

WORKING PAPERS

N° TSE - 986

January 2019

“The Economic Impact of public private partnerships (PPPs) in Infrastructure, Health and Education: A Review”

Anaïs Fabre and Stéphane Straub

The Economic Impact of public private partnerships (PPPs) in Infrastructure, Health and Education: A Review*

Anais Fabre[†] Stéphane Straub[‡]

January 19, 2019

Abstract

This paper summarizes what is known about the impact of Public-Private Partnerships (PPPs) in infrastructure (energy, transport, water and sanitation, and telecommunications), education, and health. It reviews evaluations of the effectiveness of PPPs and, whenever possible, the implications for coverage and affordability. For infrastructure, there is some evidence of PPPs leading to gains in labor productivity, as most transitions to private participation have entailed significant labor downsizing, but only mixed support for improvements in total factor productivity or quality, coverage, and affordability. Selection issues plague most available studies, calling for caution when considering the conclusions. For education, micro-level studies shows that PPPs have the potential to increase enrollment and respond to a growing demand for education, but that their impact on educational outcomes appears to depend greatly on the institutional details and the scale of the programs considered. Finally, PPPs appear to be likely to increase socio-economic and ethnic segregation in education systems. Finally, there is inconclusive evidence of an impact of PPPs on health service utilization, the quality of services, patient satisfaction and health-related outcomes. The review points to the need to identify the reasons for service under-utilization, such as lack of information on the side of users, poor quality of services, before implementing interventions, and to combine demand- and supply-side policies if needed. Overall, more work appears to be due to rigorously evaluate the impact of PPPs, especially regarding infrastructure and health.

JEL Code: L33, H54, I11, I21, D04

Keywords: Public private partnerships, evaluation, infrastructure, education, health

*We thank Laurence Carter and Fernanda Ruiz Nunez for initiating this work and the World Bank for providing funding. Ana Gazmuri and Fernanda Ruiz Nunez have provided very useful comments on earlier drafts. All errors remain ours.

[†]Toulouse School of Economics, University of Toulouse Capitole. anais.fabre@ut-capitole.fr.

[‡]Toulouse School of Economics, University of Toulouse Capitole. stephane.straub@tse-fr.eu.

1 Motivation

Public-Private Partnerships (PPPs), representing an alternative contractual arrangement to traditional public procurement, are a potentially important instrument to achieve development goals. However, there is a debate regarding their effectiveness, which has been revived by the discussion about the post-2015 development agenda. Indeed, the Sustainable Development Goals (SDG) include the promotion of “effective public, public-private, and civil society partnerships, building on the experience and resourcing strategies of partnerships”, making PPPs a central part of the development tool-box. In the developed world as well, policy makers often praise PPPs, but cases of large cost overruns and delays in highly visible projects also often generate controversies and doubts about their validity.

This literature review is an attempt to summarize the state of the empirical analysis about the impact of PPPs in key sectors – infrastructure (defined to encompass energy, transport, water and sanitation, and telecommunications), education, and health - while shedding light on what remains to be studied. This review focuses mostly on PPPs in developing countries, but it also draws on studies focusing on developed countries.

The economic literature has generally defined PPPs as long-term contractual arrangements between a public authority (which can be a local or a central-government agency) and a private supplier for the delivery of some services, in which the latter takes responsibility for building a piece of infrastructure, making arrangement towards the financing of the investment and then managing and maintaining this facility (Iossa and Martimort, 2015) [86].

For the purpose of this paper we will define PPPs as “any long-term contractual arrangement between a public entity, or authority, and a private entity, for providing a public asset or service, in which the private party bears significant risk and management responsibility”¹. In practice, this means that we consider contractual arrangements such as performance based management contracts, leasing (also known as affermage), franchise, concession, build-operate-transfer (BOT), and build-own-operate (BOO) (see Guasch, 2004) [71]. This de facto excludes, at one end of the spectrum, public procurement projects such as service contracts or turnkey construction contracts, and, at the other end, outright privatizations in which asset ownership is permanently transferred to the private sector².

Overall, this rather broad definition implies that PPP contracts are not restricted to one particular sector, such as infrastructure. On the contrary, they cover a large set of potential applications, as long as they are long-term, gather private and public actors, and feature some risk sharing between the parties. As will be discussed in the

¹See World Bank Public-private Partnership in Infrastructure Resource Center: <http://ppp.worldbank.org/ppp/overview/what-are-public-private-partnerships> (accessed May 3rd, 2018).

²We will still refer to some studies that assess pure privatization cases, as these are sometimes analyzed together with a broader range of private involvement modalities.

next session, the scope of relevant PPP arrangements varies significantly across sectors.

Our main objective is to evaluate the effectiveness of PPPs. Given the diversity of sectors covered in this review, we need to define precisely for each of them which outcomes are relevant to assess the effectiveness of PPPs. For example, while traditional measures of efficiency, broadly relating to the cost of producing a given unit of quality-standardized service, are used to study infrastructure, the outcomes assessed in the education sector are generally related to enrollment rates or schooling attainment. In addition, we also discuss potential channels at play behind those impacts. Finally, whenever it is possible, we examine their implications in terms of potential trade-offs between coverage and affordability.

In a nutshell, the conclusions are the following. Regarding infrastructure, there is some evidence of PPPs leading to efficiency gains. The evidence is stronger when looking at labor productivity, a finding related to the fact that most transitions to private participation have entailed significant labor downsizing. On the other hand, there is mixed support for private participation leading to improvements in total factor productivity or quality. Overall, there is also no clear-cut conclusion regarding how PPPs have affected coverage and affordability. One must however qualify these rather inconclusive findings. Indeed, most of the available studies are plagued by selection issues, one of the main methodological challenges in the evaluation of infrastructure PPPs. Without doubt, this is the area where more work is still due.

The nature of PPPs in education is in some way allowing for a better, if not perfect, identification of their impact. A rich picture emerges from the many programs implemented and evaluated around the world. First, it seems that PPPs have the potential to increase enrollment and respond to a growing demand for education. However, more effort should be made in the future to handle migration spillovers as it is sometimes unclear whether there is an increase in overall enrollment or in enrollment in PPP schools only. Second, the impact of PPPs on educational outcomes, focusing mostly on students' test scores, appears to depend greatly on the institutional details and the scale of the programs considered. For example, the Colombian voucher system, with its small scale and embedded incentives for students to perform, has been significantly more successful than the Chilean universal voucher program. Finally, PPPs appear to be likely to increase socio-economic and ethnic segregation in education systems, and few remedies are so far available to limit this problem.

Finally, regarding health PPPs, there is inconclusive evidence of an impact on health service utilization. This finding seems related to the fact that many interventions fail to identify whether the reasons for service under-

utilization are a lack of information on the side of users, the poor quality of services, etc. In addition, the existing evidence point to the necessity of combining demand- and supply-side policies. Regarding the quality of services, patient satisfaction and health-related outcomes, findings are again mixed and limited by the lack of data capturing health status. Pay-for-performance programs seem the most effective, thanks mostly to the additional resources being channeled to health facilities.

The rest of the paper is organized as follows. The following section describes our analytical framework by defining the outcomes we consider for evaluating effectiveness in each sector, while underlining the empirical challenge faced by such studies. Section 3 summarizes the literature covering the effectiveness of PPPs in infrastructure. Section 4 and Section 5 focus on the effectiveness of PPPs in education and in the health sector, respectively. Finally, Section 6 concludes.

2 Analytical Framework

2.1 The variety of PPP arrangements

As underlined above, the main motivations behind PPPs are common to contracts in infrastructure, education, and health. However, the details of PPP contractual arrangements vary greatly across sectors. In addition, the outcomes of interest to evaluate PPPs' impacts also differ between education, health and infrastructure. In order to identify these outcomes, we should first define what is expected from PPPs in each sector.

2.1.1 PPPs in Infrastructure

The concept of infrastructure has a wide range of definitions in the literature. In this document, infrastructure is defined as all the long lasting and sunk investments used to deliver public services in the following sectors: energy, transport, water and sanitation, and telecommunications. "Public services" include any service that the relevant government considers its responsibility to provide or ensure is provided.

The main argument in favor of PPPs in infrastructure is therefore that they can lead to efficiency gains in service delivery. A number of potential channels that are specifically relevant for infrastructure projects are worth mentioning here.³ First, given the fact that PPPs require interested private parties to step in, projects that are not profitable may fail to reach financial closure. This may represent a form of market test and go some way to

³See for example World Bank (2014) [143]

improving project selection. This argument, however, as been shown to be valid only for projects with significant user-fee funding (Engel, Fischer, and Galetovic, 2014) [55].

Second, PPPs may offer a combination of private efficiency gains and public sector-specific choice of performance indicators, such as quality or coverage. Indeed, PPPs offer the possibility for the government to contract over defined performance requirements by tying payments and penalties to them, thus incentivizing the private operator to minimize its costs given these requirements. These incentives may lead them to exert specific efforts to reduce some of the main and recurring problems in large scale infrastructure projects, namely construction delays and cost overruns. Finally, for projects carrying substantial risks, for example in terms of demand, PPPs may allow for the allocation of that risk to the party which is in a better position to manage it.

The main argument behind the promotion of infrastructure PPPs, however, is that they may lead to efficiency gains in service delivery because the long-term nature of the contracts generates adequate incentives for private partners to internalize potential positive spillovers between the building and the operating phases of the project. Such long-term contracts may be efficiency enhancing because they incentivize the use of innovative solutions to the overall service delivery chain, a process in which private know-how and previous sector experience are thought to be of crucial importance. This rationale is at the heart of the debate on the potential efficiency gains from PPPs, and will be a major focus of this review. Note, however, that such possibility is not one-sided, and there could also be losses from bundling together project stages. PPPs with weak market test as they do not rely on user fees may end up increasing costs, as has been argued for example for projects involving the construction of prisons or large scale education or health facilities (Hart, 2003 [76]; Martimort and Pouyet, 2008 [99]).

This stands in contrast to an argument often heard from policy makers, which is that PPPs can mobilize additional financial resources, allowing the government to save on scarce and costly public resources. However, as forcefully argued by Engel, Fischer, and Galetovic (2013) [54], through PPPs the private sector does not fund public projects, it only finances them. Funding, understood as who pays for building and operating the facilities, will ultimately fall on a combination of users and taxpayers. As such, PPPs only affect the intertemporal allocation of government revenues and expenditures, but do not per se create gains in terms of the fiscal space available in present value terms, a form of ricardian equivalence. Such equivalence may not hold if there are added efficiency gains from private operations, for example in the form of the above-mentioned internalization of externalities leading to improved design at the construction stage. It is thus important to review whether PPPs actually imply

efficiency gains in the delivery of key public goods, or if they only constitute a political maneuver aimed at relaxing governments' current budget constraint.

