korea. Limits on regulatory effectiveness suggest high transaction costs of enforcing separation policies in China. In addition, concerns about fake and low-quality pharmaceuticals suggest that there might be a nontrivial value of holding a physician or hospital accountable for the quality of medications

¹¹The official reform document of 2009 stipulated that “government-run health care institutions at grass-roots levels shall sell drugs with zero mark up” in both urban and rural areas. Essential medicines are to be covered by social insurance, “with the reimbursing rate much higher than that of non-essential medicines” (Yiyaoweisheng tizhi gaige jinqi zhongdian shishi fang’an 2009-2011nian 2009, as quoted in Eggleston 2010, p.257). Patients are also allowed to purchase drugs in retail pharmacies with a prescription. For public hospital reform pilots, the reform plan states that “the separation of health care services and drug sale should be promoted, gradually rescinding the drug price margin, and banning the acceptance of any drug procurement discount. The revenue reduction and losses incurred from the reform shall be resolved through introducing prescription fees, readjusting the charging criteria for some technical service, increasing government investment, and [other measures]. The prescription fees shall be integrated into the reimbursement scope of the basic medical insurance...” (Yiyaoweisheng tizhi gaige jinqi zhongdian shishi fang’an 2009-2011nian 2009, as quoted in Eggleston 2010, p.258).

¹²Gongli yiyuan gaige qiansanbu (The first three steps of [Beijing] public hospital reform) *Beijing Daily*, March 24, 2011.

¹³Data from the Ministry of Health Center for Health Statistics and Information, quote in Eggleston 2010, p.232.

dispensed, as was the case under the institution of physician dispensing (through reputation effects).

Empirical evidence from Japan (Iizuka 2007) and Taiwan (Liu, Yang, and Hsieh 2009), as well as Switzerland (Rischatsch and Trottmann 2009), confirms that integrated physicians take into account patient co-payments when prescribing medications, and thus are less likely to prescribe expensive medications when such a choice increases the financial burden on their patients.¹⁴ Iizuka's estimates suggest that to save a non-elderly patient 28 cents in co-payment, a Japanese physician in the 1990s was willing to give up one dollar of profit. Presumably this pattern arises because physicians want to defend their reputations, or are altruistic toward their patients, or both. Similarly, Rischatsch and Trottmann (2009) find that patient cost-sharing is more influential on prescribing decisions for physicians who are not allowed to dispense directly to patients, compared to dispensing physicians. This Swiss evidence is consistent with the model assumption and the evidence from Asia (further evidence that culture is not the determining factor). Overall, the fact that physicians appear to be less likely to prescribe according to profitability when doing so raises patient out-of-pocket payments supports my theoretical proposition that patient cost sharing serve to constrain physician-induced demand, and the introduction of insurance increases the potential social waste from physician-induced demand.

This evidence that patient out-of-pocket burden constrains physician prescribing also helps to explain the otherwise puzzling case of Singapore. Although in many ways similar to Taiwan and South Korea in terms of cultural roots and economic development, Singapore differs from those economies in continuing to integrate prescribing and dispensing. My hypothesis for this discrepancy is that Singapore's lack of separation reforms is closely linked to its lack of National Health Insurance systems like Taiwan and Korea. Singapore's health care financing system is heavily dependent on patient cost sharing (for example, in the form of payments from medical savings accounts or cash). Consider the contrast between Singapore and South Korea. Although per capita health spending is virtually the same (\$1,642.60 in Singapore and \$1,688.20 in South Korea according to purchasing power parity), patient out-of-pocket health expenditure as a percentage of total health expenditure was 63.3 percent in Singapore compared to 35.7 percent in South Korea in 2007 (and an even lower percentage in most other OECD countries).¹⁵ As these figures show, Singapore has a high

¹⁴This evidence is possible because the patient payment associated with a drug may differ for certain groups of the population or change during the period of observation. Iizuka's research, for example, exploits the fact that for the same medications in Japan, elderly patients pay lower co-payments (i.e., enjoy more generous insurance coverage) than non-elderly patients do.

¹⁵The statistics are drawn from the World Bank World Development Indicators, based on the

share of private financing, almost all in the form of out-of-pocket payment. This system translates into a relatively effective constraint on over-prescribing as well. Thus Singapore has never felt the pressure to disrupt patient and provider satisfaction with integration to constrain physician-induced demand. This reliance on demand-side constraints on health expenditures in Singapore can be considered the exception that proves the rule: few countries of high per capita income choose to have such high reliance on out-of-pocket financing, because of equity considerations that Singapore has addressed in other ways (see for example Lim 2004).

5.2 Weighing competing theories of separation

The theory developed in this paper represents only one among several possible explanations for separation reforms. This section weighs the evidence presented above against competing hypotheses, emphasizing which historical episodes of separation are well explained by the theory and which are not.

As noted earlier, culture strongly shapes the institutions governing physician-patient interactions. Health practices in general illustrate the strong force of cultural beliefs for institutional design, and belie the simplistic condemnation of culture as a “barrier to progress”: adherence to traditional Chinese medicine arguably was socially beneficial when the western “technology” was bloodletting, but negative when the western technological alternative was the germ theory of disease. Nevertheless, culture alone appears unable to explain the pattern of separation reforms described above, such as why South Korea enforced separation but Singapore has not.

Technological change clearly plays a large role in explaining institutions in sectors such as health care that have experienced profound technological change. The era of successful separation reforms in East Asia – roughly since Japan implemented universal coverage in the 1960s – certainly coincides with a period of rapid technological change in medical care in general, and in pharmaceuticals in particular. Specialization and division of labor leads to the regulatory question of whether organizations and functions that are now separate, such as physician and pharmacist services, can still be owned and controlled by one profession (physicians). Technology also shapes the opportunities for integration. Oncology has been a fertile area for integration because the technology of administration for chemotherapy drugs justifies integration even when separation prevails in other medical

World Health Organization National Health Account database (www.who.int/nha/en), downloaded from <http://databank.worldbank.org/ddp/home> on 23 March 2011.

specialties. However, technological change is perhaps best viewed as a necessary but not sufficient condition for separation. The pure theory of technological change, divorced from considerations of insurance and physician-induced demand, can explain neither the timing of separation reforms within countries (i.e. when Korea chose to implement separation) nor the pattern of such reforms across countries (why Taiwan separated but Singapore did not).

