Cite as

Acquatella, Angelique, Tianxu Chen, Randall P Ellis, and Taylor Watson (2024). "Provider Payment Systems and Incentives" in Stella R. Quah, (Ed.) International Encyclopedia of Public Health, Third Edition. doi.org/10.1016/B978-0-323-99967-0.00269-6

Stella R. Quah, Ed., International Encyclopedia of Public Health, Third Edition

Provider Payment Systems and Incentives

Angelique Acquatella, Tianxu Chen, Randall P Ellis, and Taylor Watson

Abstract

The financial structure within a provider payment system creates incentives that guide provider behavior, influencing the quantity, quality, and set of healthcare services provided. The key trade-off involves balancing cost-reduction incentives against quality of care and access to healthcare services. Recent trends in payment system reforms employ different financial structures across service classes (e.g. bundled, capitation, fee-for-service, value-based). This chapter covers the main financial structures used in provider payment systems, reviews the empirical literature describing the main incentive issues under each structure, and highlights the main information and agency frictions that explain the empirical evidence.

Keywords: capitation, Diagnosis Related Groups (DRG), fee-for-service, incentives, market organization, moral hazard, payment contracts, performance, provider payment, risk adjustment, selection

Contact Information:

Angelique Acquatella (Lead Author)
Toulouse School of Economics
1 Esplanade de l'Université
31080 Toulouse Cedex 06
Toulouse, France
angelique.acquatella@tse-fr.eu

Tianxu Chen
University of Connecticut Department of Economics
363-367 Fairfield Way
Storrs, Connecticut 06269 USA
tianxu.chen@uconn.edu

Randall P. Ellis
Boston University Department of Economics
270 Bay State Road

Boston, Massachusetts 02215 USA
ellisrp@bu.edu

Taylor Watson
Boston University Department of Economics
270 Bay State Road
Boston, Massachusetts 02215 USA
watson@bu.edu

1. Introduction

How do provider payment contracts affect the cost, quantity, and quality of care provided? This chapter presents a conceptual framework for provider payment systems that emphasizes the recent literature on provider payment contracts within healthcare organizations and markets. The classic decision problem facing healthcare providers frames provider treatment choices as trading off three things: patient benefit, treatment cost, and provider compensation. In this chapter, we emphasize that providers are compensated for their services through much more complex mechanisms than can be reflected in univariate functions of quantity. Instead, providers are paid through complex contracts that include salaries, fee-for-service, capitation, bundled payment, performance bonuses, or combinations of these. Provider payment contracts and treatment costs jointly determine the provider's profitability from administering a particular treatment, introducing a financial component to the provider's decision problem (Gaynor et al., 2023).

This overview is guided by recent empirical research that gives new insights into how providers make decisions – often imperfectly – and is benefiting from the availability of large digital healthcare databases and novel statistical methods. While provider decisions are made under uncertainty, medical training and skill help refine the accuracy of the diagnosis and treatment (Mullainathan and Obermeyer, 2022; Chandra, Cutler, and Song, 2011). Costs and outcomes are endogenous to provider's choices, investments, medical innovation, competition in input markets, and organizational structure (Dranove, 2011). It is the role of provider contracts to create incentives that foster efficient and equitable care for diverse patients who differ in their needs and preferences. Understanding the dimensions of these provider contracts is the focus of this chapter.

Early foundational work by Ellis and McGuire (1986) and Chalkley and Malcomson (1998) provide theoretical models of this trade-off, and emphasize the role of provider altruism; recent data and computing capabilities allow a deeper understanding of decision-making. Clemens and Gottlieb (2014) used data from the US Medicare system to explicitly link increases in payment rates (resulting from a policy reform) to increases in care provision by providers. Their rich data allows for the comparison of responses across treatment types and provider settings, further refining the role of altruism. In addition to the treatment and setting, Chen and Lakdawalla (2019) show that the characteristics of the patient also play a role in provider decisions, demonstrating that supply increases are often larger for providers treating wealthier patients versus poorer ones.

As this research has progressed, it has become increasingly able to inform health policy decisions as well. For example, in Gaynor et al. (2023), the authors use a structural model of provider choice with rich heterogeneity in altruism and costs to design optimal contracting policy for high-cost pharmaceuticals provided by physicians. They show how the existing simple

payment schemes could be adjusted to account for variations in provider altruism to better incentivize the optimal level of care provision, and avoid the current over-provision of treatment. This research and others like it can help policymakers better meet the complex challenges of payment design.

This overview of provider payment systems tries to recognize common market structures and incentives across the great diversity of provider payment systems used around the world. For example, in India, government physicians are often paid a salary while in Canada physicians are generally paid according to a government-regulated fee schedule. In the Netherlands office-based physicians receive capitated payments for much of their revenue, while in the US most doctors can negotiate fees for each service with each of many competing health plans. Similar variations are seen in payments to hospitals, which may be paid using fixed budgets, detailed fee schedules, or episode-based systems.

The focus of this chapter is on payments made to providers by social health insurance, private health insurance, other health plans, governments, or employers. While provider compensation and incentives are impacted by both supply and demand side payments, this chapter focuses on the role of supply-side payments on treatment incentives.

The rest of this chapter is organized as follows. Section 2 provides a taxonomic framework for comparing and contrasting payment systems, enabling comparisons along key system dimensions. Section 3 enumerates various incentive issues that arise, such as incentives to provide too much care (overprovision) or incentives to treat only healthy patients (skimming), while Section 4 discusses various structural and technological factors affecting provider incentives and performance. Section 5 provides a quick overview of how various countries approach provider payments and Section 6 concludes with a look at future trends and further reading.

2. Taxonomy of Provider Payment Contracts

To understand the many different kinds of provider payment approaches that abound in healthcare systems, it is helpful to establish a simple framework on which to place each system. The taxonomy in this section highlights how different payment approaches compare visually, as well as how the incentives of each may differ based on three key dimensions: Aggregation, Cost Basis, and Generosity.

Dimension 1: Aggregation

The “Aggregation” of a provider payment system refers to the bundle of provider services compensated under a single lump sum payment: with the most disaggregated being a payment for a single injection (fee-for-service), and the most aggregated being a single payment to a healthcare organization as compensation for all provided services (global budget). Under a system with “narrow” aggregation, doctors are paid explicitly for each procedure provided, and separate payments are made for each laboratory test or procedure done by other providers. Thus, a patient who sees a doctor for diagnosis and receives a test administered by a nurse would generate three separate payments - one for the diagnosis consultation, one for the test, and one for the test’s administration by the nurse.

In contrast, a "broad" payment system bundles multiple procedures together under one fee. The same patient described above might generate only one fee - that associated with a simple diagnosis and test. A system might broaden even further to include associated fees for laboratory services - or still broader to include the cost of specialists, forcing the primary doctor to internalize the cost of all physician services. The movement towards broad payment has come to be called "bundled" payment.