Finally, it is also often mentioned that an additional cost of PPPs may stem from relying on private financial arrangement that come at a higher cost than public ones. As explained in Fay, Martimort, and Straub (2018) [?], this must be nuanced on at least two grounds. First, private management and private finance are not perfectly equivalent, although of course there is some correlation. Indeed, many PPPs rely on significant shares of public financing.⁴ Second, private financiers may be better able to monitor projects, resolve key information asymmetries, and hence ensure that a more efficient financing structure is put in place. To that extent, the simple comparison of the private and public cost of debt may not be the relevant one. More important is the fact that the implicit benefit of private finance may come at the cost of raising user prices and subsidies, and hence excluding some of the poorest potential consumers of services that often exhibit a rather elastic demand schedule.

This last point opens to the broader question of whether efficiency gains, whenever they exist, imply a trade-off between access and affordability, and whether they may lead to gains along other economic and social outcomes, such as jobs creation, poverty reduction, or health indicators.

2.1.2 PPPs in Education

PPPs in education are defined by the PPP Knowledge Lab⁵ as “long-term contractual relationships between the government and a private provider for all or some part of the delivery of education infrastructure and services.” However, this definition is large and encompasses various forms of involvement of the private sector in the educative system. Following Patrinos et al. (2009) [121], who distinguish four main forms of PPPs in education, we will consider the following contracts:

- Voucher systems, in which the financing of public education is made through vouchers. In this type of systems, households are free to choose their children's school, either private or public, while the government offers a per-student subsidy to private schools. Note that we distinguish voucher systems, in which the whole financing of education is made through vouchers, from other voucher schemes for which eligible families receive vouchers from the government, that they can use in a restricted type of schools, called voucher schools, which may require tuition top-ups.

⁴The PPI database has reported information on the financing structure of PPPs since 2015, covering a subset of over 300 projects. The mean share of public financing in that sample is around 50 percent.

⁵<https://pppknowledgelab.org/sectors/education> (accessed on the June 28th, 2018)

- Private management of public schools, in which the government outsources the operation of public schools to the private sector. This system encompasses charter schools and concession schools.
- Government subsidies to private schools, which can take the form of per-student payments or subsidies covering teachers' salary or textbooks' cost.
- Private Finance Initiative (PFI), with contracts tying the building of school infrastructure to its operation and maintenance (excluding responsibility for the educational part).

The rationale behind PPPs in education is generally considered to be the following: it is a middle-ground between pure public or private provision, combining their strengths while potentially avoiding their weaknesses in the two dimensions that should be taken into account when discussing education - quantity⁶ and quality. In addition, introducing the PPP option in a country schooling system may provide better incentives for teachers, and generate competition between public and private schools, empowering parents and students by giving them more choice and hence greater leverage to monitor school performances.

These arguments mainly apply to voucher schools, private management of public schools, and subsidies. School PFI, on the other hand are unlikely to impact substantially the quality of the educational service provided since they do not impact incentives within schools (Hanushek, 2003) [75]. The rationale behind them is to allow private firms to internalize the positive externalities present between the construction and the maintenance of an infrastructure by bundling both in a single contract. Thus, the main advantage of PFI in education resides in their cost saving function - which might in turn marginally increases resources available for education - and not in their potential to impact educational outcomes or inequalities between students. For this reason, we choose to include PFIs in the section dedicated to PPPs in infrastructure rather than in the one covering education. The section on PPPs in education will thus only review evidence of the impacts of the *provision* of education by the private sector, through the first three types of contracts.

We will also follow Patrinos et al. (2009) [121] who define four main outcomes of interest: enrollment, educational outcomes such as test scores, stratification, and costs. These outcomes allow to evaluate whether PPPs successfully meet the growing demand for education without increasing geographical and socio-economic inequalities and deteriorating the overall quality of education.

⁶Quantity should be understood as overall enrollment.

2.1.3 PPPs in Health

In the health sector, PPPs have traditionally been used in the provision of infrastructures, such as hospitals, through Private Finance Initiatives (PFI) contracts, in which “the private sector is contracted to rebuild or replace a public asset and maintain this asset for 20 to 30 years” (Health Research Institute, 2010) [79]. By bundling the two stages, this type of contracts aims at incentivizing the private operator to internalize positive externalities between the construction and the maintenance of health facilities, increasing construction quality and making maintenance more efficient.⁷ As for education, the main rationale for PFIs in health is thus its potential for cost-savings rather than for improving health-related outcomes, so we will include them in the section covering PPPs in infrastructure.

It seems that the future of health PPPs is in service delivery. Indeed, the Health Research Institute argues that rather than PFIs, it is mostly substituting the provision of health services from public to private operators that has the potential to create large efficiency gains. However, the definition of health PPPs has to be adjusted accordingly. In particular, health service provision does not always fit the traditional definition of PPPs in infrastructure, which insists its long-term perspective. Following recent contributions, we will also consider as PPPs more flexible, shorter term contractual arrangements, as long as they imply that public and private entities become partners and share risks.⁸ We will thus define PPPs in health more broadly as “any formal arrangement between a government and a private entity established for the purpose of providing health services” (Marek et al., 2005) [98]. This definition then encompasses both demand-side health policies, such as voucher schemes, and supply-side health policies, such as franchising.

The relevant outcomes used to evaluate the effectiveness of health PPPs change as their scope evolves. Indeed, they were traditionally evaluated through the lenses of Value for Money (VfM), comparing the expected cost of pure public provision to the expected payments made by the government to the private operator through a PPP. With the multiplication of PPPs in service provision, new outcomes have to be taken into account, such as patient satisfaction, measures of clinical performance (infection rates,...) or measures of the staff performance (absenteeism rates,...) (Health Research Institute, 2010) [79]. We restrict our attention to four types of impacts: health service use, health outcomes, targeting/stratification, and efficiency.

⁷PPP Knowledge Lab: <https://pppknowledgelab.org/sectors/health> (lastly accessed on June 28th, 2018)

⁸In a recent report, IEG (2016) underlined that “the type of health PPPs supported by the World Bank Group tend to be more short-term arrangements, compared to infrastructure PPPs”.

2.2 Empirical Challenge

The goal of this review is to gather sound empirical evaluations of the impact of PPPs on the outcomes spelled out above. This requires expliciting what we define as *valid, causal empirical evidence*.

When evaluating the impact of a policy, one has to determine against *what* the comparison should be made. Usually, the comparison is made against *what would have happened in the absence of the policy* for the population under study, i.e. the counterfactual. The issue is that this state of the world is, by definition, unobservable, so the key to identify an impact as causal is to find the appropriate proxy for this unobserved counterfactual. Most empirical studies covering PPPs consist in anecdotal evidence and case studies, comparing outcomes before and after: these results are subject to a “time trend bias” to the extent that many other unobservable factors might evolve at the same time and impact outcomes of interest, which might be wrongly interpreted as the effect of the PPP. Such evaluations fail to identify the causal effect of PPPs and we will not include them in our review of the empirical literature.

More specifically, when studying PPPs, we are mostly interested in their impact relative to a well-defined alternative, which is what would have happened in the presence of public provision rather than PPP for the population under study. An appropriate proxy may consist in observations from a “control” group, believed to have similar characteristics than the group impacted by the PPP but for which the same service is provided by the public sector: finding such a group is at the heart of the empirical challenge faced by PPP evaluations. The issue is that most comparisons between PPPs and public providers suffer from endogeneity, which may come from three main sources: measurement errors, reverse causality and selection bias (Wooldridge, 2012) [141]. In microeconomic studies, like the ones we will focus on, the two first sources of endogeneity should be negligible. Indeed, the main explanatory variable will often be a dummy indicating whether a structure (an hospital, a school,...) is public or under a PPP contract, so that measurement errors are unlikely.⁹ On the other hand, if a measurement error affects the outcome variable (patient satisfaction, test score, etc.), this would not result in biased estimates, but simply in a less precise estimation, making it harder to detect significant effects. Reverse causality, while not excluded, should also be negligible to the extent that PPP contracts are unlikely to be causally determined by the outcomes we will focus on, but rather by more global considerations like potential profitability, global economic or social conditions, or government ideological positions. In fact this is where the key empirical challenge of PPP evaluations appears: they are likely to suffer from a selection bias as many observed and unobserved confounding factors are likely to be

⁹The remaining concern arises in studies that conflate different types of PPPs and pure privatization in a single category.

at play, impacting both the fact that a PPP has emerged and the outcomes of interest.

We can illustrate this selection bias in the different sectors encompassed in our analysis with a few hypothetical examples. Regarding education, it might be that voucher schools concentrate in urban areas while educational outcomes are lower for children living in rural areas. From this, a simple comparison between public and voucher schools would lead to overestimate the impact of voucher schools on educational outcomes. Fortunately, this bias would disappear by simply adding a variable for rural areas as control. However, many confounders are likely to be unobserved, making researchers unable to control for them. For example, parents caring more about their children's education are likely to send them to voucher schools, while their higher level of investment in their children's education is likely to result in higher test scores. The selection issue is all the more challenging because it may come from both the supply and the demand side. On the supply side, since PPPs might not be viable in any settings, they might appear only in very specific environments or circumstances and thus serve different populations than the public sector. In addition, some PPP contracts allow schools to select their students based on test scores or interviews, creating room for individual selection based on unobservables. On the demand side, students choosing to use the service provided through a PPP might have very different characteristics than the ones choosing to use public services. A related issue is that, as underlined by Patrinos et al. (2009) [121], PPP contracts creating voucher or charter schools might have broader equilibrium effects, inducing very specific students to migrate from the public to the PPP school and thus impacting the potential control schools.

Empirical evaluations of PPPs in the health sector also face what can be characterized as a double-sided selection bias issue. For example, doctors of an Indian PPP initiative designed to provide efficient and safe maternity services for poor women in remote areas of Gujarat have been shown to handle only simple deliveries, while diverting the complicated cases such as caesarean to public hospitals (Thadani, 2014) [135]. When simply comparing delivery complications in the PPP scheme and in public hospitals, one would thus overestimate the effectiveness of the former. Moreover, there is a possibility that women choosing to use the services of the PPP scheme have specific characteristics. For example, women knowing they are at-risk pregnancies and facing a higher probability of complications might be reluctant to use the free PPP scheme and might thus choose to travel to the nearest public hospital, which would reinforce the upward-bias in the estimation of PPP effectiveness.

Studies evaluating PPPs in infrastructure are at least equally likely to suffer from a selection bias. Indeed, the choice of implementing a PPP is not random and many observed and unobserved factors might be at play. For

example, PPPs have been increasingly used to address the challenge of rural electrification in developing countries, where public monopolies are often failing to connect the poor. Then, PPPs in electricity might operate in poorer areas than the ones served by the public sector. Geographical characteristics are correlated to the type of operator that self-selects in these areas (PPP or public), as well as to poverty levels. Alternatively, it might be that PPPs self-select in areas that are expected to grow faster than others, which would be unobserved by the econometrician.