A third hypothesis competing with the model developed here is that separation reforms are predominantly a story of professional rent-seeking behavior. In other words, integration or separation are determined by the political economy of guilds and other physician and pharmacist interest groups vying for dispensing revenues. This hypothesis focuses exclusively on the political power of physician interest groups, which in my model is only one factor (Ω) in the costs of enforcing separation, balanced against the costs of incentive for physician-induced demand.

The professional rent-seeking hypothesis has several strengths in explaining the historical experience of Europe, where separation dates back at least to the Edict of Palermo in 1231. Several European states began regulating separation fairly early, indicating that costs of overcoming interest groups to enforce separation (Ω) were low. Indeed, the Edict of Palermo stipulated that physicians and apothecaries were not to enter into business relationships, and that the government had authority over the number of apothecaries, their locations and prices. Since at that early date neither technology nor the economic costs of its wasteful overuse were very large, separation policies that spread across continental Europe would appear to be better explained by the dynamics of guilds and rent-seeking behavior over the control of dispensing revenues (Sonnedecker 1986; Savage 1994; Anderson 2005).¹⁶

While there is substantial evidence that the role of guilds in China (*huiguan* or *gongsuo*) differed substantially from that of Europe,¹⁷ it is less clear that guild dynamics fully explain the inertia of physician-dispensing through to the modern era in China and elsewhere in East Asia. Mao Zedong’s health policymakers, newly installing a communist government for China and expropriating hospitals for government ownership, did not appear to be concerned with physician or pharmacist

¹⁶I thank an anonymous reviewer for emphasizing this point: “The separation of diagnosis and dispensing in Europe was the result of guild regulations. Most “state” regulation in the medieval and early modern periods, like the Edict of Palermo, were actually codifications (or applications on a larger geographic scale) of already existing guild regulations, whose origins lay in the dynamics of rent-seeking among specialists in different stages of the production process – certainly not in any concern about physician-induced demand. In turn, more separation in Europe perhaps meant a more professional and less homespun pharmaceutical tradition. England and the US had weaker guild power, and so not surprisingly had a more varied organizational pattern than continental Europe.”

¹⁷For recent accounts that describe Chinese merchant and craft guilds (*huiguan* and *gongsuo*) and their role in China’s economy, see Liu 1988; Fairbank and Goldman 2006; Shiue and Keller 2007.

associations when they officially sanctioned a 15 percent mark-up on drugs for hospitals. Moreover, there is no evidence to suggest that weakening of physician interest groups triggered later separation reforms. In fact, physician interest groups continue to be strong in most of East Asia, and they have repeatedly played a powerful role in delaying separation and extracting concessions from regulators when separation was enforced. South Korea's reforms are particularly instructive on this point. Kwon and Reich (2005) analyze the politics of Korea's separation reforms in detail, compared to two other recent episodes of health reform (financial integration and provider payment reform). They show that physicians blocked separation reforms for a long time, and constituted a more powerful interest group in the 2000 reforms than the pharmaceutical industry, pharmacists, patients, or other affected parties. According to their analysis, the separation reforms eventually passed only because "the change of government, the president's keen interest in health policy, and democratization in the public policy process toward a more pluralist context opened a policy window for reform" (Kwon and Reich 2005, p.1003).¹⁸

Other separation reforms in East Asia also illustrate the continuing influence of physician organizations in shaping policy. In Taiwan, physicians extracted concessions so that they were able to hire on-site pharmacists (which pharmacists did not oppose), and physicians increased orders for other profitable services, emasculating the reform's objective to control expenditures (Chen, Gertler, and Yang 2011). Japan incrementally reformed dispensing practices, since attempts to legislate strict separation failed, arguably because physician interest groups opposed more stringent reforms. In China since 2009, separation reforms have only been imposed (with some difficulty) on the weakest providers, the grassroots clinics and community health centers in rural and urban areas, rather than on the major players, the hospitals. Two years after national reforms emphasized the desirability of separation, for example, the 2011 official plan for hospital reform in Beijing municipality calls for separation as a long-term goal but acknowledges that "conditions are not yet ripe for full implementation of separating prescribing and dispensing" (*shishi yiyao fenkai de*

¹⁸ "The strikes organized by physicians (both office-based and private hospital-based) in 2000 panicked the entire health care system. . . . After a series of nationwide strikes, physicians gained a much stronger voice in policy decisions—and the government had to deny civic groups a seat at the negotiation table. Physicians pushed the government to change the original version of the pharmaceutical reform package. Physicians blocked the use of generic prescriptions, protected their right to prescribe brand-name drugs, increased the proportion of prescription drugs relative to nonprescription drugs, and overturned the government plan of including injection drugs in the reform package (which the government hoped to include in order to reduce the chronic problem of overuse). Most notably, physician strikes drove the government to raise the reimbursement fees for physician services by 44 percent, as compensation for income loss caused by the pharmaceutical reform. The strong influence of physicians on the Korean pharmaceutical reform is in contrast to pharmacists, who did not develop a political power as strong as physicians and accepted the reform reluctantly" (Kwon and Reich 2005, pp. 1017-1018).

tiaojian shang bu chengshu; Gongli yiyuan gaige qiansanbu 2011). Importantly, in none of these cases have pharmacist associations played a significant political role.

Thus, in the case of modern East Asia, separation has been associated less with physician-pharmacist professional jockeying over dispensing revenues than with a confluence of forces overpowering physician interests to impose separation. The theory developed in this paper – emphasizing the interplay between cultural expectations (T_t), interest groups (Ω_t), technological change, insurance expansion and government financing (P^H) – appears to offer a plausible explanation of separation reforms in Asia. This argument emphasizes a mechanism complementary to, but distinct from, the technological change and rent-seeking hypotheses, by which technological change shaped the political economy of forces impinging on the doctor-patient relationship. In this telling, insurance and government financing play the trigger role in overcoming the inertia of the old physician-dispensing regime: insurance coverage reduces the demand-side constraint on over-prescribing, and since government budgets now bear a substantial fraction of the associated expenditures, government officials develop ways to encourage (Japan) or impose (Korea) separation.