Similar issues of the aggregation of the provider payment system arise with hospital payments. Hospital DRG (Diagnostic-Related Groups) or per diem rates can be designed to narrowly cover only the room and board cost (US Medicare) or broadened to include laboratory and other support services. Even more broadly, they may include hospital physician services (France, Germany). In the US, certain DRG payments have even been broadened to include the cost of post-acute services provided after discharge. In other countries, hospitals are even given responsibility for certain kinds of primary care services (China).

In the empirical literature on bundling, the aggregation of payments is often discussed in terms of "bundling" of services. While new bundling innovations are being tried in many countries, much of the research on this topic comes from the US, where bundling is at the core of a long-developing series of aggregated payment approaches. Earlier studies of bundling focused on the US Medicare program, which adopted DRGs for hospital payments in the early 1980s. Coulam and Gaumer (1991) reviewed this early experience and concluded that DRGs had the predicted effect of lowering costs and intensity of input usage, consistent with the incentives of a change from lower to higher aggregation. The authors cautioned, however, that early studies were not able to inform sufficiently on the potential negative effects of this change in payment structure. Increased payment aggregation might lead to decreases in care quality, as hospitals seek to use less inputs in order to increase their profits relative to the fixed DRG rates. Another concern was changes in patient admission rates, known as patient selection (also elaborated in Section 3). Given unpriced patient heterogeneity, a fixed DRG payment makes it profitable for hospitals to prefer treating the healthiest, lowest-cost patients rather than more complicated cases, creating incentives for patient selection on who they admit.

Dimension 2: Cost Basis

The "Cost Basis" of a provider payment system refers to information used by the insurer to predict treatment costs incurred by the provider. The information structure of the provider contracting problem is asymmetric: true treatment costs are known only to the provider and not to the insurer that pays the provider for these costs. In light of asymmetric information, the insurer has to either impute costs using statistical predictions based on past information or rely on reported information from the provider. The first is subject to measurement error and the second to moral hazard. Moreover, payments can be based on the expected expenditure required to complete treatment ex-ante (before the service is provided), or on realized expenditure outlaid ex-post (after the service is provided). Some provider payment contracts combine the two approaches via outlier payment or mixed payment schemes that pay a blend of predicted and actual costs.

Ex-ante payments are commonly used for hospital DRGs. The DRG amount paid for a given diagnosis is often based on what the treatment expenditures have been for that diagnosis in previous episodes. The payer and hospital look at previous treatment costs of the same DRG and

calculate an average, agreeing to use that expected amount as the payment for future cases. The hospital then receives that DRG amount for each patient with the relevant diagnosis, regardless of how much it actually costs to treat each patient with the condition. The incentive problems arising from ex-ante payment hinge almost entirely on measurement error.

Consider, for example, the case of dialysis treatment for chronic kidney disease in the US Medicare program. These dialysis patients often received the drug Epogen to treat anemia, whose high cost made it Medicare's largest drug expense for Medicare Part B in 2010. To reduce this expenditure, Medicare switched from a fee-for-service approach to DRGs. As shown in Eliason et al. (2023), the fee-for-service payments were more than the Epogen costs for providers, creating a positive profit margin per dose which incentivized providers to administer more Epogen than necessary. Once the drug was bundled into the hospital's DRG and providers received a fixed payment regardless of whether Epogen was administered, rates of use fell, thus saving the program significant expenditure.

Ex-post payment systems are sometimes called cost reimbursement systems. Rather than the payer and hospital agreeing up-front to a payment amount on a given diagnosis, the provider treats the patient and then submits their realized costs and outlays to the payer for reimbursement. The incentive problems arising from ex-post payment arise from moral hazard (Cutler, 2010). If provider compensation depends on reported costs, providers have an incentive to report (inaccurately) large costs and keep the profit margin. Even if providers do not retain the profits, ex-post payments still fail to give providers the correct incentive to choose the most cost-effective treatments.

With these two dimensions in hand, it becomes easy to compare and contrast different provider payment systems quickly and visually. This can be seen in Figure 1: The y-axis is the first dimension of a system, its Aggregation, while the x-axis plots the Cost-Basis of payments. Countries can be plotted in this space to compare internationally or across systems, to understand how incentives existing in one system may be more or less important in other systems.

Dimension 3: Generosity

Not only is the unit of payment important, but also the overall generosity or level of payment. Empirical research shows that relative prices as well as their levels matter. If payment levels to providers are set too low, this reduces the incentive for quality for almost any payment system.

The generosity of payments can have both short-term and long-term effects, changing both the treatment decisions of individual providers as well as the market structure through its effect on entry, exit, and technology adoption decisions. With larger payments, there can be more providers because, for the same cost structure, revenues are larger. In a study for inpatient services, Ciliberto and Lindrooth (2007) estimated the probability of hospital exit as a function of reimbursement level throughout the late 1980s and 1990s. As expected, they found that more generous payments decreased the probability of hospital exit. Moreover, hospitals with a higher share of Medicare beneficiaries were more likely to exit the industry, but the effect dissipated in the late 1990s.

This framework for considering the dimensions of Aggregation, Cost Basis, and Generosity provides a foundation for examining payment system effects across different outcomes of interest, as well as across countries and their payment systems. The next section considers contracting and incentives related to specific outcomes of interest, while the following section examines experience across various countries and payment systems.

<Figure 1 near here>

Note: The x-axis measures the dimension of cost-basis of payments, and the y-axis is the dimension of aggregation in the system. This framework of the two dimensions could be used to compare across different provider payment systems. Source: Authors original interpretation.

3. Contracts and Incentive Issues

Building on this taxonomy of fundamental dimensions of provider payment systems, this section explores the efficiency and fairness implications of payment design on three outcomes of central interest: quantity of care provided, selection incentives for providers to prefer certain types of patients and efficient and fair treatment of patients who differ in their medical complexity. After providing examples of each problem to illustrate the issues involved we examine two currently popular methods that try to improve incentives: pay for performance and risk adjustment.

Incentive 1: Quantity

The first metric used to assess health system incentives is the quantity of care provided. As is clear from the previous discussion, the provision of quantity can be highly sensitive to payment incentives. In any provider system, when payments exceed marginal costs, providers will have an incentive to over-provide care (i.e., give too many services relative to the socially optimal amount). At the other extreme, some payment systems can lead to under-provision (where they give too few).