In theory, an appealing way to evaluate the relative merits of PPPs and public provision would be to randomly assign public and PPPs projects to different areas. For service provision, it would in addition require to randomly assign individuals to either the public or private provider, through a two-step randomization, while taking steps to limit migration. While such Randomized Controlled Trials (RCT) might be feasible in some context, but they remain largely impracticable in most of the cases discussed below.

One thus needs to find ways to circumvent selection issues. Some of the potential sources of selection in either infrastructure, education, or health can easily be managed. Indeed, the potential confounders that are observed by the econometrician should be controlled for. However, the main sources of selection are likely to be unobservable. Usually, panel data allow for the use of fixed effects that get rid of unobserved confounding factors that are constant over time, but several issues remain. First, a non-negligible part of confounding factors are likely to vary over time. This would require resorting to alternative strategies, such as the use of instrumental variables, conditional on the fact that the instruments chosen do not violate either the relevance or exclusion restriction assumptions. Second, in many cases the variable of interest, contrasting PPPs and public operators, may be time invariant and can thus not be identified when fixed effects are included. Finally, for service provision, one solution to circumvent this issue has been to use individual-level panel data to exploit the fact that some users switch from a public to a private operator over time, but this again requires to be able to extract the exogenous part of such moves.

Overall, as underlined by Estache and Saussier (2014) [63], “the econometric evaluation of various types of PPP experiences shows indeed that the careful choice of control variables, the proper framing of the PPPs institutional and sectoral context and the careful avoidance of selection biases in sample choices matter to the conclusions reached by empirical tests of the impact of PPPs”. Identifying the causal impacts of PPPs in either education, health or infrastructure thus represents a serious empirical challenge, and it should come as no surprise that there is a dearth of econometrically convincing studies.

3 PPPs effectiveness in Infrastructure

To the extent that most economists and practitioners agree that the main potential benefits of PPPs are maximizing efficiency gains, we will first focus on the literature covering this outcome. Efficiency gains are crucial with respect to the further improvements that they allow and that are directly part of the post-2015 development agenda: coverage, accessibility, quality of services, or poverty reduction. We will study the impact of PPPs on the latter variables in a second time.

Through this review, it will appear that the academic empirical literature on the economic impact of PPPs is very limited, with most of the existing findings based on anecdotal evidence and case studies - which we will exclude - comparing outcomes before and after without a well-defined counterfactual.

3.1 Efficiency

Infrastructure efficiency is usually measured through a proxy for productivity. A commonly used measure, partly thanks to its wide availability, is labor productivity, generally using a simple ratio of labor per unit of output. When efficiency is measured with ratio of output to inputs, it is straightforward to consider it can be improved by producing the same output at a lower cost or producing more and better outputs at the same cost, although technological progress and changes in the mix of inputs used to produce the service, for example through capital intensification, can sometimes confound the comparisons.

In addition, quality is sometimes used as another proxy of efficiency. It can be measured in a number of ways. In the energy and water sectors, interruption of services and losses in the network are a first way to gauge the efficiency of service delivery. Compliance with service standards, such as voltage stability, or water potability also matters. In the transport sectors, analogous measures are delays in the case of ports, airports, or railways, and congestion in the case of roads. Finally, user satisfaction as measured through surveys or the rate of complaints, is sometimes used.

One important restriction that limits the number of studies included in this review is that while we focus as much as possible on PPPs, and not on the more broadly defined Private Sector Participation (PSP) which also encompasses other types of private involvements such as privatization, studies often do not distinguish between these different types of contracts. This makes it difficult to draw conclusions about the specific impact of PPPs, but we still discuss these results.

In this section, we summarize findings from several studies, using either single- or multi-country samples pooling a large number of utilities. We also draw on several surveys of the literature, including Gunatilake (2008) [73], Berg and Marques (2011) [25], and more recently Porcher and Saussier (2018) [126].

Methodologies include standard panel econometrics, usually difference-in-differences either controlling for firms specific trends or applying some type of matching methodology. In the context of the diff-in-diff approach, authors have implemented several techniques to ensure basic hypothesis underpinning the validity of the methodology hold. For example, Andres (2004) [6], which was then extended by Gassner et al. (2009) [67], circumvent the two problems that traditionally plague the literature on PSP in infrastructure - namely the fact that most of the existing evidence comes from case studies and that the large majority focuses on privatization - by exploiting a large panel dataset, spanning from 1973 to 2005 and covering 301 utilities that have undergone a shift to PSP and 926 State-Owned Enterprises (SOEs) in 71 developing and transition economies. Interestingly, the information contained in their data enables them to distinguish between four different types of PSP: full divestiture (privatization), partial divestiture, concession, and lease and management contracts. They use different specifications, among which the most robust is the combination of a difference-in-differences with propensity score matching that allows them to control for both constant observable and unobservable differences between the SOEs that shifted to PSP and those who did not.

Andres et al. (2008) [7] and Andres et al. (2013) [9] implement diff-in-diffs estimations, looking at the effect of both the transition and the post-privatization entry, and controlling for firm fixed effects and firms specific trends. In addition, they envision the use of an IV procedure to address the potential selection bias, using country-level macroeconomic variables as instruments based on the assumption that private participation decisions are driven by the evolution of such indicators, but these do not affect variables like utility-level efficiency, number of connections, quality, and prices. This approach fails to capture within-country, within sector variability in the timing of changes in ownership, and thus appears unsuitable for both water, the transport, and the electricity sector. When implemented in the telecom sector, however, the IV estimates appear to be larger than the OLS ones, suggesting that better-performing firms are likely to have been privatized first, as indeed suggested by the literature. Their results also suffer from the fact that, first, although their dataset includes management and lease contracts, concessions, greenfield projects and divestitures, they are not able to disaggregate the impact by type of contract: they then use interchangeably the terms PPI (Private Participation in Infrastructure) and privatization along their study, and do not conduct any heterogeneity analysis. Second, their sample is constituted by 118 firms working in

the electricity sector that switched from public to PPI during the 1990s. They are thus isolating the impact from PPI by adding firm fixed effects to control for constant confounding factors and firm-specific time trends. However, as underlined before, because they focus only on firms undergoing a transition to PSP, their estimates are likely to suffer from a sample selection issue. Their results should thus be interpreted with these caveats in mind.

Another large set of studies relies instead on trying to identify “frontier” firms in terms of efficiency through economic modeling techniques such as data envelopment analysis (DEA) or stochastic frontier analysis (SFA), and looking at the situation of firms with private participation with regards to the best-performing counterparts. This is a highly sophisticated literature addressing the estimation of productivity and other dimensions of efficiency. However, as discussed in Section 2, the problem at hand is not only to derive meaningful measures of efficiency but mostly to address the fundamental selection issues that plague the comparisons between PPPs and their public counterparts. For this reason, we mention this strand of literature in sectors where they have been used (mostly for ports and airports, and water and sanitation), but stop short of doing an exhaustive review of such contributions.

3.1.1 Energy

Gassner et al. (2009) [67] find that, overall, a greater degree of private participation is correlated with larger gains in productivity. In particular, PSP is associated with a 50% increase in electricity sold per worker and a 42% increase in residential connections per worker, both being statistically significant. However, the authors note that results vary by type of contract. Indeed, in the electricity sector, the impact of PSP on electricity sold per worker is largely positive and significant for privatizations only, while it is non significant for partial divestitures and even significantly negative for concession contracts. These findings hold for the number of connections per workers, at the difference that this time neither partial divestitures nor concessions have a significant impact on this outcome. The only result that could signal an increase in efficiency for PPPs is that the electricity lost in distribution decreased by 14.3% for partial divestitures. Overall, if privatizations seem to create efficiency gains, it does not appear to be the case for PPPs in electricity. This result is particularly important in the sense that it underlines how we should abstain from extrapolating the estimated impact of privatizations - or non-disaggregated PSP - on efficiency gains to the case of PPPs.

We also have to note that one should be particularly careful when interpreting the results from Gassner et al. (2009) [67]. Indeed, the findings from their difference-in-differences combined to propensity score matching are different from the ones from their simple panel analysis adding utility fixed effects. The latter conclude that partial

divestitures did significantly increase efficiency gains, measured by connections per workers and electricity sold per workers. However, this analysis only uses observations for which the type of the utility ownership varies over time, from SOEs to PSP. However, such utilities represents a selected sample and their ex-ante performance might have been significantly different from that of other SOEs which did not undergo a change of status. For example, it might be that only poor performing SOEs are considered for a PSP, which would bias the estimates of the impact of a PSP upwards. This led us to restrict our focus on the results from the difference-in-differences.

Andres et al. (2008) [7] find that the number of connections per employee as well as the energy sold per employee first significantly increase, in the transition period,¹⁰ before decreasing in the post-transition period. Then, PPI do not appear to have any significant productivity impacts in the long-run. Moreover, even these observed short term improvements might hold only for privatizations, as shown by Gassner et al. (2009) [67].

Similarly to Andres et al. (2008) [7], Estache and Rossi (2004) [62] are not able to differentiate between privatization and other types of reforms which would fall under the definition of PPPs. They first estimate a parametric labor requirement function in order to evaluate the impact of PSP on firms' efficiency, before checking the robustness of their results through the estimation of a stochastic frontier model, a between-firms variations model and a first-difference model. They use a sample of 110 firms in the electricity distribution sector from 14 countries in Latin America during 1994-2000. These firms had different levels of private sector participation in the ownership structure as well as different types of regulation applied to them for electricity distribution (price cap or rate of return). The authors find that private firms are on average significantly more labor efficient than the public ones, a result which is consistent across different specifications. However, again, it is not possible to determine if these results would hold for PPPs alone.

The work of Mollisi (2016) [107] has the particularity to focus exclusively on PPP in the energy sector. He relies on data collected from the Italian district heating industry in order to estimate the impact of PPP on efficiency through a structural model. He shows that firms under a PPP perform significantly better: PPPs are associated with a 14% increase in output. In addition, as expected in theory, the author finds "that a PPP contract allows a technological externality between the different phases of a project to be internalized, leading to a positive effect on total factor productivity". In particular, it is estimated that, for firms under a PPP, a unit increase in capital quality raises output by 15%.