An outside observer might naively suggest that physician dispensing endures in some parts of East Asia only because patients and regulators are ignorant about the incentive for over-prescription that physician dispensing implies. To the contrary, one implication of the theory is that patients are aware of doctors' propensity to induce demand; physician dispensing is self-enforcing despite physician-induced demand because patients are willing to compensate physicians for effort through extra payments for medications. Indeed, payment of a bonus for inducement, $\beta > 0$, serves to relax the reneging constraint for physician dispensing and increase the incentive for quality effort, so that the physician chooses higher levels of effort on both tasks compared to what would have been chosen under separation (see physician dispensing reneging constraint (13) above and Baker, Gibbons and Murphy's discussion of their Result 2, pp.66-67). Consistent with this rational tolerance of physician-induced demand (up to a point), patients in East Asia appear to be aware of doctors' incentives to over-prescribe.¹⁹

Ironically, separation may not reduce drug expenditures or overall health expenditures, for several reasons. First, drugs with the highest profit margin for dispensing physicians are not always the drugs that are most expensive. For example, in both Taiwan and Switzerland (Liu, Yang, and

¹⁹For example, some hospitals in China have found it profitable to self-adopt a fixed price for specific hospitalizations to attract more patients, clear evidence that patients are aware of hospitals' incentives to over-treat and that fixed prices mitigate that problem (Eggleston 2009).

Hsieh 2009; Rischatsch and Trottmann 2009) generic drugs appear to have been more profitable for physicians than brand-name drugs were, so integration pushes toward generic substitution, which actually helps to constrain overall health expenditures. Second, well-insured patients seek the highest quality treatment; with profit considerations removed, physicians may simply indulge patient moral hazard or compete on perceived quality by prescribing brand-name drugs. Such behavior would tend to increase drug spending after separation. Third, physicians may extract other concessions that tend to increase expenditures, such as delay of provider payment reform, as well as induce demand for non-pharmaceutical services to compensate for lost drug revenues, as evidenced by empirical studies of oncologists in the US (Jacobson et al 2010) and general practitioners in Taiwan (Chen, Gertler, and Yang 2011).

Finally, although the theory developed here was motivated by separation reforms in Asia, it is worth considering how experience in other parts of the world reinforce or conflict with this account. As already acknowledged, the rent-seeking hypothesis probably better fits the experience of Europe, at least prior to the 20th century, and may well provide a satisfactory account of jockeying among physicians and pharmacists in other regions such as South Africa (Gilbert 2001). In the UK and the US, some of the forces at play in the model appear to have some traction. Separation came later than in Europe, presumably at least in part because of weaker guilds and larger contracting and travel costs. Interestingly, the English apothecary “understood that he was not permitted to charge for his consultation, and was quite prepared to rely on the sale of drugs and medicines for his profit” (Worling 2005, p.60); these apothecaries developed into the general practitioners of today. As the theory suggests, insurance coverage did help to spur separation: the National Insurance Bill of 1911 and the insurance scheme implemented two years later separated prescribing and dispensing for insured patients, becoming nationwide after the establishment of the National Health Service (Hunt 2005). At least four elements of the theory suggest that separation would have come later in the US, as indeed it did: lower organizational power of physician guilds, relatively smaller capacity and willingness to impose regulation, high travel costs, and lack of universal coverage. Even so, the US experience with Medicare revising reimbursement policies for chemotherapy drugs clearly illustrates the nexus between government insurance, regulatory oversight, and suspicions of over-prescribing when physicians directly dispense medications to patients (see for example the careful empirical study by Jacobson et al. 2010 and the Government Accountability Office reports cited therein). Further evidence that integrated diagnosis and treatment gives rise to physician-induced demand

comes from Afendulis and Kessler (2007), who econometrically show that diagnosis of an elderly American by a heart specialist who also provides surgical treatment (“integrated cardiologist”), compared to a non-integrated cardiologist, increases health spending without improving health outcomes. Thus physician-induced demand and moral hazard from insurance arrangements appear to be health policy concerns in many parts of the world, as well as central determinants of separation reforms.

6 Conclusion

Why did policymakers throughout East Asia systematically choose to legitimize drug sales as the profit center for doctors and hospitals until recently? Important institutional elements from the past – especially shared beliefs about the proper form of a physician-patient encounter and legitimate source of revenue for providers – shaped the options available. In the words of Greif (2006), there is a “fundamental asymmetry” between the structure inherited from the past and the technologically feasible alternatives. These costs of enforcing separation will be greater, the longer the time period over which provider interest groups have coalesced around the rent stream associated with physician dispensing and the more generations of patients have come to expect receiving a medication as the ‘payoff’ from visiting a physician.

However, shifts of ‘quasi-parameters’ can render long-standing practices self-undermining. Eventually the expanding capabilities of medicine – both endogenously produced and exogenously imported through development of the global pharmaceutical and medical care technology industries – undermined the physician dispensing institution by increasing the social opportunity costs of induced demand. The rise of the welfare state and third-party payers tended to remove demand-side constraints on health spending and prompt governments to search for social coalitions to overcome the interests of physician dispensers and impose separation. Although remolding patient and provider beliefs and norms often takes time, these potent forces led to new institutions governing the patient-physician relationship even in societies with long traditions of physician dispensing, like those of East Asia.

This case study of physician dispensing serves to illustrate that health systems provide a rich field for testing hypotheses of institutional economics. Patterns of incentives, so vital to economic performance, are neither random nor perfectly malleable to the design of social planners. The

incentive structure of current healthcare delivery systems, for example, have deep historical and cultural roots, yet must cope with rapid technological change as well as market and government failures. Understanding the evolution of health sector institutions can thus provide valuable illustrations of how and why some institutions are more tenacious or resilient than others, and how these qualities are associated with economic performance.