Over-provision of care is common in systems where provider payments are very disaggregated: in other words, when payments are provided for every treatment, test, or service individually, rather than in bundles or groups. The US commercial market is a prime example of this kind of system: providers usually bill insurance companies using a fee schedule that has specific amounts for every component of the care process. This is known as “fee-for-service” or FFS in the US. Given this structure, it is natural for providers to increase the amount of care they provide as much as possible, as this will directly increase their revenues.

Newhouse (2002) nicely summarizes the conventional economists’ and policymakers’ views that fee-for-service payment mechanisms create poor incentives for controlling quantity. Fee-for-service rewards physicians with more revenue for rendering more services, whether or not these services improve the health or well-being of the patient. Under fee-for-service reimbursement, services that have little or no value to the consumer may nonetheless be provided merely because they increase the provider's net income. Overprovision is especially likely to be a problem if the fee for a particular service is more than the incremental cost of that service to the provider. Carrin and Hanvoravongchai (2003) describe several instances in which fee-for-service payment systems caused the overprovision of services and thus caused a country’s healthcare costs

to rise. A classic example from 1987 was when general practitioners in Copenhagen began to receive fees for some services, and the provision of those services increased significantly.

The opposite quantity incentive problem is under-provision. This is especially likely in salaried or budgeted facility payment systems, where provider payments are highly aggregated to the level of a single payment per physician or hospital. Because revenues are fixed regardless of the volume of care rendered, physicians and hospitals have a financial incentive to inappropriately keep costs and costly effort at levels that are too low. Fixed revenues can also motivate physicians and hospitals to preferentially treat low-risk patients, to try to avoid high-cost patients, to make too many referrals, to minimize the number or intensity of services provided, and in general to underprovide quality.

The one countervailing force is the possible threat of losing one's job or shutting down a facility, but countries that rely on salaried or fixed facility budget payments rarely do this. Peer review and payer monitoring may also be effective at promoting quality. Payments based solely on provider characteristics do not tend to reward care coordination across providers. Salaried or fixed hospital budget systems often rely on provider altruism to ensure that appropriate and high-quality services are provided. However, only if providers realize immediate cost-savings from preventive care will a payment system encourage providers to engage in such care.

Under-provision can also occur in a disaggregated FFS system if fees are too low relative to provider costs. Many costly services are under-compensated with fee-based reimbursement. End-of-life planning, patient lifestyle counseling, or smoking, drug, and alcohol consumption typically require unobservable provider effort and rarely are compensated in existing payment systems. Competition and consumer information, also affect incentives of fee-based payments. It is important to assess each system in its own context to understand the incentives it may create.

Incentive 2: Patient Selection

If the quantity of treatment is the intensive margin of care, the decision of whether or not to treat a patient can be considered the extensive margin - this is where the second incentive comes into play. Payment systems often create incentives for providers to try to attract certain types of patients (e.g., healthier ones) and avoid treating or admitting others (e.g., sicker ones). This is called patient selection, also sometimes referred to as "cream skimming".

To see how this incentive varies with payment design, let us again consider the two extremes of payment aggregation as in quantity. For highly aggregated payment systems, the same pressure that can lead to underprovision of care will also lead providers to select healthier patients and avoid sicker ones. Given a fixed budget, the kind of patients that are treated is a second margin on which providers can minimize their total expenditure and thus maximize their profits.

On the other side, one might expect that very disaggregated payments might lead to the opposite kind of selection (preferring sicker patients who require a higher volume of treatment), but the reality is a bit more nuanced. For disaggregated payments, the other dimensions of payment can play a greater role in selection incentives. For example, if payments are highly generous, then one could imagine this kind of opposite selection: generous payments will more than compensate

for the high costs of treating sick patients, making them potentially highly profitable. On the other hand, if payments are not generous, and margins are very thin, then the incentive to treat sicker patients over healthier ones will decrease. Considering the interplay between different payment dimensions is crucial for understanding how providers will behave under different payment systems.

The cost basis dimension is also important in selection incentives. Again, considering the extremes, if payments are based entirely on ex-post realized costs, then the incentive to select patients should disappear. Providers know that they will be reimbursed based on what it costs to treat the patient, and so they face no risk of losses if a patient is more expensive to treat. The same cannot be said for payment systems where costs are calculated ex-ante. In such situations, providers will have a strong incentive to select patients with less risky outlooks.

On the other hand, a very generous payment system based on realized ex-post costs will incentivize competition among providers for all patients, even higher-cost ones, to gain from their margins as in the quantity scenario. The riskiness of a patient is mitigated when their realized treatment costs are the basis for payments, as that incorporates their outcomes directly.

The impacts of patient selection incentives on provider behaviors are an active area of research. Considering the impacts on overall healthcare efficiency, Layton et al. (2017) examine the distortions caused by this incentive type in prices and benefit designs. Using robust simulations, they consider how additional payment features may mitigate or exacerbate these kinds of incentives.

Incentive 3: Complexity

The incentive effects of complexity are subtler than those of quantity or patient selection, making them more difficult to identify in outcomes data; however, the implications in system costs and efficiency are still highly significant. The complexity of a payment system mostly refers to the degree of differentiation made between different treatments, procedures, or diagnoses.

Consider, for example, a hospital system using DRG payments. A highly complex version of such a system would have a specific DRG for each of numerous different diagnoses, even specifying how the amounts should change with finer details of the disease within a single diagnosis. A less complex system would have fewer DRGs, such that patients with similar diagnoses and slightly different presentations of the disease would still result in the same payment to the hospital. Note the distinction from the aggregation dimension here - while the aggregation addresses how payments for individual components of care are combined, the complexity refers to how many of those bundled payments there are / how complex their derivation is.

The incentive distortions caused by low complexity are relatively similar to those discussed above regarding patient selection. A hospital that is paid using a relatively sparse or unrefined schedule of DRGs would prefer to treat patients within each DRG who are healthiest, as patients who are sick and require further care are not differentially billable. By adding more detailed DRGs or by otherwise expanding the payment system to account for the complexity of treatment scenarios, systems can mitigate this particular kind of patient selection incentive.

However, with more classifications, there are greater monitoring problems and more incentives to game the system. Providers may upcode, or classify patients into a higher payment category, to receive a larger compensation. In the literature, this is often referred to as "DRG creep" when referring to hospitals and "code inflation" when referring to providers paid based on diagnosis codes. Finer systems can be more difficult to monitor since more patients will be near the margin where upcoding can make a difference, on the other hand, coarser systems reward upcoding with greater increases for those cases where upcoding is possible. Selection incentives within a coarse system may also be a concern. In summary, increasing complexity will tend to mitigate patient selection and increase the fairness of payments to providers, but if it goes too far it can create new negative incentives as well.