¹⁰The transition period spans from approximately two years before the concession is awarded - i.e. when the reform is announced - to one year after

Conducting a very different type of work, Hoppe et al. (2013) [80] also “find support for the theoretical prediction that, compared to traditional procurement, a public–private partnership provides stronger incentives to make cost-reducing investments”. They reach this conclusion through a large-scale public procurement laboratory experiment, in which participants have to choose investment levels for an infrastructure they build and operate under a traditional procurement contract or a PPP. They find that PPPs induce very strong incentives to invest in cost reductions compared to traditional procurement contract, which is likely to imply efficiency gains: high level of investments are chosen more often under a PPP contract. Of course, conclusions arising from laboratory experiments might be extremely difficult to extrapolate to real world behaviors. Still, when the empirical literature is as limited as it is to evaluate PPPs, such studies offer a useful way to test whether the theoretical arguments backing PPPs have the potential to hold in reality.

3.1.2 Transport

Some evidence on the efficiency impact of transport PPPs also exists. Contrary to the energy sector, most of them have focused on the analysis of a single country. For example, Estache et al. (2002) [61] study the effect of port reforms in Mexico, relying on a stochastic production frontier approach. The authors underline how during the 1990s the need for an increased efficiency in ports brought most of the LAC countries to implement reforms aiming at introducing more competition by allowing the participation of the private sector. Specifically, in Mexico, reforms resulted in the strong penetration of the private sector in port services, through the attribution of concession contracts. The authors note that the last phase - consisting in the privatization of the port administration - had not been successfully implemented five years after the start of the reform. Taking advantage of a panel dataset of 11 ports spanning from 1996 to 1999, they estimate that the set of reforms increased the average efficiency level by approximately 5 percentage points, while the average growth rate in port efficiency was 2.8-3.3%.

Estache et al. (2004) [60] complement this analysis by calculating and decomposing the changes in productivity that happened for Mexico’s 11 main ports between 1996 and 1999, two years after the reforms. Using a DEA (Data Envelopment Analysis) decomposition of a Malmquist TFP (Total Factor Productivity) index,¹¹ the authors show that TFP in Mexican ports rose by an average of 4.1% a year during 1996–1999. Their decomposition enables them to conclude that “while labor adjustments were a factor in the improvements in TFP — as suggested by many critics of privatization — they were clearly not the only one. Indeed, adoption of new technologies and increases in

¹¹The Malmquist TFP index is defined by the authors as a measure of “the TFP change between two data points by calculating the ratio of the distances of each data point relative to a common technology”.

capacity resulting from large investments made right after the reforms began were the main contributors to these improvements”. However, the authors also underline that DEA is not free from issues. For example, operators identified as the best practice through the analysis may not be efficient. In addition, the technique fails to properly identify the impact of the Mexican reform, since it picks up the impact of confounding factors such as major shocks to international trade, like the East Asian crisis.

Despite these limitations, other researchers rely on DEA in order to study changes in efficiency in ports. This is the case of Wanke and Pestana Barros (2015) [138], who use this technique to assess the impact of PPP on scale efficiency in Brazilian ports. Interestingly, this study focuses specifically on PPP: a number of Brazilian ports have undergone a transition to PPP between 1995 and 1996, so that they are today public ports with concession agreements to private terminal operators. Using data for the year 2012 over 27 major Brazilian ports, among which one third is under a PPP, the authors conclude that “PPPs seem to be an effective way, under the Brazilian regulatory system, to move port operations towards their most productive scale size” since they find that PPPs have a positive impact on port scale efficiency. However, this result should be interpreted very carefully to the extent that the PPP estimate is only significant at the 10% level, and it is likely to be biased due to the previously mentioned DEA limitations regarding sample selection.

Apart from these country-specific studies, some researchers have also taken advantage of cross-country data in order to conduct the same type of analysis. Regarding efficiency in road construction, Blanc-Brude et al. (2009) [31] focus on the cost of constructing road infrastructure assets, using a database of road projects financed by the European Investment Bank (EIB) between 1990 and 2005. They find that a 24 percent excess cost of PPP roads over traditionally procured ones, and conclude that this corresponds to the inclusion in the ex ante PPP construction contract prices of the construction risk that is transferred to the private sector. The validity of that conclusion rests however on the disputable assumption made by the authors that the choice of procurement method is exogenous within the road sector, i.e., that roads built under PPP and traditional procurement have similar observables and unobservables characteristics.

Cheon et al. (2010) [43] use a panel dataset for the years 1991 and 2004, covering 98 major ports in the world in order to evaluate how port reforms impact efficiency gains. They direct their attention to institutional reforms, often based on an increased private sector participation through leasing, concessions or privatization. Again, in this case, PPPs are mixed with pure privatization, making it impossible to disentangle the impact of the former on

efficiency gains. Similarly to Estache et al. (2004), the authors perform their analysis by implementing a Malmquist Productivity Index model through a DEA approach. Comparing the changes in port efficiency between ports that have restructured ownership through private sector participation and those that have not, they find that the former have benefited from larger efficiency gains between the two dates. However, this finding is only significant at the 10% level. Unfortunately, the same caveats as above prevent us from inferring any robust conclusions from the further decomposition of efficiency gains into scale efficiency changes, technical efficiency changes and technology progress.

Finally, Trujillo et al. (2013) [136] use the same technique than Estache et al. (2002) - a stochastic production frontier approach - to assess the impact of reforms on African ports' efficiency. Those reforms are a black box, with likely differences between countries, but most of them seems to imply an increased participation of the private sector and some form of PPPs. Their results show that port efficiency has increased between 2004 and 2007. Looking at efficiency by type of ports, the authors find that efficiency is the highest in ports under a PPP. However, although the authors interpret these results as the potential influence of port reforms, many other variables could be at play in those relationships.

Airports have undergone the same type of reforms than ports, implying an increased participation of the private sector, as underlined by Perelman and Serebrisky (2012) [122] for the case of Latin America. Implementing a DEA analysis, they analyze data covering 22 LAC airports in the years 2005 and 2006. They also use data from 23 airports from Asia-Pacific, 40 from Europe and 63 from Canada and the US in order to be able to make some comparisons. After having shown that the majority of the LAC airports in their sample seem to suffer from inefficiency, the authors estimate a truncated regression in order to study the drivers of airport efficiency. They find that "airports which are privately operated tend to stand closer to the efficient frontier than their publicly operated counterparts", but this result is not robust to different model specifications. Finally, denoting with the previous results regarding port efficiency, Perelman and Serebrisky (2012) find that privately operated airports are not associated with higher productivity gains than public ones.

These findings are actually similar to the ones from Pestana Barros and Dieke (2008) [124] who complement their previous work (Pestana Barros and Dieke, 2007) [123], and assess the efficiency of 31 Italian ports using a DEA analysis, by regressing efficiency on different contextual variables. Indeed, although they interpreted their results as a confirmation that "airports that are partially private contribute to efficiency", they are only significant at the

10% level, calling for caution when concluding on the relative efficiency of (partially) privately managed airports and publicly managed airports.

In the same vein of these previous findings, the results from Fung et al. (2008) [65] can be interpreted as the absence of a significant impact on efficiency from the increased private participation which followed the reforms implemented in Chinese airports from the 1990s. Indeed, after having performed a DEA analysis, their decomposition of the Malmquist index illustrates that, although the average annual growth in airport productivity was above 3% in 25 Chinese regional airports between 1995 and 2004, the major source of productivity growth is technical progress, and not improvement in efficiency. Still, the authors find that airports listed on the stock exchange are more efficient than the non-listed ones. However, it is unclear whether this was triggered by private sector participation or if it captures the fact that only the most efficient airports were considered to be listed. In any case, despite their lower performance in level, non-listed airports actually benefited from a higher increase in productivity between 1995 and 2004 than listed ones. Chi-Lok and Zhang (2009) [44] perform a similar analysis and, while confirming that listed airports appear to be significantly more efficient than non-listed ones, show that being listed is not significantly related to efficiency gains.

A last set of papers gives more insights on the relative efficiency of PPPs in particular to the extent that they differentiate between various degrees of private involvement. Oum et al. (2006) [119] use a sample of 116 airports located in Asia-Pacific, Europe and North America, spanning from 2001 to 2003, to develop a measure of their productivity. Then, they regress this productivity measure on ownership/governance types.¹² Their results are very surprising with respect to PPPs: they find that airports operated by 100% government-owned corporations are significantly more efficient than airports under a PPP. In addition, the former do not perform significantly less efficiently than airports with a private majority. This policy recommendation is further confirmed by Oum et al. (2008) [120] who apply a stochastic frontier analysis to a panel data OF 109 of the world's major airports. They conclude that, given their lack of efficiency, "mixed ownership of airports with a government majority should be avoided in favor of even 100% government owned public corporation, despite the fact that many countries regard PPP with government majority as a politically acceptable model to raise private funds for infrastructure capacity expansion without losing government control". However, although the authors state that endogeneity should not be a problem in their analysis, we have to remind that these results might just reflect the fact that, for example,

¹²The categories are the following: government agency or department operating an airport; mixed private-government ownership with private sector owning a majority share; mixed government-private ownership with government owning a majority share; government ownership but contracted out to an airport authority under a long term lease; multi-level governments form an authority to own/operate airports in the region; 100% government corporation ownership/operation.

governments are more likely to undertake PPP for inefficient airports.

Finally, although PPP contracts in road projects can be found, the literature covering them is much more scarce. This is actually what is underlined by Chen et al. (2017) [42], who conduct the evaluation of the socio-economic impact of a US highway PPP project, using a Dynamic Computational General Equilibrium (CGE). They do not focus on efficiency gains but rather on broader welfare impacts, so that we comment below on their results.

Overall, private participation seems to have different impacts on efficiency according to the specific transport sector it is applied to. For example, an increased participation of the private sector in ports appear to be more often correlated with efficiency gains than in the case of airports. One reason for the failure of the latter might be the high rates of contract renegotiations, underlined by Guasch et al. (2008) [72], who explain that 53% of the concessions in the transport sector in LAC were renegotiated, on average only 3.1 years after the contract was initiated. This illustrates that private operators often behave opportunistically, in order to benefit later on from modifications in contracts' or regulatory terms. For example, they may chose to not maximize efficiency investments in the short run to benefit from more favorable regulation later on, an occurrence known as the ratchet effect in the literature (Laffont and Tirole, 1993), or they may strategically underbid at the award phase, to renegotiate and modify the contract's terms later on. In addition, it is not possible to robustly conclude that private participation in a specific transport sector significantly impact efficiency gains, as several issues weaken most of the analyses presented. First, the relationships presented are likely to be spurious due to the presence of sample selection. Second, for most of the studies reviewed here, various degrees of private sector participation are blurred together, although they would require an heterogeneity analysis to properly account for the impact of PPP on efficiency.

3.1.3 Water

When it comes to the water and sanitation sector, evaluating efficiency using labor productivity is expected to produce sharp results. Indeed, for water and sanitation utilities, labor generally represents a large fraction of costs, and presents very large variations across public and private firms, as the former have often been use as employment handle by politicians.