Physician-dispensing also teaches about institutional transplantation (Acemoglu, Johnson, and Robinson 2001; Yoo and Steckel 2010). Governments have tried to impose separation in East Asia (as Japan attempted in colonial Korea) or to import the practice. Japan readily adopted western ideas and technologies during the Meiji era and after World War II, so why did the separation efforts in both Meiji Japan (1871) and under the Allied occupation (1951) fail? These aborted efforts at institutional transplantation show the importance of understanding the historic context of social institutions. In the case study explored here, physician dispensing appears to have been a self-reinforcing institution prior to the confluence of universal coverage with rapid technological change and population aging in the later decades of the twentieth century.

References

- [1] Abbott, Andrew. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.
- [2] Acemoglu, Daron, Johnson, Simon, Robinson, James A. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91, 1369-1401.
- [3] Afendulis, Christopher C., Kessler, Daniel P. 2007. Tradeoffs from integrating diagnosis and treatment in markets for health care. *American Economic Review* 97, 1013-1020.
- [4] Aoki, Masahiko. 2001. *Toward a Comparative Institutional Analysis*. Comparative Institutional Analysis Series; Cambridge and London: MIT Press.
- [5] Arrow, Kenneth J. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* 53, 941-973.
- [6] Baker, George, Gibbons, Robert, Murphy, Kevin J. 2002. Relational contracts and the theory of the firm. *Quarterly Journal of Economics* 117, 39-84.

- [7] Bisin, Alberto, and Thierry Verdier. 2010. The economics of cultural transmission and socialization. National Bureau of Economic Research Working Paper No. 16512, November 2010.
- [8] Campbell, John C., Ikegami, Naoki. 1998. *The Art of Balance in Health Policy: Maintaining Japan's Low-Cost, Egalitarian System*. Cambridge: Cambridge University Press.
- [9] Cavalli Sforza L.L., and Marc Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.
- [10] Chen, Brian K., Gertler, Paul J., Yang, Chun-Yuh. 2011. Physician ownership of non-physician medical services. Working paper, February 21, 2011.
- [11] Chou, Yiing J., Yip, Winnie C., Lee, C. H., Huang, N., Sun, Y. P., Chang, H. J. 2003. Impact of separating drug prescribing and dispensing on provider behavior: Taiwan's experience. *Health Policy and Planning* 18, 316-329.
- [12] Cochran, Sherman. 2006. *Chinese Medicine Men: Consumer Culture in China and Southeast Asia*. Cambridge and London: Harvard University Press.
- [13] Currie, Janet, Wanchuan Lin, Wei Zhang. 2010. Patient knowledge and antibiotic abuse: Evidence from an audit study in China. National Bureau of Economic Research Working Paper 16602, December 2010.
- [14] Darby, Michael, Karni, Edi. 1973. Free competition and the optimal amount of fraud. *Journal of Law and Economics* 16: 67-88.
- [15] Deaton, Angus. 2004. Health in an age of globalization. Brookings Trade Forum 2004, 83-130.
- [16] Eggleston, Karen (Ed.), 2009. *Prescribing Cultures and Pharmaceutical Policy in the Asia Pacific*. Stanford University Shorenstein Asia-Pacific Research Center, distributed by Brookings Institution Press.
- [17] Eggleston, Karen, Yang, Bongmin. 2009. Physician and pharmacists in comparative historical perspective: The case of South Korea. In: Eggleston, Karen (Ed.), *Prescribing Cultures and Pharmaceutical Policy in the Asia Pacific*. Stanford University Shorenstein Asia-Pacific Research Center, distributed by Brookings Institution Press, pp. 267-280.

- [18] Eggleston, Karen. 2010. ‘Kan bing nan, kan bing gui’: Challenges for China’s healthcare system thirty years into reform. In: Jean C. Oi, Scott Rozelle, and Xueguang Zhou (Ed), *Growing Pains: Tensions and Opportunities in China’s Transformation*. Stanford, CA: Walter H. Shorenstein Asia-Pacific Research Center, distributed by Brookings Institution Press.
- [19] Fairbank, John King, and Merle Goldman. 2006. *China: A New History*. Cambridge, MA: Harvard University Press.
- [20] Fong, Yuk-fai. 2005. When do experts cheat and whom do they target? *Rand Journal of Economics* 36, 113-130.
- [21] Fujii, Mitsuru, Reich, Michael. 1988. Rising medical costs and the reform of Japan’s health insurance system. *Health Policy* 9, 9-24.
- [22] Gilbert, Leah. 2001. To diagnose, prescribe and dispense: Whose right is it? The ongoing struggle between pharmacy and medicine in South Africa. *Current Sociology* 49, 97
- [23] Glaeser, Edward L., Shleifer, Andrei. 2003. The rise of the regulatory state. *Journal of Economic Literature* 41(2), 401-425.
- [24] Gongli yiyuan gaige qiansanbu (The first three steps of [Beijing] public hospital reform) *Beijing Daily*, March 24, 2011, available at http://www.bjd.com.cn/10bjxw/ss/201103/t20110324_667472.html (accessed 24 March 2011).
- [25] Gorodnichenko, Yuriy, Roland, Gerard. 2010. Culture, institutions and the wealth of nations. National Bureau of Economic Research Working Paper No. 16368, September 2010.
- [26] Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Political Economy of Institutions and Decisions series; Cambridge and New York: Cambridge University Press.
- [27] Grossman, Sanford, Hart, Oliver. 1986. The costs and benefits of ownership: A theory of vertical and lateral ownership. *Journal of Political Economy* 94, 691-719.
- [28] Hall, Robert E., Jones, Charles I. 2007. The value of life and the rise in health spending. *Quarterly Journal of Economics* 122, 39-72.