How complexity affects incentives is particularly apparent when payments go from physician fee schedules to DRG payments. This was illustrated in Korea when inpatient physicians switched from a relatively complex and narrowly defined payment formula to a simple and broad system. Until 1997, under the national health insurance program, a fee-for-service (FFS) payment system had been in place for approximately twenty years. The FFS payment system has led to a high volume of low-intensity health care, characterized by frequent but short visits and hospital stays. Under the FFS system, inpatient physicians had financial incentives to choose treatments with greater profits, and hence increase the use of medical supplies and pharmaceuticals and avoid hospitalizations. In 1997, in hopes of controlling quantity and selection, a DRG pilot program was introduced for voluntarily participating providers. By its third year, the pilot program covered nine disease categories with twenty-five DRG codes which depended on the severity and age of the patient. The program began with disease groups that had low expenditure variation, little disagreement among providers on treatment methods, a low degree of uncertainty about treatment outcomes, a high frequency of utilization, and a lower possibility of DRG creep. Within each of these DRGs, there were three types of patients: normal case, outlier below the lower limit, and outlier above the upper limit. The pilot program succeeded in lowering expenditure on medical care, reducing length of stay, and reducing the use of antibiotics. Early evidence suggests that the program did not harm the quality of care as measured by complications and re-operations (Kwon 2003).

The effects of incentives were also demonstrated when the National Health Insurance was established in Taiwan in 1995. At this time, Taiwan switched from paying office-based doctors a salary to paying them according to a FFS schedule. A relatively simple FFS was used based on a national fee schedule. The switch to FFS was accompanied by an increase in the volume of services with a shortened average visit length. According to one source, this led to misdiagnosis, improper treatment, and delays; the government responded by changing the fee structure to try to limit the number of patients each provider can treat during a given day (Cheng 2003).

The US Medicare system compensates providers according to a more complicated payment system (i.e., has more payment categories) than the systems in Korea and Taiwan. Under the US Medicare system, hospitals receive a prospectively determined price depending upon the patient's DRG. Initially using only 470 DRGs, in FY 2022 the US Medicare system had 767 DRGs for inpatient admissions. As mentioned above, providers have a financial incentive to select the most profitable patients within one DRG group. Even with numerous DRGs, Dranove (1987) discusses the large differences in costs that occur within one DRG in the US Medicare system. In three Chicago area hospitals Dranove (1987) finds that one-sixth of the DRGs have a standard deviation

that exceeds the mean. Large differences in the cost of treating patients classified in one DRG encourage hospitals to prefer treating less costly patients within a DRG. Thus, Dranove believes that a finer payment system is needed to avoid selection. However, one might worry that a finer US Medicare system would create more incentives to upcode patients. The classic study of this phenomenon in the US by Carter et al. (1990) found that upcoding or DRG creep accounted for less than one-third of the change in Medicare's Case Mix Index between 1986 and 1987.

The complexity of the Medicare payment system results in about 5% of physician revenue being lost to billing issues and details (Dunn et al., 2018), yet this is less than a third of the loss physicians take from Medicaid, nearly 18%. This loss in revenue has led to many physicians refusing to accept Medicaid patients in states with more complex billing procedures.

Australia also uses a complex payment system to fund its public hospitals; the 2022 Australian Refined Diagnostic Related Groups (AR-DRG) has 801 categories in its Release 11.0, a system used by several other countries as the foundation of their DRG systems. The German G-DRG system was carefully phased in over 5 years from 2005-2010 to ultimately have 1200 G-DRGs (Busse et al., 2013). In 2020 Germany unbundled nursing care from the facility DRGs to enable its trajectory to differ from that of the remainder of facility charges. One 2022 University of Hamburg study found that German hospitalizations had increased but not lengths of stay. Because the payment systems in Australia and Germany are relatively fine, there are fewer opportunities for providers to select low-cost patients. On the other hand, there are potentially more opportunities for providers to upcode patients to increase facility payments, although, unlike the US system, this changes the allocation of total funds, not the sum totals.

Pay for Performance

To improve specific objectives, regulators, and administrators are increasingly using pay for performance (P4P) systems for provider payment to overcome the problem whereby doctors who provide inadequate care receive the same compensation as doctors who provide excellent care. For example, doctors may receive bonuses based on performance targets, such as high immunization rates, or low surgical complication rates. Other payment systems reward providers according to how well they perform relative to their peers on various cost or quality measures. Rich research on P4P has been conducted in the US, (Mendelson et al. 2017) Europe (Emmert et al. 2012), and low and middle countries. (Singh et al., 2021). In the US, the fundamentals of P4P have been transformed into a broader concept of value-based payment, discussed below.

In the US, early P4P systems were commonly used by health maintenance organizations (Rosenthal et al. 2006; Dudley and Rosenthal 2006), while later variants have been applied in a wide variety of other settings, including hospitals and primary care payment, accountable care organizations (ACOs), and new incentive systems called value-based payment. In the US and UK, primary care physicians as individuals or practices, are grouped into risk pools that share financial rewards and penalties; the size of the risk pool will influence a physician's responsiveness to incentives. Larger pools may reduce risk and make the detection of improvements easier but will also tend to dilute the incentives to perform well.

For some medical conditions, such as diabetes and asthma, there are easily recognizable quality measures available for gauging provider performance. For example, if bonuses are allocated to providers who track diabetics' blood sugar levels, then doctors will be more likely to

have diabetic patient's blood sugar monitored. Similarly, asthmatic patients will receive high-quality treatment if physicians are given bonuses for prescribing the correct asthma medication. Performance-based payment can encourage providers to give high-quality care to patients with these two conditions.

However, there are challenges to performance-based payment systems. P4P programs may increase selection if physician performance is based on claims data that does not fully reflect a patient's risk factors. If so, bonuses create incentives for doctors to drop risky patients within a payment category. Evidence of this behavior was found by Oxholm et al., (2021), who used a controlled laboratory experiment with Danish medical students to see how different kinds of patients fared under P4P. They found that patients who had high potential to achieve the performance targets received more care, but that patients with little potential may receive less care. Similarly, many doctors oppose performance-based payment systems because their payments rely too heavily on their patients' risk factors as well as their patients' actions (e.g., whether their patients take their prescriptions or return for a follow-up appointment). Another concern is that only a relatively small number of performance measures are typically used. Incentives that do well on these measures may not carry over to other actions that are not measured or rewarded.

One widely cited study examined doctors who were offered bonuses if they complied with basic public health guidelines, including guidelines on preventative care. In 2003, California doctors were evaluated according to levels of breast cancer screening, cervical cancer screening, and hemoglobin testing. The top-rated doctors split a bonus pool of $3.4 million. Compared to doctors in Oregon and Washington who were not offered bonuses, California doctors offered more cervical cancer screening. Although the quality of care increased in all three areas, only for cervical cancer screening was the improvement greater in California than in Oregon and Washington; this may have been because the financial rewards for quality were too low or because substantial quality improvements take time. Physician groups whose performance was initially the lowest improved the most, whereas physician groups who had previously achieved the targeted level of performance improved the least (Rosenthal et al. 2005).