Water losses, also referred to as non-revenue water (NRW), correspond to the water that enters the distribution system, but is finally not billed. This may be due to physical losses, such as spills and leakages, and to commercial losses, such as theft, unbilled consumption, and metering errors. NRW is usually considered to be a good proxy for inefficiency. Depending on the source of data, it may or may not be possible to disentangle the physical losses

component from the commercial dimension and to identify for example the bill collection ratio.

Other things equal, improvements in water quality can also be considered to proxy for efficiency gains. Quality can be measured in a number of ways. These include first proxies of the continuity of the water service, such as daily hours of availability, or the frequency of service interruptions. Regarding the quality of the product itself, standard measures are compliance with potability tests or more generally with water composition requirements.

For water services, Gassner et al. (2009) [67] rely on 141 utilities under private operation and 836 SOEs, most of them in Latin America and the Caribbean, Europe and Central Asia, and East Asia and Pacific. They show that PSP is associated with gains in performance and labor productivity: it leads to a 54 percent estimated increase in connections per worker, and a 18 percent increase in water sold per worker (for concessions), relatively to SOEs. These results are partly driven by large employment decreases (22 percent) under private management, but also by an increase in the average number of connections. Interestingly, and contrary to the results for electricity, these results appear to hold for all types of private involvement, and specifically for PPPs. For example, residential connections per worker increase by 47 percent for concessions and by 78 percent for partial divestitures, while water sold per worker increases by 18 percent for concessions, 40 percent for lease and management contracts, and 62 percent for partial divestitures. On losses and quality, the study is inconclusive, with either null results or results failing to replicate across different methodologies, regarding the impact of PSP on the bill-collection rate, water distribution losses, and number of hours with water daily. One limitation of the study is that results for subcategories of PSP cannot be robustly established with the matching procedure, as disaggregation leaves insufficient data.

Andres et al. (2013) [9] use an unbalanced panel data including 116 firms from Latin America and the Caribbean and 1,103 firm-year observations. They uncover a significant drop in the employment level of water utilities, both during the transition to PSP, and after it. Given the excess employment levels mentioned above, governments often started downsizing the labor force before involving private partners, in order to make the companies more attractive. As a result, labor productivity, measured with water connections per employee, increased 30.7 percent during the transition and another 42.5 percent after the transition. However, once firm-specific trends are controlled for, the increase in productivity growth during the transition appears to be largely reverted post-transition. On the other hand, they find a significant decrease in NRW, of 3.8 percent during the transition, and 14.4 percent in the posttransition period.

Jiang and Zheng (2014) [87] analyze a panel data set of water utilities from 200 Chinese cities between 1998 to 2007. Starting in the 1990s, there was entry of foreign water companies, such as Veolia and Suez, in particular through build-operate-transfer (BOT) schemes. They use a panel model, controlling for firm- and time-fixed effects, for utility and city-level variables, for factors explaining the probability of a shift to PSP (it is shown to respond to utilities' profitability, liability level, and to cities road infrastructure in the prior year), and distinguishing PSP with a minority versus majority of private shareholding. They conclude that PSP in water utilities has led to cost savings through employment downsizing and reduction in managerial expenses, both effects being significant economically and statistically. In addition, they find positive, although not significant effects on labor productivity and TFP.

Porcher and Saussier (2018) [126] provide a useful summary of a few key studies applying DEA and SFA techniques. Their main conclusion is that in developed countries there is either no gap in efficiency between the privately owned and the public companies or that the latter actually outperform their private counterparts, while in developing countries some papers do find an efficiency advantage for private companies, but this does not seem to be robust. Papers looking at industrialized countries cross-section samples include among others Bhattacharyya et al. (1995) [28], who apply a translog variable cost function to the data of 221 U.S. water utilities in 1992, Shih et al. (2004) [132] who apply DEA to more than 1,000 observations of US water suppliers. Both find public utilities to be more efficient. Using a panel of ten UK companies privatized in England and Wales in 1989 and using cost function and TFP analyses, Saal and Parker (2001) [129] find that the transition to the private sector did not lead to a statistically significant reduction in the growth rate of total costs or to an increase in productivity. Regarding developing countries samples, Kirkpatrick et al. (2006) [89] use 110 water utilities in African countries, finding higher relative efficiency of privately owned utilities when applying the DEA method, but no statistically significant difference when using SFA. Estache and Rossi (2002) [61] apply a stochastic cost frontier model to 50 developing and transition countries water utilities in the Asian and Pacific region, failing to find significant differences between efficiency under public and private ownership.

Caution is in order when interpreting these results. Indeed, many of the reasons why privatization may have an impact on efficiency, by changing the incentive structure of managers and workers, limiting capture and political pressures, and hardening the soft-budget constraints, also point to potential selection effects related to the decision to implement PSP based on utilities' observables and unobservables and on specific shocks. Indeed, several studies have documented such selection at the privatization stage, based on sector- or firm-level profitability, as well as

quality of countries' governance (Auriol and Straub, 2011) [15]. Equally important may be the influence of factors that change along with the ownership shifts, such as the regulatory environment.

Overall, given the methodological difficulties in addressing identification issues, these mixed results should come as no surprise: another recent survey, Berg and Marques (2010) [25], review 47 studies, of which 18 find private water utilities to be more efficient than the public ones, 12 find the opposite result, and 17 fail to find a difference.

Regarding quality, studies from France and the US provide contradictory results. For example, Porcher (2012) [125] finds private management to lead to a slightly higher percentage (1 percent) of the tests complying with required microbiological standards in a sample of 2,200 French municipalities over the 1998 to 2008 period. For the US, Lyon et al. (2017) [95] analyze how changes in water system ownership (both privatization and municipalization) affect the number of Safe Drinking Water Act (SDWA) violations for a system in a given year, using difference-in-differences applied to matched pairs of water systems. They conclude that both types of transition (privatization and municipalization) lead to statistically significant reductions in water quality violations. They interpret this as supporting a "change would do you good" message, and infer that improvements are specific to the economic and institutional context rather than related to the market- or public-nature of management. In a related paper, Montgomery et al. (2018) [108] argue that water quality outcomes are mostly driven by the type of stakeholder attention that water system attract: national social movement organizations for large, privately owned systems, and local stakeholders for smaller, and publicly owned systems.

In developing countries, Andres et al. (2013) [8] find improvement in water potability during the transition, while a well-known paper by Galiani et al. (2005) [66] document that the privatization of the Buenos Aires water system in Argentina led to large reduction in water turbidness and in spillages, and as a result in child mortality.

3.1.4 Telecommunication

The telecommunication sector has mainly undergone full privatizations rather than PPPs, explaining the scarcity of studies we can include in this review. Indeed, Andres et al. (2008) [7] show that, between 1990 and 2004, no management and lease contracts or concessions were undertaken in the Latin American telecommunication sector. It appears that they only exploit a panel data of 16 telecommunication companies that all undergone a privatization during the 1990s. In this case, their results only reflect the impact of pure privatization on efficiency.

To the best of our knowledge, only one study focuses on contracts that at least partly fall under the definition of PPPs. Fink et al. (2003) [64] make clear that, between 1985 and 1999, in the countries they study, the privatization

of the telecommunication sector was often only partial, so that it can in some cases be considered as shedding light on PPPs.¹³ Taking advantage of a database covering 86 developing countries in Africa, Asia and Latin America and spanning from 1985 to 1999, they estimate a cross-country specification, adding country and time fixed effects in order to control for constant differences between countries and time trend, such as technical progress, respectively. They find that the number of mainline per employees significantly increased after a country decided to privatize its telecommunication sector. Interestingly, their results also suggest that the presence of a regulator impacts positively productivity when the sector is privatized. In addition, productivity also increases when privatization is completed by reforms increasing competition. Actually, the authors advance that the “beneficial effect of competition primarily occurs through its interaction with privatization”. These findings might not hold under an heterogeneity analysis, as the positive and significant impacts might be driven by the countries that have undertaken full privatizations. Again, one should thus be very careful when extrapolating these results to the case of PPPs.

3.2 Coverage, Affordability, and Social Outcomes

The main point in studying efficiency gains is that they can ultimately translate into improved coverage and accessibility, as well as increased affordability. However, this impact is not immediate or systematic as it requires specific actions from the private entity involved in the PPP, such as a decrease in prices. For this reason, outcomes related to coverage and affordability should be studied more carefully.

3.2.1 Energy

While we have seen that it is unclear whether PPPs in the energy sector generate efficiency gains, it is useful to review the few evidence at hand with respect to coverage and affordability. The results from the difference-in-differences combined to propensity score matching conducted by Gassner et al. (2009)[67] show that partial divestitures are the only type of PSP in electricity that are correlated with a decrease in supply interruptions, suggesting an increase in service quality. In addition, this contract might also improve coverage, to the extent that partial divestitures are also associated with an increase in the number of residential connections. However, a simple panel estimation, adding utility fixed-effects, reach the opposite conclusion. Unfortunately, results regarding residential coverage are not available for the difference-in-differences strategy because of the insufficient size of the

¹³Although there is no consensus on the definition of PPPs, the World Bank’s Public-Private-Partnership In Infrastructure Resource Center considers that partial divestitures fall under the scope of PPPs: <https://ppp.worldbank.org/public-private-partnership/agreements> (lastly accessed: June 27th, 2018).

sample at hand. Regarding affordability, the authors do not find any significant impact from PSP on the average residential tariff. This indicates that the potential efficiency gains failed to translate into lower tariff. In addition, PSP were associated with a 28% drop in employment. Importantly, this impact varies across types of contract: indeed, while employment decreases by 47% for full divestitures, the results from the difference-in-differences show no significant differences between utilities under partial divestitures or concession contracts and state-owned peers. Still, non significance should not be interpreted as the absence of impact of these types of PSP on employment. Indeed, these findings might simply be explained by the lack of precision of the estimation.

Although they are not able to disaggregate them by type of contracts, Andres et al. (2008) [7] also find that PPPs significantly decreases employment in the transition period. Their results with respect to coverage are not robust to controlling for time trends. Indeed, while both the number of connexions and coverage seem to significantly increase under a PPP, the authors note that “these outcomes would have occurred in the absence of privatization”. In addition to this failure to improve coverage, PPP also results in a decrease in affordability to the extent that it is associated with a significant increase in electricity prices. The only improvement that PPPs implies appear when looking at quality indexes: both the average duration of interruptions per consumer and the average frequency of interruptions per consumer fell, at least during the post-transition period.

Finally, Estache and Rossi (2004) [62] - although they warn that their results are only partial to the extent that many variables are not taken into account in their analysis - explain that “a comparison of the changes in prices and labor productivity reveals that, in most cases, final prices to customers did not fall to reflect the huge labor productivity gains that were achieved during the period under analysis”, implying that some of the gains were channeled into rents and higher tax revenues for the private operators and the state, respectively. Overall, given the limited evidence at hand, it seems that PSP in electricity did not substantially improve neither coverage or affordability.¹⁴

3.2.2 Transport

Unfortunately, the number of studies of PSP in infrastructure focusing on outcomes related to coverage and accessibility is extremely limited. For example, while Estache et al. (2004) [60] stress that efficiency gains are not an end in themselves by underlining that “competitiveness will improve only if the efficiency gains are shared with users”, for example through scheduled tariff adjustments, they fail to study such outcomes.