- [29] Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Oxford University Press.
- [30] Hsieh, Chee-Ruey. 2009. Pharmaceutical policy in Taiwan, In: Eggleston, Karen (Ed.), *Prescribing Cultures and Pharmaceutical Policy in the Asia Pacific* Shorenstein Asia-Pacific Research Center series with Brookings Institution Press, pp.109-128.
- [31] Hubbard, Thomas N. 2002. How do consumers motivate experts? Reputational incentives in an auto repair market. *Journal of Law and Economics* 45, 437-468.
- [32] Iizuka, Toshiaki. 2007. Experts' agency problems: Evidence from the prescription drug market in Japan. *RAND Journal of Economics*, 38(3): 844-862.
- [33] ———. 2009. The economics of pharmaceutical pricing and physician prescribing in Japan. In: Eggleston, Karen (Ed.), *Prescribing Cultures and Pharmaceutical Policy in the Asia Pacific* Shorenstein Asia-Pacific Research Center series with Brookings Institution Press, pp.47-60.
- [34] Jacobson, Mireille, Craig C. Earle, Mary Price, and Joseph P. Newhouse. 2010. How Medicare's payment cuts for cancer chemotherapy drugs changed patterns of treatment. *Health Affairs* 29, 1391-1399
- [35] Kim, Hak-Ju, Prah Ruger, Jennifer. 2008. Pharmaceutical reform in South Korea and the lessons it provides. *Health Affairs* 27, w260-w269.
- [36] Kleiner, Morris M., Park, Kyoung Won. 2010. Battles among licensed occupations: Analyzing government regulations on labor market outcomes for dentists and hygienists. National Bureau of Economic Research Working Paper No. 16560, November 2010.
- [37] Kwon, Soonman, Reich, Michael R. 2005. The changing process and politics of health policy in Korea. *Journal of Health Politics, Policy and Law* 30, 1003-1025.
- [38] Lim, Meng-kin. 2004. Shifting the burden of health care finance: A case study of public-private partnership in Singapore. *Health Policy* 69, 83-92.
- [39] Liu, Kwang-Ching. 1988. Chinese merchant guilds: An historical inquiry. *Pacific Historical Review* 57, 1-23.

- [40] Liu, Ya-Ming, Yea-Huei Kao Yang, and Chee-Ruey Hsieh. 2009. Financial incentives and physicians' prescription decisions on the choice between brand-name and generic drugs: Evidence from Taiwan. *Journal of Health Economics* 28, 341-349.
- [41] Lundin, Douglas. 2000. Moral hazard in physician prescription behavior. *Journal of Health Economics* 19, 639-662.
- [42] McGuire, Thomas G. 2000. Physician agency. Chapter 9. In: A.J. Culyer and Joseph P. Newhouse, eds. *Handbook of Health Economics*. Amsterdam: Elsevier Science.
- [43] Murphy, Kevin M., Topel, Robert H. 2006. The value of health and longevity. *Journal of Political Economy* 114, 871-904.
- [44] Newhouse, Joseph P. 1992. Medical care costs: How much welfare loss? *Journal of Economic Perspectives* 6, 3-21
- [45] North, Douglass C. 2005. *Understanding the Process of Economic Change*. Princeton: Princeton University Press.
- [46] Porter, Roy. 1997. *The Greatest Benefit to Mankind: A Medical History of Humanity from Antiquity to the Present*. London: HarperCollins.
- [47] Posner, Richard A. 2010. From the new institutional economics to organization economics: with applications to corporate governance, government agencies, and legal institutions. *Journal of Institutional Economics* 6(1): 1-37.
- [48] Rischatsch, Maurus, Trottmann, Maria. 2009. Physician dispensing and the choice between generic and brand-name drugs—Do margins affect choice?" University of Zurich Socioeconomic Institute Working Paper No. 911, July 2009.
- [49] Rodwin, Marc A., Okamoto, Etsuji. 2000. Physicians' conflicts of interest in Japan and the United States: Lessons for the United States. *Journal of Health Politics, Policy and Law* 25, 343-376.
- [50] Savage, Deborah Anne. 1994. The professions in theory and history: The case of pharmacy. *Business and Economic History* 23, 129-160.

- [51] Shiue, Carol H, Keller, Wolfgang. 2007. Markets in China and Europe on the eve of the industrial revolution. *American Economic Review* 97, 1189-1216
- [52] Sonnedecker, Glenn. 1986. *Kremers and Urdang's History of Pharmacy*. Philadelphia: American Institute of the History of Pharmacy.
- [53] The standing conference of State Council of China adopted Guidelines for Furthering the Reform of Health-care System in principle. [<http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohbgt/s3582/200901/38889.html>, accessed 20 August 2009].
- [54] Starr, Paul. 1982. *The Social Transformation of American Medicine*. Basic Books.
- [55] Sun, Qiang, Santoro, Michael A., Meng, Qingyue, Liu, Caitlin M, Eggleston, Karen. 2008. Pharmaceutical policy in China. *Health Affairs* 27, 1042-1050.
- [56] Taylor, Curtis R. 1995. The economics of breakdowns, checkups, and cures. *Journal of Political Economy* 103, 53-74.
- [57] Tomita, Naoko. 2009. The political economy of incrementally separating prescription from dispensation in Japan. In: Eggleston, Karen (Ed.), *Prescribing Cultures and Pharmaceutical Policy in the Asia Pacific* Shorenstein Asia-Pacific Research Center series with Brookings Institution Press, pp.61-78.
- [58] Truffer, Christopher J., Sean Keehan, Sheila Smith, Jonathan Cylus, Andrea Sisko, John A. Poisal, Joseph Lizonitz and M. Kent Clemens. 2010. Health spending projections through 2019: The recession's impact continues. *Health Affairs* 29, 522-529.
- [59] Wolinsky, Asher. 1993. Competition in a market for informed experts' services. *Rand Journal of Economics* 24, 380-398.
- [60] Worling, Peter. 2005. "Pharmacy in the Early Modern World, 1617 to 1841 AD," chapter 5 in *Making Medicines: A Brief History of Pharmacy and Pharmaceuticals*. Stuart Anderson, ed. London: Pharmaceutical Press, 2005.
- [61] Yoo, Dongwoo, and Richard H. Steckel. 2010. Property rights and financial development: The legacy of Japanese colonial institutions. National Bureau of Economic Research Working Paper No. 16551, November 2010.