One of the most prominent uses of P4P is in Accountable Care Organizations (ACOs). ACOs are an organizational tool used by Medicare to try to decrease costs while maintaining high levels of quality of care. Traditionally, Medicare operated under a fee-for-service payment system and doctors operated alone or in small groups with little or no integration. Besides the issue of overprovision due to fee-for-service, the lack of integration added the problem of difficult coordination, leading to duplication of treatments, with different providers prescribing treatments that had already been done by others. ACOs were created to overcome this problem. An ACO is an institution comprising doctors who work together to treat patients. Integrating care between different specialties is expected to increase professional ties, overcome the duplication of treatment problems, and reduce costs. Providers in ACOs also agree to provide high-quality care and their payments are tied to performance measures. The payment model in ACOs is mainly based on global budgets subject to performance indicators, while incentives for cost control are given by allowing them to share their savings.

The first few years of ACO experiences show some cost reductions. McWilliams et al. (2015) show a reduction of quarterly spending per beneficiary of around 1.2% among certain early

adopters of ACOs compared to those enrolled in traditional fee-for-service. Nyweide et al. (2015) find similar results with no evidence of harm due to cost savings. Both studies considered ACO programs that were voluntary for providers. Given that the most willing providers are likely to also be lower cost these results should be used with care, since a more widespread program would likely result in lower savings.

In the US Medicare Merit-based Incentive Payment System (MIPS) started in 2017, physicians providing Medicare Part B services were scored 100-point cost and quality metrics (e.g., interoperability, costs, improvements). Depending on these scores, physicians could receive substantial bonus payments. Despite the simple design, McWilliams (2022) found that the scores were not predictive of clinical performance and outcomes.

While the US is a leader in P4P in some areas, P4P measures have also been introduced in other countries, notably for primary care in the UK starting in 1990. A variety of studies found modest effectiveness of such payments on immunizations (Kouides et al., 1998; Ritchie et al. 1992), and larger impacts on chronic care quality indicators where bonuses could increase practice income by up to 25% (Doran et al. 2006).

P4P schemes are increasingly used in low and middle-income countries, though much like in high-income countries there are mixed results on their efficacy. In a literature review by Kovacs et al., (2020), the authors note the significant heterogeneity in how the programs are designed as driving the variation in successful outcomes. A later review by Singh et al., (2021) goes further, to discuss which underlying mechanisms and contexts are likely to produce more consistent positive outcomes in these countries.

The evidence continues to be mixed on the effectiveness of P4P on quality and quantity of care. The theoretical and empirical research literature has not kept up with recent innovations, and many innovations are still being implemented that have not yet been validated or the incentives modeled. Unfortunately, it does not seem that the lessons of early research are being effectively built upon in more recent efforts: Zaresani and Scott (2021) found in a meta-regression analysis across P4P studies that the effectiveness of P4P schemes is evolving slowly, and that new design and evaluation approaches face similar challenges to those tried in the past.

Since 2000 in the US many payers have switched over to emphasizing value-based purchasing (VBP), as reviewed in Pandey et al. (2023) in which contracts reward providers for improving health care quality while reducing spending. This can be seen as a variant of P4P in which spending is a key part of the performance, cost reduction incentives are higher powered, and larger fractions of revenue are at risk from the point of view of providers and payers. VBP contracts often have nonfinancial supports that include data reports, technical assistance, leadership training, and much more. Pandey et al. summarize the key finding of their meta-analysis of 24 VBP programs with sufficient quality to assess as “VBP is a durable policy trend despite mixed results and somewhat halting progress toward its goals”. It seems too early to assess the long-term success of this new approach.

Risk Adjustment

Risk-adjusted payments are used to mitigate (to some extent) the drawbacks of performance-based systems. By rewarding providers according to the characteristics of the patient – how risky they are – they would less likely to dump risky patients as it would not harm their chances of achieving their performance indicators.

For example, if the indicator is low surgical complication rates, then doctors would want to avoid operating on patients who are more likely to have complications. Risk-adjustment handles this issue by giving risky patients a higher weight in the calculation of this indicator, rewarding providers that treat riskier patients.

Risk adjustment tools can be used in a variety of healthcare settings, even if not directly used to pay providers. In countries such as the US, it is used to level the risk of insurance companies that provide health plans for Medicare beneficiaries, for example, to avoid selection occurring from plan characteristics. In countries where there is a National Health System, risk adjustment can be used to allocate resources across regions, where selection can still occur due to differences in condition prevalence, age composition, or socioeconomic conditions.

The amount of information available also affects the extent to which risk adjustment can be used. While the initial methods to calculate risk scores used mainly age, gender and disability status as cost predictors, including diagnosis information greatly increased the power of cost predictions. For example, Ash et al. (2017) studied the addition of social determinants of health (SDH) to diagnosis-based payment formulas in Medicaid. They found that adding SDH could eliminate or significantly reduce underpayments for vulnerable groups. For a comprehensive description of the evolution of risk adjustment utilization and different methods see Cid et al. (2016).

Unfortunately, adding more variables to a risk adjustment score calculation is not without cost, as each new variable potentially presents an opportunity for upcoding or other issues. Balancing the performance of risk adjustment with a smaller number of variables was explored by McGuire et al. (2021), who used a combination of constrained regression and machine learning to produce alternative formulas. They can achieve fit performance as good or better than calculators used in the US Marketplace using fewer variables, offering potentially efficiency-enhancing simplifications.

At the intersection of this topic and the previous discussion of pay-for-performance, using risk adjustment to calculate payments while taking into account performance measures has been studied by Ash and Ellis (2012) for primary care doctors. In their study, they calculate risk-adjusted costs for primary care services that doctors should provide, rather than all the services that they currently provide, thus rewarding doctors who remain close to what should be done. The definition of the services that primary care doctors should be providing takes into account other types of care that might signal the need for primary care. Their approach rewards doctors who achieve outcomes better than expected. Moreover, the methodology is adaptive in the sense that results can be re-estimated when new data is available, thus creating a new benchmark for the expected outcomes. Should the average outcomes increase from one year to the next, so does the new benchmark, thus creating incentives for doctors to keep improving and not accommodate the current benchmark.

4. Structural and technological factors that affect incentives and performance

Market Structure & Competition

The way payments are defined along the previously mentioned dimensions will affect the incentives for providers to enter or exit the market, as well as incentives for providers to merge or divide. If payment systems change the number or size of providers over time that exist in a given market, the resulting changes in the degree of competition can change market outcomes.