¹⁴This is reminiscent of the argument made by Martimort and Straub (2009) [100], and Bonnet et al. (2011) [32], to explain the high level of popular discontent with privatization in Latin America in the 1990s and 2000s.

Although not directly related to coverage and accessibility, Trujillo et al. (2013) [136] assess how the corruption and quality of 37 African ports are correlated to different types of ownership. They find that the highest quality is found in ports under a PPP contract, while they also display the lowest level of corruption. However, and especially given the small size of the sample, these relationships are nothing more than correlations and do not allow to infer any conclusions about the causal impact of PPP on these outcomes.

Interestingly, Oum et al. (2006) [119], using their sample of 116 airports, are able to evaluate the impact of the type of airport ownership on user charges. They find that “airports with a private majority have significantly lower average aeronautical charges than other airports”, suggesting that privatization did not result in monopoly pricing.

An alternative approach to gain more insights on the welfare impacts of PPPs in infrastructure is proposed by studies using a computable general equilibrium (CGE) model. This is the case of Chen et al. (2017) [42], who use a dynamic CGE model in order to investigate the socio-economic impacts of a highway PPP in the USA. This approach allows them to measure the infrastructure capital expenditure and tax shock effects, which they then compare to two public sector comparators (PSCs) by assuming that both differ in investment amount and the type of funding source. They estimate that both public sector scenarios generate greater welfare losses than the PPP, to the extent that traditional procurement models’ rely much more on public funding, implying larger negative tax effects than the PPP contract. Still, one should be careful in interpreting these results since they might be highly dependent on the set of assumptions made by the authors, such as the one stating that “public financing depends exclusively on tax revenues”.

For transport, the bottom line is that it is hard to draw conclusion on access and affordability as no general, universally agreed access metrics exist for the transport sector, contrary to the water or energy sectors, in which percentage of households connected to the service provide a simple way to analyze such changes.

3.2.3 Water

The issue of coverage is only relevant for developing countries, as most developed countries have already reached universal access. A number of studies using developing countries sample have found an increase in connections under private participation. For example, Gassner et al. (2009) [67] find that PSP leads to an increase in the total number of residential connections during the transition period (8 percent) and the post-PSP period (12 percent), an outcome mostly driven by concessions. Andres et al. (2013) [9] find that the number of water and sewerage connections increased during both the transition and post-transition periods, but that this evolution can be accounted for by

firm-specific trends. Analyzing water privatization in Bolivia, McKenzie and Mookherjee (2003) [101] conclude that there was an improvement in access for the poor.¹⁵ In Argentina, Galiani et al. (2005) [66] also show that the proportion of households connected to the water network increased significantly with privatization (by 2.8 percent in a diff-in-diff setting comparing municipalities which privatized before 1997 to those who did it later, and by 5.8 percent when excluding Buenos Aires from the sample), but it is also the case that privatization appears more likely in larger and less well off municipalities, and in those with a local government from a party aligned with the federal government pushing the privatization program.

Regarding affordability, the results are generally mixed. As a matter of fact, one should not expect increased efficiency to translate automatically in a price decrease, as the transition to private participation has often occurred in a context of prices that were set well-below cost recovery levels.

For developing countries, McKenzie and Mookherjee (2003) [101] find a decrease in prices in Bolivia, while Barbosa and Brusca (2015) [19] find no significant differences between public and private management, conditional on some type of regulation being implemented. Similarly, in developed countries results vary widely. As nicely discussed in Porcher and Saussier (2018) [126], the case of France, which has been widely studied, illustrates the difficulties of reaching unbiased conclusions. Indeed, private management is typically associated with a higher average price, but this difference disappears once controlling for both technical characteristics of the water supply conditions, such as density, required treatment, etc., and for the contractual characteristics of the water concessions, and addressing the endogeneity of private management through a fixed effect model assuming that any unobserved heterogeneity that leads a municipality to opt for a specific organization form is time-invariant (Chong et al., 2015) [45]. The only exception seems to be for small municipalities under 10,000, where prices under private provision exceed those under public provision by approximately 11 to 13 euros per 120 cubic meters. Chong et al. (2015) [45] interpret their results as indicative of the greater ability of larger cities to challenge private incumbents by attracting competitive bids or bringing the service in-house at the time of renewal time.

Finally, regulatory arrangements add to the difficulty of identifying the specific effect of management arrangements. For example, Barbosa and Brusca (2015) [19] do find privately provided water to be more expensive when unregulated. This points to the relevance of regulatory arrangements such as increasing block tariffs, or social tariffs, to limit potential price increases in case of private participation.

¹⁵This is not taking into account the impact of the Cochabamba failed privatization experience.

3.2.4 Telecommunication

The literature evaluating PPPs in the telecommunication sector is highly limited, with only the study by Fink et al. (2003) [64] covering some PPPs. Exploiting the panel structure of their database by adding time and country fixed effects, they estimate that countries that have (at least partially) privatized their telecommunication sector have benefited from a significant increase in mainline penetration per individual, which can be understood as a proxy for coverage. Similarly to what they found when studying productivity, they also show that the presence of a regulator and an increased degree of competition have positive impacts on coverage in countries where the telecommunication has been privatized. However, it is unclear whether we can extrapolate these findings to conclude that PPPs in telecommunication have a positive impact on coverage: indeed, no heterogeneity analysis has been conducted by Fink et al. (2003), so that these results might be valid for full privatizations only.

3.3 Private Finance Initiatives in Health and Education

Private Finance Initiatives are not restricted to the traditional infrastructure sectors. Indeed, this type of contract has also been used in the health and education sectors. They consist in contracts tying the building of an health or education infrastructure to its operation and maintenance. In theory, bundling these phases in a single contract should, as underlined above, allow private firms to internalize the positive externalities present between the construction of the infrastructure and its maintenance, resulting in a higher efficiency.

Unfortunately, the literature evaluating the impact of PFIs in health and education consists mainly of case studies, which fail to provide robust estimates. Through their review of the literature of PPPs in health, Roehrich et al. (2014) [127] tend to conclude that the “hospital build quality is not unambiguously better for PFIs, and facilities management services provide actually lower value for money (VfM) when compared to non-PFI hospitals”. For example, Liebe and Pollock (2009) [93], focusing on the case of the UK, advance that the cost of finance is higher under PFI and that facilities management services provide less value for money compared to non-PFI hospitals. Regarding the education sector, Patrinos (2009) [121] notes that “there is little actual evidence that private finance initiatives lead to cost savings”. The French Comptroller General office reaches a similar conclusion when evaluating prisons built as PFIs (Cours des Comptes, 2017) [48].

In addition to this lack of robust evidence, the way the outcomes of PFI are measured might be problematic. Indeed, they are traditionally assessed through Value for Money (VfM), comparing the expected cost of pure public

provision to the expected payments made by the government to the private operator through a PPP. However, as underlined in a report conducted by the Health Research Institute (2010) [79], VfM might be hypothetical, “subjective, difficult to measure, intangible and misunderstood”. They thus recall that other tools should be used for evaluating PFIs.

Finally, a certain consensus seems to have emerged regarding the potential impact of PFIs on educational and health-related outcomes. Authors broadly underline that the rationale of PFIs should not be extended beyond their cost-savings function. Indeed, Patrinos (2009) [121] notes that it is largely unclear how PFIs could affect learning. In the health sector, the Health Research Institute (2010) [79] even argues that substituting the provision of health services from public to private operators would create larger efficiency gains than PFI in infrastructure provision.

Overall, no robust evidence is available to assess the impact of PFI on efficiency in the health and education sectors. As of October 2018, the British government, which had led the way into the PFI experiment, announced it was giving up on this modality for new infrastructure investments because of their fiscal implications and the lack of evidence about their value-for-money.¹⁶

4 PPPs effectiveness in Education

In this section, we review the empirical evidence on the relative impacts of PPPs on enrollment, educational outcomes, stratification and costs. We have to underline that studies covering PPPs in education use a broad range of counterfactual definitions. Some, for example, use randomization at the student level in order to evaluate the impact of *receiving a voucher*. Others exploit variations at a broader level in order to compare privately managed schools to public schools, or voucher schools to regular private schools. In fact, the empirical strategy implemented is largely dependent on the scale of the program, which determines the level of the treatment variable (student, school, municipality, etc.) that is of interest. In particular, universal programs, that affect all students in a population, are particularly difficult to evaluate as there is no clear comparison group that can be used. In addition to these heterogeneous definition of counterfactual, one should be careful when defining what can be understood as a PPP in education. Indeed, while many papers compare private and public schools, we have to remember that the presence of private schools alone does not signal that PPPs are implemented in the educational system: here, we review only

¹⁶<https://www.theguardian.com/uk-news/2018/oct/29/hammond-abolishes-pfi-contracts-for-new-infrastructure-projects>.

papers that study contracts falling in one of the three first categories of PPPs in education defined above: voucher systems, private management of public schools, and government subsidies to private schools.

Before reviewing the empirical evidence, it is useful to discuss the broader context of education policies and the promises of the private provision of education. Regarding quantity, the success of policies aimed at increasing the enrollment rates in primary and secondary education in developing countries has created a new challenge: the offer has to match the growing demand for education. For example, as underlined by Barrera-Osorio et al. (2016) [21], many governments have abolished schooling fees, resulting in “sudden and massive increases in enrollments that the public education systems are generally not equipped to absorb.” Fee-charging private schools have emerged in response, absorbing the excess demand for education. Private schools are thus a way to extend the supply of education services when governments reach their capacity limit.

Maybe more importantly, the private provision of education may have the potential to solve the inefficiency of the public service. Indeed, the latter largely fails to provide the right incentives to solve moral hazard issues. In particular, the extent of teachers’ absenteeism has been identified as one of the key reasons behind the low quality of public education in developing countries (World Bank, 2017) [144]. To give a few examples, a survey conducted by Chaudhury et al. (2006) [41] reports that teachers’ absence rates¹⁷ reach 25% in India and 27% in Uganda. Even worse, among teachers present in government primary schools in India, half were found not working.