7 Appendix

This appendix (1) presents the basic model in formal mathematical terms; (2) describes spot market contracting between patient and physician; (3) shows the conditions necessary for relational contracting to dominate spot market transactions for both physician dispensing and separation; (4) considers how physician altruism impacts equilibrium outcomes, using the case of physician dispensing to illustrate; and (5) extends the model to include a physician role in drug discovery (e.g. for traditional herbal medications).

7.1 A simple model of physician dispensing

7.1.1 Physician efforts

Each period, the physician chooses two actions, quality-related effort e_1 and effort to induce demand, e_2 , at cost $c(e_1, e_2)$. No one except the physician herself can observe the actions e_1 and e_2 .

Patient benefit from physician quality effort, Q_i , can take two values: Q_o or Q_H . The minimum contractible quality is normalized to zero: $Q_o \equiv 0$. With probability $q(e_1)$, patient benefit increases to $Q_H > 0$. In other words, the surplus associated with physician quality effort is probability $q(e_1)$ of benefit $Q_i = Q_H$. Note that $\Delta Q = Q_H - Q_o = Q_H$.

Let P_j represent additional use of resources on treatment (such as multiple and excessive prescriptions) that do not benefit the patient, but bring additional revenues to the physician and for which the patient may be willing to pay as long as dispensing a medication is the expected outcome of the physician-patient interaction. This ‘physician-induced demand’ can also represent monopoly pricing (above marginal benefit) for innovative drugs. P_j can range from 0 to P_H . We capture this range by assuming that with probability $p(e_2)$, the physician succeeds in inducing demand to $P_j = P_H$; otherwise (with probability $1 - p(e_2)$), $P_j = 0$.

In institutional equilibrium, patient and physician both observe Q_i and P_j . Of course the full benefit (Q_i) and the extent of physician-induced demand (P_j) will not be evident at the time the physician writes a prescription or dispenses the medication to the patient. However, in equilibrium with multiple interactions (and the patient’s recourse to self-diagnosis and self-treatment), we assume the patient can eventually observe (average) treatment effectiveness and deduce the extent of physician-induced demand.

7.1.2 Physician and patient objectives

Assume the patient seeks to maximize utility given by perceived benefit from treatment, Q_i , less payment to the physician.

$$D^K = E[Q_i]^K - \text{payment}^K = q(e_1) Q_H - \text{payment}, \quad (2)$$

where K represents the institutional setting: $K = MDD$ (physician-dispensing) or SPD (separation of prescribing from dispensing). The payment includes R for the contractible minimum quality, and additional bonuses that depend upon the institutional setting.

To be sustainable, the equilibrium payment must also fulfill an ability-to-pay or liquidity constraint. Let the patient's fraction of the healthcare payment (the co-insurance rate) be θ , where $1 \geq \theta \geq 0$. The maximum patient out-of-pocket spending cannot exceed the patient's ability to pay, W :

$$\theta * (\text{payment}) \leq W. \quad (3)$$

An uninsured patient has $\theta = 1$ and thus faces the tightest constraint.²⁰

The upstream physician seeks to maximize revenue less the cost of effort,

$$U^K = \text{Payment}^K - c(e_1^K, e_2^K).$$

A provider might also put some weight on patient benefit, so that provider utility would be $U = \text{Payment} + \alpha D - c(e_1, e_2)$. As one might intuitively expect, higher provider 'benevolence' or altruism, α , increases quality effort for any given payment, and decreases physician-induced demand to the extent that physician-induced demand decreases patient net benefit (see appendix section below on physician altruism).

7.1.3 The socially efficient benchmark

²⁰With "first dollar" insurance ($\theta = 0$) the constraint would never bind even for the lowest wealth levels. If the population were evenly distributed between maximum and minimum wealth levels, then the higher the co-insurance requirement θ , the larger the fraction of the population for which the ability to pay constraint (3) will bind. We focus on a single representative patient of wealth W , abstracting from financing of the health insurance or health service premiums (presumably tax-based) when $\theta < 1$.

The socially efficient effort choices maximize social welfare, $S \equiv D^K + U^K$ given by

$$S \equiv q(e_1)Q_H - c(e_1, e_2). \quad (4)$$

The value of the physician-patient relationship would be highest if the physician invested only in quality-improvement effort e_1 and never induced demand ($e_2 = 0$). Assume that effort costs take the simple functional form $c(e_1, e_2) = \frac{(e_1)^2 + (e_2)^2}{2}$, and that $q(e_1) = qe_1$ ($q > 0$) and $p(e_2) = pe_2$ ($p > 0$). Then the socially efficient physician efforts are

$$\begin{aligned} e_1^* &= qQ_H, \text{ and} \\ e_2^* &= 0, \end{aligned} \quad (5)$$

leading to social surplus $S^* = \frac{q^2 Q_H^2}{2}$.

7.1.4 Relational contracting

Under physician dispensing, the patient pays the physician in three components: R for the contractible minimum quality; bonus b_H if $Q_i = Q_H$; and payment β_H if $P_j = P_H$ (otherwise, $b_i = \beta_j = 0$). The patient is unable to distinguish P from Q in all cases, and is willing to pay the physician a premium or monopoly price ($\beta_H > 0$) to assure quality effort. We assume that $Q_H \geq b_H > Q_H/2$ and $P_H \geq \beta_H > P_H/2$ under physician dispensing.

If the physician believes the patient will follow through with these promised payments, the physician chooses efforts accordingly:

$$\underset{\langle e_1, e_2 \rangle}{Max} [R + qe_1b_H + pe_2\beta_H - c(e_1, e_2)] \equiv U(e_1^{MDD}, e_2^{MDD}).$$

The equilibrium efforts under relational physician dispensing would then be

$$\begin{aligned} e_1^{MDD} &= qb_H, \\ e_2^{MDD} &= p\beta_H, \end{aligned} \quad (6)$$

yielding physician payoff $U^{MDD} = R + \frac{q^2 b^2 + p^2 \beta^2}{2}$. The patient expects a payoff of $D(e_1^{MDD}, e_2^{MDD}) =$

$q^2b(Q_H - b) - R - p^2\beta^2$. Total surplus under relational physician dispensing is therefore

$$S^{MDD} = q^2bQ_H - \frac{q^2b^2 + p^2\beta^2}{2}. \quad (7)$$

Comparing (6) with (5), we see that the socially efficient levels of effort require $\beta_H = 0$ (which is not possible under physician dispensing unless physician-induced demand harms the patient and the physician is altruistic enough) and either $b_H = Q_H$ or sufficient physician altruism ($\alpha > 0$) to choose e_1^* despite $b_H < Q_H$ (see sub-section on altruism below).