In the US, market structure and degree of competition have changed substantially since the late 1980s. In general, the market has consolidated and experienced a decrease in competition. Gaynor and Town (2011) show that the mean Herfindahl–Hirschman Index – a measure of market concentration defined by summing the square of market shares – for US hospitals increased by almost 40% between 1987 and 2006, primarily due to mergers and acquisitions during the 1990s. The degree of competition also seems to have decreased for physicians, according to Liebhaber and Grossman (2007). They show that the proportion of physicians working alone or in groups of up to five has decreased while the proportion working in groups of 6 or more has increased. The continuation of this trend was confirmed by Muhlestein and Smith (2016), who used data from 2013-2015 to show that the share of physicians working in groups 100-499 has overtaken the share of physicians working alone or in pairs.

Innovation and Technology

Provider payment can also influence the nature of technology and innovation. If doctors and hospitals are rewarded for adopting new expensive technologies with higher revenues, this may spur the development and adoption of such new technologies, whereas if the revenues remain constant, they may underinvest. Because insurance with low patient cost sharing may make patients overeager to choose expensive rather than less expensive technologies, supply-side payment incentives may be useful to offset this tendency. *Ex Post* (cost-based) and Narrow payment systems will potentially incur overinvestment in cost-increasing investments relative to *ex ante*, Broad systems.

Mullainathan and Obermeyer (2022) provide a useful modern framework for thinking about how technology, patient beliefs, and diagnostic uncertainty interact with payment system incentives to affect physician decision-making. Consider a patient arriving in an emergency room with symptoms that might indicate a heart attack - chest pain and nausea. The physician must choose whether or not to administer the test to confirm it is a real heart attack, weighing the costs and benefits: while failure to identify a real heart attack could have grave consequences for the patient, the test is costly and invasive, and the same symptoms which indicate heart attack may also be caused by relatively harmless conditions. The physician's choice is further complicated by their financial incentives: in some systems, they may receive payment for every test they run (incentivizing more testing), while in others they may have a fixed budget for running such tests each quarter (incentivizing less testing). How might the patient react to knowing the role of the personal financial incentives of their physician in their treatment choice?

Mullainathan and Obermeyer (2022) are pathbreaking in that they use machine-learning approaches to compare real physician testing choices to the data-predicted risk of those same patients. They find that when systems use financial incentives to avoid over-testing low-risk

patients, physicians often end up under-testing high-risk patients as well, resulting in missed heart attacks and greater future costs. The incorporation of machine-learning techniques like this one will likely continue to affect provider incentives and shape the landscape for years to come. This is especially true given the massive amounts of data that are available in the healthcare space.

Prevention versus curative care

Many payment systems over-provide treatment relative to curative care. If providers get more money for treating sicker than healthier patients, then they may have less incentive to provide optimal prevention effort. In the US, the underpayment for primary care is often blamed for the inadequate attention given to prevention relative to specialty treatment for chronic illnesses. Undercompensation of primary care providers for providing COVID-19 vaccines early on in the pandemic in several European countries has also been blamed for delays in achieving widespread vaccination rates.

Government regulations

Most discussion of payment systems has focused on parameters describing the transfer of funds to providers by payers. Other government regulations can also interact with payment system features. For example, if primary care doctors or hospitals are regulated to accept any patient requesting care, or ambulances are required to take patients only to certain specific hospitals based on geography this can also have direct consequences on costs, quantity and quality. It can also affect the other dynamic considerations discussed here.

5. Global Outlook

As is clear from the number of different examples in the preceding sections, there is significant heterogeneity in provider payment approaches around the world. This section provides a convenient reference in Table 1 to summarize the dominant payment method used in a selection of countries and markets. Additional references with helpful international comparisons are available from Ellis et al. (2014) and McGuire and van Kleef (2018).

Table 1: Provider Payment Systems Global Outlook

Country	**Inpatient**	**Outpatient**	**Primary Care**	**Hospital**	**Office Physician**	**Notes**
North America						
Canada	Global Budgets	FFS	FFS	Global Budgets	FFS	Some provinces consider activity-based hospital funding
USA - Private	DRG	FFS	DRG-FFS Hybrid	Public - DRG	FFS	These are modal forms among great diversity

USA - Medicaid	FFS	FFS	FFS	FFS	FFS or Capitation	Varies by state
USA - Medicare	DRG	DRG (Prospective Payment System)	FFS	DRG	FFS or Capitation	Different for Medicare Advantage capitation
Europe						
Austria	DRG		FFS + P4P	DRG	FFS	
Czech Republic		FFS	FFS	DRGs		
Denmark	Fixed Budget + DRGs		Capitation + FFS	Fixed Budget + DRGs	FFS	
France	Global Budget + DRG / T2A	FFS + Salary	FFS + capitation in some chronic cases	Global Budget + DRG / T2A	FFS + Salary	
Germany	DRG	FFS	FFS Hybrid (with cap by specialty)	DRG	FFS Hybrid (with cap by specialty)	Standardized features of statutory plans
Italy	DRG	FFS	Capitation + FFS	DRGs; however locally run often use global budgets	Salary	Prospective
Nether-lands	Budget (case mix)	Salary+ private FFS	Capitation + copayments	DRG	FFS Hybrid with Capitation	Cost-Based
Norway	DRG	DRG	Capitation + FFS	Global Budget	FFS	
Spain	Global Budget + DRG	Global Budget	Global Budget + P4P	Global Budget	Salary + P4P + Capitation	
Sweden		Global budgets +	Capitation + FFS	Global budgets + DRGs + P4P	Salaries	

		DRGs with some volume ceilings				
UK	DRG	Salaries + FFS	A mix of capitation, FFS, and P4P	Budget + PPF	Capitation	
Asia						
China	FFS	FFS	Salaries + FFS	FFS/Public - Subsidies	FFS	Experimentation with DRG, capitation, and global budgets growing for inpatient care
India	Public: Salary	Public: Salary/Private: FFS	Public: Salary	Public: Annual Budgets/Private: FFS	Public: Salary	
Israel	Medical: Per Diem/Procedure: PRG	FFS or PRG	Salaries and capitation, vary by health plan		Capitation + FFS up to ceiling	
Japan			FFS with some per-caes exceptions in the national schedule	Choose either FFS or DPC	FFS with some per-case exceptions in the national schedule	DPC diagnosis-procedure combination uses case-mix per-diem payment with FFS for more expensive components
Korea	DRG	FFS		FFS		
Singapore	DRG	FFS	FFS	Overall amount ceiling	FFS	
Taiwan	FFS & DRG	FFS & DRG	Global Budget divided across six regions	FFS & DRG	FFS	Uniform national fee schedule for hospitals/P4P across 12 target diseases

Other						
Brazil	MoH pays localities based on DRGs, who in turn reimburse hospitals FFS	FFS	Salaries	MoH pays DRGs to localities that pay hospitals FFS & volume-based complex service payments	FFS	
New Zealand	Case-mix with some P4P		Capitation + minor P4P	Budgets	Public: Salaries / Private: FFS	

Note: This table is the authors' original interpretation of this information. This information underlying the table is primarily sourced from The Commonwealth Fund's series of International Health Care System Profiles, available here: https://www.commonwealthfund.org/international-health-policy-center/countries

6. Trends for the Future & Recommendations for Future Reading

This chapter has shown that provider payment systems vary in many dimensions, all of which should be considered when evaluating reforms or studies. Incentives created by provider payment vary along key dimensions of provider payment systems, but many structural and technological factors play a role. As health systems around the world continue to seek to mitigate negative incentives and enhance positive ones, many are expected to shift towards systems that better use technology and data, particularly in more complicated contracting schemes such as P4P and beyond.