In comparison, Muralidharan and Kremer (2009) [110] find that, in rural India, teachers in private schools were 2-8 percentage points less absent and 6-9 percentage points more likely to be teaching than teachers in public schools during random unannounced checks. The main reason behind these gaps is likely to come from the different incentives faced by teachers in public and private schools. In fact, the higher flexibility of private schools with respect to teachers’ contracts and tuition fees allows them to play on the different levers identified by Banerjee and Duflo (2006) [17] to address absence in public education or health facilities: financial incentives and beneficiary control over service providers. First, private contracts give schools more leeway with respect to financial incentives, implementing performance-based or attendance-based financial bonuses and penalties to teachers, which have been shown to have the potential to significantly decrease teachers’ absenteeism and increase test scores (Duflo et al., 2012) [52]. Second, since private schools often charge fees, they are subject to a higher degree of accountability to the parents (Barrera-Osorio, 2016) [21]. This later point directly addresses the fact, mentioned by Banerjee and Duflo, 2006) [17], that beneficiary control is limited for public services: the only option for beneficiaries to express

¹⁷A teacher was reported as absent if he could not be found at school at the time of a random unannounced check.

their discontentment is to vote with their feet, which might be costly.¹⁸ On the contrary, the existence of tuition fees in private schools makes the latter dependent on the beneficiaries' satisfaction to be viable, so that they should be highly responsive to beneficiary control. In summary, tuition fees increase the bargaining power of parents, which is low under a public system.

Then, one could expect that fee-charging private schools should adequately respond to the increase in demand for education while offering higher quality services, making the private provision of education a relevant policy option. Unfortunately, the multiplication of private schools has created a highly stratified educational system in many developing countries: private, fee-charging schools often fail to reach the most deprived fractions of the population, while free public schools and low-cost, low-quality private schools concentrate children from the poorest households. This would not be an issue if these schools were providing services of equal quality and if peer-effects were absent. However, as underlined above, public schools are likely to suffer from the traditional lack of efficiency of public service providers while peer effects are likely to be significant (Hoxby, 2000) [82], participating in what has been called the "learning crisis" (World Bank, 2017) [144], i.e. the fact that schooling does not result in adequate learning for a large number of children. This stratified system - both between the public and private sector and within the private sector - would then contribute even more to the already existing socio-economic inequalities in educational outcomes.

PPPs may seem to be a relevant alternative, likely to combine the efficiency of the private sector with the equity consideration of the public sector. In addition, as underlined by Patrinos et al. (2009) [121], the presence of affordable private schools, induced by PPPs, might bring public schools to increase the quality of their services as a response to the increased competition. In fact, the increased school choice created by PPPs may allow beneficiary control, despite the fact that most schools under a PPP contract are not allowed to charge fees, by increasing the number of free alternatives parents can use to vote with their feet. These features make beneficiary control both possible and likely to impact schools' quality. Of course, this potential benefit relies on the assumption that parents indeed value schooling quality and choose the school where they will send their children accordingly. However, quality might not be the only factor taken into account by parents when choosing a school. Actually, heterogeneity in preferences is likely to generate different degrees of demand-side pressure for quality improvement across schools. Indeed, Hastings et al. (2009) [77] show that higher-socioeconomic status (SES) parents are more likely to choose higher-performing schools compared to minority families, who might place less weight on schooling quality, so that

¹⁸Although they do so, since more and more children switch to private schools, despite the fact that they charge tuition.

schools serving the disadvantaged families will face little pressure to improve. School choice may thus “widen rather than narrow the gap in achievement”, causing greater educational inequality and stratification over time. Still, Patrinos et al. (2009) advance that PPPs offer to the government the possibility to define specific requirements in term of the quality of the service that will be offered by the private provider.

4.1 Enrollment

Subsidies and voucher systems are particularly likely to impact school enrollment as they give schools incentives to attract as many students as they can. In addition, voucher systems increase the supply of schools available for a non-negligible number of households by relaxing their budget constraint: they might thus be effective in making the school supply match the unmet demand for education. This outcome is particularly relevant in developing countries, where the achievement of the Millennium Development Goal (MDG) in universal primary education has increased the need of primary but also secondary schools, as more and more students are eligible for secondary education. Many governments have then taken steps to increase school supply, while sometimes making private schools more affordable for parents. Unfortunately, although the relevant outcome of interest is overall enrollment, it will become clear in this section that most studies reviewed here are only able to capture changes in enrollment in treated schools, due to migration issues.

This is the case of a policy evaluation taking place in Uganda, where the government has launched a program in which participating private schools are offered a per-student subsidy, as part of its strategy to absorb the demand for secondary education. Participating schools were not allowed to charge any additional fees to children from the public system. The policy was randomly phased-in, allowing Barrera-Osorio et al. (2016) [21] to compare participating private schools to private schools which applied to the program but entered it after a one-year delay. It has to be noted that the sample of schools selected for the study was not random: on the 250 private schools that applied to the program, 150 schools considered as being in critical location with respect to local demand for secondary education were directly enrolled in the PPP program. These 150 schools were thus excluded from the sample. Then, when studying enrollment rates, one should keep in mind that although the estimates are internally valid, they might be lower than if the randomization had included schools in areas with a large unmet demand. The authors find that average enrollment increased by 35% in participating compared to non-participating private schools and conclude that the initiative was successful in absorbing the increased demand for secondary schools.

However, while the authors underline that they only perform a partial-equilibrium analysis, their estimate is likely to include the impact from the potential migration of students from non-participating to participating private schools, in which case the positive impact on enrollment in participating schools would not translate in net enrollment gains, making the study inconclusive.

Similarly, Barrera-Osorio and Raju (2011) [22] study the impact of public subsidies to low-cost private schools in Pakistan. They exploit the fact that, in order to be eligible, private schools have to apply and achieve a minimum student pass-rate in a standardized academic test. They are then able to use a Regression Discontinuity Design, comparing outcomes of schools scoring just above the threshold to the ones scoring just below. They are able to conclude that receiving the subsidy increases the number of students enrolled by 85 students for schools that were just above the threshold, a 37% increase compared to baseline. However, these estimates should be interpreted with caution as it seems that test scores have been manipulated around the threshold. In addition, to the extent that schools participating in the program were not allowed to charge fees to students, it is possible again that the increase in enrollment mainly captures migration from students between non-program schools and program schools.

Angrist et al. (2002) [10] provide evidence from a quasi-experiment that should better deal with migration issues. From 1991, Colombia has used a voucher system with the objective of increasing secondary school enrollment for children of low income families in the country's biggest cities. The authors exploit the lottery used by cities to allocate vouchers when demand exceeded supply - with supply determined by municipality budget constraint and an estimation of the discrepancy between primary-school enrollment and public secondary schools supply - by comparing lottery winners and losers in order to evaluate the voucher system. The fact that the treatment status is at the individual level, as opposed to Barrera et al. (2016) [21] and Barrera-Osorio and Raju (2011), allows them to abstract from migration issues when studying enrollment: they will be able to recover net gains in enrollment between lottery winners and losers. The authors conclude that private school subsidies did not directly increase overall secondary school enrollment, suggesting that the household decision of whether sending their children to school or not is not price-sensitive. On the other hand, the authors demonstrate that the choice between public and private schools is depending on price, since lottery winners are more likely to attend a private school than lottery losers. From this, the authors advance that even if they do not detect any impact on enrollment between lottery losers and winners, the switch from public to private schools induced by the voucher system might have allowed public schools to accept students that would have been turned away due to overcrowding in the absence of

the program. This argument implies that migration between schools alone has the potential to help increase net enrollment in an overcrowded public system. However, this hypothesis is speculative and should be tested rigorously before drawing any conclusion.

Some interventions are directly aimed at providing incentives to private entrepreneurs so that they build more schools. Barrera-Osorio et al. (2013) [20] study such initiative in the Sindh province of Pakistan. According to them, while low-cost private schools have the potential to help the government achieve universal primary school enrollment by increasing the number of schools available, it has been limited to certain areas of the country. In others, like the Sindh province, such schools were constrained by the low educational attainment of the population, making them unable to find enough skilled but low-cost teachers. Hence, the government has decided to launch a program in which financial incentives are offered to private entrepreneurs for establishing and operating public schools, through the payment of a per-student subsidy.

This study has also a gender dimension since some schools were offered a higher subsidy for girls than for boys. A nice feature of the program is that it was clustered at the village level, which should limit migration issues between villages. Still, it is possible that children in control villages are sent to schools in treatment villages, resulting in an underestimation of the treatment effect. The authors report that the program was effective at increasing school enrollment: it increased average enrollment in treated villages by 30 percentage points compared to control villages. However, we have to underline the fact that, again, the counterfactual is not public schools but rather no additional schools in the village: it is thus impossible to conclude anything on whether these voucher schools were more effective at increasing enrollment than public school. Still, this study is highly informative on the potential for voucher schools to be used in order to increase school availability in previously neglected areas. An earlier study, evaluating a similar program, offering to private schools a subsidy per girl enrolled in order to incentivize them to construct new facilities in lower-income neighborhoods of Pakistan, reaches an analogous conclusion: the program increased both girls and boys' enrollment (Kim et al., 1999) [88]. However, due to the small size of the sample, the randomization failed to balance covariates between control and treatment neighborhoods, casting doubts on the robustness of these results. Through a difference-in-differences, controlling for constant confounding factors, the authors are still able to conclude that the program significantly increased girls enrollment in treated neighborhood by 28%, though the point estimate is of smaller magnitude than in the previous analysis.

Overall, the few studies available tend to confirm that PPPs in education have the potential to increase net

enrollment in developing countries. However, more evidence is needed as most of the studies reviewed in this section focus on the impact of PPPs on enrollment in treated schools rather than on overall enrollment, and fail to control for migration from non-participating to participating schools, which may bias the results. Future research should take steps to deal with migration issues: analyzing student enrollment decision or clustering randomization at a broader level than the school are solutions that should be explored.

4.2 Educational Outcomes

While enrollment rates increase in many countries, the true education emergency seems to lie in the so called learning crisis (World Bank, 2017) [144], i.e. the fact that even though children are attending schools, they are not learning. Public schools, lacking the appropriate incentives to be efficient, play an important role in this situation. PPPs could be a potential solution to this crisis as they allow to benefit from the efficiency of the private sector in a system that acknowledges the fact that education is a public good.

In addition, voucher systems increase the degree of competition on the school market - as they allow school choice - so that they are particularly likely to be effective at increasing the quality of education. One should however be careful when studying voucher systems as they particularly suffer from the “double-sided selection bias”, i.e., supply and demand side selection bias. Randomization at the school level, for example, might only solve the bias coming from the self-selection of schools in the voucher system. This is the case with the study from Barrera-Osorio et al. (2016) [21]. They exploit the fact that private schools willing to participate in the voucher system were randomly phased-in to get rid of part of the supply side selection bias. However, when comparing test scores in participating and non-participating private schools, they omit to mention that students’ characteristics in these schools might be very different. For example, budget and credit-constrained households particularly caring about their children education are likely to choose participating private schools, so best performing public schools students would then transfer to participating private schools, increasing their average test scores, while letting unchanged the student composition of non-participating private schools.¹⁹ Non-random selection is all the more likely to happen to the extent that schools are allowed to select the public students they accept. Due to this characteristic, the randomization does not even get rid of the full supply side selection bias. As expected, the authors find that students in participating private schools perform significantly better than students in non-participating private

¹⁹Indeed, one should remember that the policy did not target students already in the private system, so that it did not give students in non-participating private schools incentives to switch to participating private schools.

schools.