To be sustainable, relational physician dispensing must fulfill not only the reneging constraint but also the ability-to-pay or liquidity constraint, (3). The maximum patient out-of-pocket spending ($\max(b_i + \beta_j)$) cannot exceed the patient's ability to pay, W : $\theta * (Q_i + P_j) \leq W$, or

$$P_H \leq \frac{W}{\theta} - Q_H. \quad (8)$$

In addition, we assume that doctors' maximum ability to induce demand depends on available technology, $TECH$.

Claim 7 $P^H(\theta; TECH)$ with $\frac{\partial P^H}{\partial TECH} > 0$.

7.2 Separation of prescribing and dispensing

Under a one-time physician-patient interaction with independent pharmacies, the patient utility is $D^{SSPD} \equiv Q_o - R$ and the physician utility is $U^{SSPD} = R$; the resulting social surplus is $S^{SSPD} = Q_o - R + R = Q_o = 0$.

Since the physician no longer reaps benefits from dispensing, the patient need not promise to pay for P_j : $\beta_H = 0$ and therefore $e_2^{SPD} = 0$. Physician utility becomes $U^{SPD} = R + qe_1b_H - c(e)$, yielding quality effort $e_1^{SPD} = qb_H$. Thus $U^{SPD} = R + \frac{(qb_H)^2}{2}$.

Patient utility under separation relational contracting is $D^{SPD} \equiv E[Q_i - R - b_i \mid e_1 = e_1^{SPD}] - \Omega - T_t$. $T_t + \Omega$ is the per-period cost of a self-enforcing relationship under separation (which need not be paid again if the relationship breaks down and the parties revert to spot market interactions). The patient expects a payoff under relational separation of $q^2b_H(Q_H - b_H) - R - \Omega - T_t = D^{SPD}$. The patient can choose to honor the contract, paying b_i and receiving $\frac{1}{r}D^{SPD}$ in the future. Alternatively, the patient can renege, which would save the current bonus payment but

imply no physician effort above the minimum contractible level in the future. Using the simplified functional forms,

$$S^{SPD} = q^2 bQ - \frac{q^2 b^2}{2} - \Omega - T_t, \quad (9)$$

which can be written as $bQ - \frac{b^2}{2} - \Omega - T_t$ if $q = 1$.

For simplicity assume $q = p = 1$. The self-enforcing institution of physician dispensing dominates the self-enforcing institution of separation if and only if $bQ - \frac{(b^2 + \beta^2)}{2} > bQ - \frac{b^2}{2} - \Omega - T_t$, or $\frac{\beta^2}{2} < \Omega + T_t$. Assuming $\max(\beta_H) = P^H$, then physician dispensing dominates separation when (1) holds true.

7.3 Spot market contracting

Assume a one-time physician-patient interaction under physician dispensing ('spot outsourcing') takes the following form: the patient offers to pay Q_i for diagnosis and prescription, and the physician recommends a higher amount P_j consistent with incentives to induce demand. Nash bargaining determines the final payment (amount of medication and its price): $\frac{Q_i + P_j}{2}$. Anticipating this outcome, the physician chooses the vector of 'spot physician-dispensing' efforts $e^{SMDD} \equiv (e_1^{SMDD}, e_2^{SMDD})$ according to

$$\underset{\langle e_1, e_2 \rangle}{Max} \left[\frac{1}{2} (qe_1 Q_H + pe_2 P_H) - c(e_1, e_2) \right],$$

so that

$$\begin{aligned} e_1^{SMDD} &= \frac{qQ_H}{2}, \\ e_2^{SMDD} &= \frac{pP_H}{2}, \end{aligned} \quad (10)$$

yielding physician payoff $U^{SMDD} = \frac{q^2 Q_H^2 + p^2 P_H^2}{8}$. The patient then receives payoff of

$$\begin{aligned} D^{SMDD} &= E \left[\frac{Q_i - P_j}{2} \mid e = e^{SMDD} \right] \\ &= \frac{q^2 Q_H^2 + p^2 P_H^2}{2}. \end{aligned}$$

Spot market physician dispensing thus yields surplus $S^{SMDD} = \frac{3}{8} (q^2 Q_H^2 - p^2 P_H^2)$.

7.4 When relational contracting dominates spot market transactions

First consider physician integration. The physician upholds relational contracting under physician dispensing if, for all i and j ,

$$\begin{aligned} b_i + \beta_j + \frac{1}{r}U^{MDD} &\geq \frac{1}{2}(Q_i + P_j) + \frac{1}{r}U^{SMDD}, \text{ or} \\ b_i + \beta_j - \frac{1}{2}(Q_i + P_j) &\geq \frac{1}{r}[U^{SMDD} - U^{MDD}]. \end{aligned} \quad (11)$$

The patient will honor the relational contract when the current net benefit $(Q_i - b_i - \beta_j)$ and the present value of the future relationship $(\frac{1}{r}D^{MDD})$ exceed the alternative, the net benefit from paying the spot market price $(Q_i - \frac{1}{2}(Q_i + P_j) = \frac{1}{2}(Q_i - P_j))$ and forever after having spot market physician-patient interactions $(\frac{1}{r}D^{SMDD})$. In other words, the patient renegeing constraint under physician dispensing is

$$\begin{aligned} Q_i - b_i - \beta_j + \frac{1}{r}D^{MDD} &\geq \frac{1}{2}(Q_i - P_j) + \frac{1}{r}D^{SMDD}, \text{ or} \\ b_i + \beta_j - \frac{1}{2}(Q_i + P_j) &\leq \frac{1}{r}(D^{MDD} - D^{SMDD}). \end{aligned} \quad (12)$$

Following Baker, Gibbons and Murphy's reasoning, if (11) holds for all i and j , then it must hold for the smallest value of $b_i + \beta_j - \frac{1}{2}(Q_i + P_j)$, which is 0. Moreover, if (12) holds for all i and j , then it must hold for the largest value of $b_i + \beta_j - \frac{1}{2}(Q_i + P_j)$, which arises when $b_H = Q_H$ and $\beta_H = P_H$, implying $b_i + \beta_j - \frac{1}{2}(Q_i + P_j) = \frac{1}{2}(Q_H + P_H)$. These two conditions can be combined to form the following necessary condition for physician dispensing to be self-enforcing:

$$\frac{1}{2}(Q_H + P_H) \leq \frac{1}{r}(S^{MDD} - S^{SMDD}), \quad (13)$$

where S^K denotes the total surplus achieved under each contracting arrangement, $K \in \{MDD, SMDD\}$.