In addition to the many works cited in this chapter, there are further readings that offer supplemental insights into the topic of provider payments and incentives. For those seeking a broad view of how different nations have structured their systems, the European Observatory on Health Systems and Policies provides helpful reviews in a standardized format (WHO, 2023). Similar cross-country comparisons are available from Carrin and Hanvoravongchai (2003), who focus on high-income countries, and Jegers et al., (2002), who provide an alternative typology for organizing payment systems. Finally, a recent publication from Rebitzer and Rebitzer (2023) explores the many challenges facing cost-containment strategy in healthcare, including many of the difficult incentives discussed in this chapter.

Within this publication, the following chapters may also provide insight on these topics:

- International Classification Systems for Health
- Economic Models of Hospital Behaviour
- National Health Systems: A Historical OverView
- Health-Care Delivery Systems

- Comparative health systems
- Demand for health care

References

Ash, A., and Ellis, R. (2012). Risk-adjusted payment and performance assessment for primary care. *Medical Care* 50(8), 643.

Ash, A.S., Mick, E.O., Ellis, R.P., Kiefe, C.I., Allison, J.J., and Clark, M.A. (2017). Social Determinants of Health in Managed Care Payment Formulas. *JAMA Internal Medicine* 177(10):1424-1430.

Busse, R., Geissler, A., Aaviksoo, A., Cots, F., Häkkinen, U., Kobel, C., Mateus, C., Or, Z., O'Reilly, J., Serdén, L. and Street, A., (2013). Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals?. *BMJ*, *346*.

Carrin, G., and Hanvoravongchai, P. (2003). Provider payments and patient charges as policy tools for cost containment: How successful are they in high-income countries? *Human Resources for Health* 1(6), 1-20.

Carter, G., Newhouse, J., and Relles, D. (1990). How much change in the case mix index is DRG creep? *Journal of Health Economics* 9, 411-428.

Chalkley, M., and Malcomson, J. (1998). Contracting for health services when patient demand does not reflect quality. *Journal of Health Economics* 17(1): 1-19.

Chandra, A., Cutler, D., and Song, Z. (2011), Who Ordered That? The Economics of Treatment Choices in Medical Care. In Pauly, M., McGuire, T., and Barros, P. (eds.) *Handbook of Health Economics,* Vol 2, pp 397-432, Elsevier.

Chen, A. and Lakdawalla D.N. (2019). Healing the poor: The influence of patient socioeconomic status on physician supply responses. *Journal of Health Economics* 64: 43-54

Cheng, T. (2003). Taiwan's new National Health Insurance program: genesis and experience so far. *Health Affairs* 22, 61-76.

Cid, C., Ellis, R., Vargas, V., Wasem, J., and Prieto, L. (2016). Global Risk-Adjusted Payment Models. In Scheffler, R. (eds.) *Handbook of Global Health Economics and Public Policy*, Vol 1, World Scientific.

Ciliberto, F., and Lindrooth, R. C. (2007). Exit from the hospital industry. *Economic Inquiry* 45(1), 71-81.

Clemens, J., and Gottlieb, J. (2014). Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health? *American Economic Review* 104 (4): 1320-49.

Coulam, R., and Gaumer, G. (1991). Medicare's prospective payment system: A critical appraisal. *Health Care Finance Review* (Suppl): 45-77.

Cutler, D. (2010). Where Are the Health Care Entrepreneurs? *Issues in Science and Technology* 27 (1): 49-56.

Doran, T., Fullwood, C., and Gravelle, H. (2006). Pay-for-performance programs in family practices in the United Kingdom. *The New England Journal of Medicine* 355, 375-384.

Dranove, D. (1987). Rate-setting by diagnosis related groups and hospital specialization. *The RAND Journal of Economics* 18(3): 417-427.

Dranove, D. (2011). Health care markets, regulators, and certifiers. In Pauly, M., McGuire, T., and Barros, P. (eds.) *Handbook of Health Economics*, Vol 2, pp 639-690, Elsevier.

Dudley, R.A., and Rosenthal M.B. (2006) *Pay for Performance: A Decision Guide for Purchasers*. Report prepared for the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services Contract No./Assignment No. 290-01-0001/298 AHRQ Publication No. 06-0047, April.

Dunn, A., Gottlieb, J., and Shapiro, A. (2018). The complexity of billing and paying for physician care. *Health Affairs* 37(4): 619-26.

Eichler, R., Auxila, P., and Pollock, J. (2001). Promoting preventive health care: Paying for performance in Haiti. In Brook, P.J., and Smith, S.M. (eds.). *Contracting for public services: Output-based aid and its applications*, pp 65-71, World Bank Publications.

Eliason, P., Heebsh, B., League, R., McDevitt, R., and Roberts, J. (2023). The Effect of Bundled Payments on Provider Behavior and Patient Outcomes: Evidence from the Dialysis Industry. Unpublished.

Ellis, R.P., and McGuire, T.G. (1986). Provider behavior under prospective reimbursement: cost sharing and supply. *Journal of Health Economics* 5, 129-152.

Ellis, R.P., Chen, T. and Luscombe, C.E. (2014). Comparisons of Health Insurance Systems in Developed Countries, In *Encyclopedia of Health Economics*, Anthony J. Culyer (eds.), Elsevier.

Emmert, Martin, Frank Eijkenaar, Heike Kemter, Adelheid Susanne Esslinger, and Oliver Schöffski. (2012). "Economic evaluation of pay-for-performance in health care: a systematic review." *The European Journal of Health Economics* 13: 755-767

Gaynor, M., Mehta, N., and Richards-Shubik, S. (2023). Optimal Contracting with Altruistic Agents: Medicare Payments for Dialysis Drugs. *American Economic Review* 113 (6): 1530-71.

Gaynor, M., Town, R. (2011). Competition in Health Care Markets. In Pauly, M., McGuire, T., and Barros, P. (eds.) *Handbook of Health Economics*, Vol 2, Elsevier.

Jegers M., Kesteloot, K., De Graeve, D., and Gilles, W. (2002). A typology for provider payment systems in health care. *Health Policy* 60 255–273.

Kouides, R.W., Bennett, N.M., Lewis, B. *et al* (1998). Performance-based physician reimbursement and influenza immunization rates in the elderly. *American Journal of Preventative Medicine* 14, 89-95.

Kovacs, R.J., Powell-Jackson, T., Kristensen, S.R. *et al.* (2020). How are pay-for-performance schemes in healthcare designed in low- and middle-income countries? Typology and systematic literature review. *BMC Health Services Research* 20, 291.

Kwon, S. (2003). Payment system reform for health care providers in Korea. *Health Policy and Planning* 18(1), 84-92.

Layton, TJ., Ellis, RP., McGuire, TG., and van Kleef, R. (2017). Measuring efficiency of health plan payment systems in managed competition health insurance markets. *Journal of Health Economics* 56: 237-255.

Liebhaber, A., and Grossman, J. M. (2007). Physicians moving to mid-sized, single- specialty practices. Tracking report/Center for Studying Health System Change (18): 1-5.

McGuire, T. G., and van Kleef, R. C. (2018). Regulated competition in health insurance markets: Paradigms and ongoing issues. In T. G. McGuire and R. C. van Kleef (eds.), *Risk adjustment, risk sharing and premium regulation in health insurance markets: Theory and practice,* pp. 55–104, Elsevier.

McGuire, TG., Zink, AL., and Rose S. (2021). Improving the Performance of Risk Adjustment Systems: Constrained Regressions, Reinsurance, and Variable Selection, *American Journal of Health Economics* 7(44): 497–521.

McWilliams J.M. (2022). Pay for Performance: When Slogans Overtake Science in Health Policy. *JAMA* 328(21): 2114–2116.

McWilliams, J. M., Chernew, M. E., Landon, B. E., and Schwartz, A. L. (2015). Performance Differences in Year 1 of Pioneer Accountable Care Organizations. *New England Journal of Medicine* 372: 1927-1936.

Mendelson, A., Kondo, K., Damberg, C., Low, A., Motúapuaka, M., Freeman, M., O'neil, M., Relevo, R., and Kansagara, D. (2017). The effects of pay-for-performance programs on health, health care use, and processes of care: a systematic review. *Annals of Internal Medicine* 166(5): 341-353.

Muhlestein D.B., and Smith N.J. (2016). Physician consolidation: Rapid movement from small to large group practices, 2013-15. *Health Affairs* 35: 1638-1642.

Mullainathan, S. and Ziad Obermeyer. (2022). Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care. *The Quarterly Journal of Economics* 137(2): 679–727.

Newhouse, J.P. (2002). *Pricing the Priceless: A Health Care Conundrum*. MIT Press.

Nyweide, D. J., Lee, W., Cuerdon, T. T. *et al* (2015). Association of Pioneer accountable care organizations vs traditional Medicare fee for service with spending, utilization, and patient experience. *The Journal of the American Medical Association* 313(21): 2152-2161.

Oxholm, A., Di Guida, S., and Gyrd-Hansen, D. (2021). Allocation of health care under pay for performance: Winners and losers. *Social Science & Medicine* 278.

Pandey, A., Eastman, D., Hsu, H., Kerrissey, M., Rosenthal, M., and Alyna Chien. (2023). Value-Based Purchasing Design And Effect: A Systematic Review And Analysis. *Health Affairs* 42(6): 813-821.

Rebitzer, J., and Rebitzer, R. (2023) *Why Not Better and Cheaper? Healthcare and Innovation.* Oxford University Press.

Ritchie, L.D., Bisset, A.F., Russell, D., Leslie, V., and Thomson, I. (1992) Primary preschool immunization in Grampian: Progress and the 1990 contract. *British Medical Journal* 304(6830): 816-819.

Rosenthal, M.B., Frank, R.G., Li, Z., and Epstein, A.M. (2005) Early evidence with pay- for-performance: From concept to practice. *The Journal of American Medical Association* 294: 1788-1793.

Rosenthal, M.B. Landon, B.E., Normand, S.T, Frank, R.G., and Epstein, A.M. (2006). Pay for performance in commercial HMOs. *The New England Journal of Medicine* 355: 1895-1902.

Singh, NS., Kovacs, RJ., Cassidy, R., Kristensen, SR., Borghi, J., and Brown, GW. (2021). A realist review to assess for whom, under what conditions and how pay for performance programmes work in low- and middle-income countries. *Social Science & Medicine* 270.

Soeters, R. and Griffiths, F. (2003). Improving government health services through contract management: a case from Cambodia. *Health Policy and Planning* 18, 74- 83.

WHO. (2023). Health in Transition (HiT) Health system reviews, European Observatory on Health Systems and Policies, World Health Organization.

Zaresani, A., and Scott, A. (2021). Is the evidence on the effectiveness of pay for performance schemes in healthcare changing? Evidence from a meta-regression analysis. *BMC Health Services Research* 21, 175.

Author Biographies

Angelique Acquatella

Angelique Acquatella is an Assistant Professor at the Toulouse School of Economics. While a PhD student in Economics at Harvard University, she was an NBER Aging and Health Fellow and an NSF Fellow. Her research studies the optimal design of health care policy, with a focus on public health insurance systems and pharmaceuticals. She is interested in policy designs that advance health equity, minimize risk for the most disadvantaged individuals, and incentivize socially valuable investments. Her work at the intersection of health economics and public finance combines methods from optimal tax theory with traditional cost-effectiveness analysis in health economics.

Tianxu Chen

Tianxu Chen is an assistant professor in residence of Economics at the University of Connecticut. She received her PhD degree in Economics from Boston University. Her fields of interest include Health Economics, Labor Economics, Chinese Economics, and Applied Econometrics.

Randall P. Ellis

Randall Ellis is a professor of economics at Boston University with a background in industrial organization and econometrics that he applies primarily to health economics, spanning both US and international economics topics. His recent work focuses on healthcare payment systems, insurance, innovation, and predictive modeling using big data. A past associate editor at JHE and AJHE, and the US country coordinator of the Risk Adjustment Network, his recent research funding is from AHRQ, HRSA, and SSRC.

Taylor Watson

Taylor Watson is an economics PhD candidate at Boston University. His research covers healthcare pricing and regulation, with a particular focus on the pharmaceutical sector. He completed his B.A. at Dartmouth College and M.A. at the Ludwig Maximilian University of

Munich. Before entering graduate studies, he spent several years working in the healthcare industry.