Now consider separation (SPD) as an equilibrium institution. From the patient's perspective, separation can be self-enforcing when $-b_i + \frac{1}{r}D^{SPD} \geq \frac{1}{r}D^{SSPD}$, which can be re-written as $\frac{1}{r}(D^{SPD} - D^{SSPD}) \geq \max(b_i) = b_H$, or

$$\frac{1}{r}(qe_1^{SPD}Q_H - T_t - \Omega) \geq b_H. \quad (14)$$

The physician will honor the relational contract if $b_i + \frac{1}{r}U^{SPD} \geq \frac{1}{r}U^{SSPD}$, or

$$\min(b_i) = 0 \geq \frac{1}{r} (U^{SSPD} - U^{SPD}). \quad (15)$$

Combining these two reneging constraints, we find that the institution of separating prescribing from dispensing is self-enforcing if the relational (noncontractible) bonus payment b_H is not too great (relative to the payment for contractual minimum quality):

$$\begin{aligned} \max b_i - \min b_i &= b_H \leq \frac{1}{r} (D^{SPD} - D^{SSPD}) - \frac{1}{r} (U^{SSPD} - U^{SPD}), \text{ or} \\ b_H &\leq \frac{1}{r} (S^{SPD} - S^{SSPD}) = \frac{1}{r} S^{SPD}. \end{aligned}$$

The reneging constraint for separation is fulfilled when $b_H \leq \frac{1}{r} \left(bQ - \frac{b^2}{2} - \Omega - T_t \right)$, implying that r must be small enough to support reputational enforcement of separation.

7.5 Physician altruism

Assume that the physician places weight α on the patient's net benefit from treatment, $0 < \alpha < 1$. Physician utility under physician dispensing then becomes

$$U^{MDD} = qe_1b_H + pe_2\beta_H + \alpha D^{MDD} - \frac{(e_1)^2 + (e_2)^2}{2},$$

yielding choices of quality effort and physician-induced demand given by

$$\begin{aligned} e_1 &= q[(1 - \alpha\theta)b_H + \alpha Q_H], \\ e_2 &= (1 - \alpha\theta)p\beta_H. \end{aligned}$$

An altruistic physician ($\alpha > 0$) considers the impact of prescription decisions on both patient health (Q_H) and on patient out-of-pocket payment burden ($\theta(b_i + \beta_j)$). Then physician-induced demand is held in check not only by patient ability to pay but also by physician consideration of patient co-payment burden ($\frac{de_2}{d\theta} < 0$) and any potential adverse health impact or risk for the patient (i.e., if $\frac{dQ_H}{de_2} < 0$). Physician altruism thus relaxes the constraint on physician dispensing as a self-enforcing institution, allowing physician-dispensing to remain the institutional equilibrium over a wider range of parameters. This extension can also help to illuminate why separation

reforms are often undertaken within a public discourse that excoriates providers for unethical and unprofessional behavior.

7.6 Incentives for drug discovery

Physician dispensing clearly yielded a stream of rents to physicians that reinforced the organizational structure protecting those rents, such as the political power of physician associations. It is less clear whether physician dispensing in traditional society could have been justified in terms of giving physicians more incentive to innovate with drugs than separate pharmacies or druggists, who often had comparative advantage in chemistry and testing.²¹ However, since pharmacies integrated into or owned by physicians (which is included in the definition of physician dispensing in my model) have also been common historically as well as in contemporary Taiwan and mainland China, it is worthwhile considering how the model can encompass incentives for drug discovery.

Consider an extension with an additional stage: after physician dispensing or separation is chosen, but before the physician sees a patient and chooses efforts, the physician may choose to invest in drug discovery. Patient benefit is increasing in investment in drug discovery, with diminishing returns. Then relational payments would represent combined payments for drug discovery investments and diagnostic quality effort. In societies with herbal medicine traditions, physician dispensing provided incentive for developing effective medications, and allowed doctors to cover the up-front investment with revenues from future dispensing above marginal cost, the traditional society's equivalent to patent rights for an innovation.

Following Grossman and Hart (1986) and Hart (1995), we assume that asset ownership – in this case, the right to dispense the *mifang* – conveys incentive for *ex ante* investments, and that the efficient ownership structure allocates ownership to the party whose investment is more valuable. In traditional society, innovation in herbal medicine was intimately linked to doctors' human capital and observation of clinical results. Integration of the three-stage process of drug discovery, patient diagnosis, and drug dispensing gave the physician incentive for effort in the first two stages, while preserving the cultural norm that patient payment to the physician was for the third stage, the dispensing of medication. To the extent that physician dispensing yielded a greater variety of effective medications than separation would have achieved, this drug discovery incentive expanded the circumstances over which physician dispensing remained a self-enforcing institution.

²¹I thank an anonymous reviewer for this point.

Physician investment in drug discovery was naturally limited by wealth constraints of individual physicians. Since biomedicine de-couples the drug discovery incentive from human capital in traditional Chinese medicine, and has expanded to require massive up-front investments, the drug discovery incentive of physician dispensing has been eclipsed by the development of a separate pharmaceutical industry even in China and India.

Figure 1.
MD-Dispensing (MDD) and Separation of Prescribing from Dispensing (SPD)
as Self-Enforcing Institutions