The lotteries used to attribute vouchers in case of excess supply is likely to provide more robust estimates. Indeed, the way the randomization is performed allows to get rid of both the supply and demand side selection biases: the selection of student by voucher schools has been made *ex ante*,²⁰ and only voucher applicants are considered for the lottery. In addition, individuals are directly compared rather than schools, so differences in, say, private and public schools location do not imply a selection bias but are only different channels when studying the impacts of *receiving a voucher* - which is the treatment unanimously chosen when evaluating the Colombian system. All these reasons explain why this system has extensively been studied.

Angrist et al. (2002) [10] studied educational outcomes three years after the start of the program. They compare lottery winners and lottery losers and find highly positive impacts of winning the lottery. For example, lottery winners completed 0.12-0.16 additional years of schooling and were 5-6 percentage points less likely to repeat a grade than lottery losers. Since the authors warn that education may not automatically imply learning, they use standardized tests in mathematics, reading and writing. They conclude that lottery winners scored 0.2 standard deviations higher than lottery losers, although this difference is only marginally significant. Unfortunately, the analysis suffers from a low power since only 60% of lottery participants agreed to take the test, potentially explaining the low significance of results. More importantly, the small share of students who participated in the test might reflect selection. Still, attrition rates between control and treatment groups are not significantly different so that the parameters estimated are internally valid. One should also note that compliance with “voucher status” is imperfect: some lottery losers actually receive some form of government scholarship, while some lottery winners do not actually receive any scholarship. The authors provide 2SLS estimates of the impact of receiving a scholarship, instrumenting the reception of a public scholarship with lottery status, in order to capture the impact of the voucher on the subset of scholarship users who would not have used one without the lottery.

Angrist et al. (2006) [11] use the same identification strategy in order to evaluate the long-term effects - after 7 years - of the Colombian voucher program. They find that voucher winners were 5-7 percentage points more likely to graduate from high school than voucher losers, although they only have an imperfect proxy for high school graduation, making the estimation less precise. This result implies that a smaller number of lottery losers took the college-entrance test. They thus have to correct for this when studying test scores, as directly comparing test scores of lottery winners and losers might suffer from a downward selection bias. Indeed, taking the test is likely to be

²⁰Voucher allocation is conditional on having been accepted to private school.

correlated with both voucher status and test scores if non-voucher students who expect to perform poorly are the ones not taking the test. The different strategies used in order to circumvent this issue - artificial censoring and nonparametric bounds - allow to conclude that the difference in test scores between lottery winners and losers is significantly positive in the long-run. In addition, Bettinger et al. (2010) [26] are able to disentangle the channels at play in the Colombian voucher system and conclude that vouchers have a real positive impact on the outcomes selected by Angrist et al. (2006) [11], outside potential peer effects. They reach this conclusion by studying a subpopulation of voucher applicants: applicants to vocational schools. They exploit the fact that voucher winners who had applied to vocational schools attended schools (vocational schools) with lower observable peer quality than voucher losers, who were more likely to attend an academic school. Thanks to this strategy, they are able to show that “despite having worse peers, voucher winners in this population had significantly better outcomes than voucher losers”. One of the potential channel for this positive impact is that the characteristics of the Colombian voucher system give incentives to students to be more serious in order to avoid repeating and thus, losing the voucher. In addition, the authors suggest that vocational private schools are more likely to adapt the skills they teach to students’ preferences and economic conditions, thanks to their flexibility.

Overall, although results from studies covering the Colombian program seem positive, it should be reminded that it was a very specific policy to the extent that it was implemented on a small scale and that the vouchers were renewable conditional on grade completion. This latter characteristic, implying incentives for students to perform better, could actually explain why voucher winners performed better as suggested by Bettinger et al. (2010) [26], while private schools might not necessarily perform better in the absence of this feature.

Voucher systems have also been used in large US cities in order to target low income households. Over-subscription was often handled through a lottery, getting rid of both supply- and demand-side selection biases as in the Colombian case. This is for example the case for the first federally funded private school voucher program in the country, the DC Opportunity Scholarship Program, launched in 2004. Wolf et al. (2010) [139] exploit the random allocation of vouchers in order to identify the impact of winning a voucher. They report that, although lottery winners were significantly more likely to graduate from high-school (by 12pps) after at least four years, they did not perform significantly better in term of reading and math test scores. Similarly, Rouse (1998) [128] exploits the random allocation of voucher among 2,300 applicants through a lottery to evaluate the Milwaukee Parental Choice Program, targeting poor households from 1990 in Wisconsin. She warns that the sample of lottery losers might not

be representative due to attrition, and that compliance with voucher status was not perfect. She finds that lottery winners only start scoring significantly higher in maths after 4 years into the program, while there is no statistically significant differences in reading. More recently, Abdulkadiroğlu et al. (2018) [2] use the same strategy in order to evaluate the Louisiana Scholarship Program, providing private school vouchers for disadvantaged students attending low-performing public schools. Instrumenting the use of the voucher with lottery status, they find highly negative results: among the 1,412 first-time applicants in grades three through eight, participation in the program lowers test scores in math (0.41 s.d.), reading (0.08 s.d.), science (0.26 s.d.) and social studies (0.33 s.d.) for compliers, after one year. These results are robust to controlling for differential attrition rates using trimming and bounds. Mills and Wolf (2017) [104] extend the previous analysis to two years after the start of the program. They argue that results are more encouraging in the second year of the program, although estimates are still negative.

If the voucher programs studied above were restrictive in the way they were allocating vouchers, there also exist some *universal* voucher system, in which every student is eligible for a voucher, irrespective of parental income. This is for example the case in Chile, where the government decided to offer a per-student subsidy to both public and voucher schools from 1981. The impacts of such universal voucher system on educational outcomes can only be identified through weaker strategies than the previous randomization, as constructing a control group is made extremely challenging. Still, the case of Chile has been extensively studied. In this review, we will only select the strongest identification strategies, while underlining how their results differ from studies using cross sectional data.

Lara et al. (2011) [92] summarize that “most studies using cross-sectional individual level data have found that students attending private voucher schools have higher educational outcomes than those from public schools; the estimated impact is typically between 0.15 and 0.2 standard deviations”. They underline that most of these results are based on a Heckman two-stage approach in order to control for non-random selection in private schools, often using school density as an instrument. However, they argue that since school density may actually reflect unobserved characteristics impacting educational outcomes, such approaches are likely to provide invalid estimates. Similarly, propensity score matching (Anand, Mizala and Repetto, 2009 [5]) is unlikely to account for the full set of potential confounders. Lara et al. (2011) [92] exploit the fact that they gain access to the previous students’ test scores: they show that they can replicate the positive and highly significant impact of private schools on student test scores (test score gain of 13% and 14% of one standard deviation for language and mathematics, respectively), but that these estimates become insignificant once they control for previous test scores. This suggests that most

of the previous results suffered from an upward bias, as more able students seem to select into private voucher schools. They also use a novel strategy which is likely to avoid at least part of the demand side selection bias. They exploit the fact that most schools provide either primary or secondary education, implying that students have to switch school. They restrict their sample to eighth-grade students in public schools that have to switch school: the treatment group is composed of students who switch to private schools while the control group is composed of students who switch to another public school. They assume that such approach is likely to make the two groups more similar. Using both propensity score matching and changes-in-changes, they find that private voucher education has a small and sometimes insignificant impact on test scores, between 4% to 6% of a standard deviation. Still, their identification strategy might fail to account for confounders determining the selection into different school types at the end of primary school.

Hsieh and Urquiola (2006) [84] circumvent the issue of student selection by evaluating the impact of school choice on education instead of the impact of *attending a private school relatively to a public school*. They use municipal panel data, exploiting the fact that private schools developed more in urban and wealthier municipalities, in order to implement a difference-in-differences: under the assumption that this differential impact is driven by characteristics that are fixed over time, they can compare the “change in educational outcomes in urban and wealthier communities to that in communities where private schooling increased by less”. Using this strategy, they find that increased school choice fails to impact test scores, while it has a negative impact on average repetition rates and grade-for-age measures, even when controlling for pre-existing trends. However, as with any difference-in-differences estimates, their validity is conditional on whether the parallel trend assumption holds, i.e. on whether the outcomes of interest would have evolved similarly in the two groups in the absence of the program. In addition, they also complement the difference-in-differences approach with different instruments for private school growth and show that, “if anything, the OLS estimates overstate the impact of the voucher program.” They also provide more anecdotal evidence that Chilean students’ performance in international test did not improve between 1970 and 1999. We will see later that the authors explain this poor impact of school choice by the stratification that resulted from the creation of the voucher system in Chile.

Finally, Aguirre (2017) [3] takes advantage of a recent reform of the voucher system which aims at preventing segregation of students based on their socio-economic status. In 2008, the government increased the voucher school subsidy by 50% for students in the bottom 40% of the income distribution, giving more incentives to private schools

to enroll them. The reform further forbade voucher schools to use screening - although the author notes that this was not implemented in practice - and to charge add-ons to parents eligible for these targeted vouchers. Through this, Aguirre (2017) can identify the impact of being eligible for a targeted voucher on educational outcomes, exploiting the discontinuity in eligibility with respect to income levels created by the reform through a RDD. Results are again disappointing: any positive impact above 0.04 standard deviation on math and language test score can be rejected. We will come back to the potential explanations when analyzing evidence over socio-economic segregation.

We now turn to the evidence regarding the private management of public schools, encompassing charter schools - not allowed to charge tuition or screen students - which have extensively been studied due to their popularity in the US. Again, such PPP offer a nice way to identify the impact of charter schools as oversubscribed schools use a lottery to determine which students will be enrolled. Most authors choose to instrument enrollment in private school with lottery status, allowing to estimate the impact of being enrolled in a private school, even though the validity of these estimates is restricted to lottery compliers. Epple et al (2015) [56], in their review of the literature over charter schools, find at least 7 papers that fall into this category. Overall, they observe that attending a private school seems to have positive and significant impacts on test scores in math and language, although there are differences in magnitude. For example, Hoxby and Rockoff (2003) [81] find that attending a charter school in Chicago significantly improves math - between 6 and 7 pps - and reading - between 5 and 6 pps - test scores of students who apply to kindergarten through grade five, using a sample of 3,407 students. In New York, Hoxby et al. (2009) [83] analyze the impact of being enrolled in grades 3 to 12 of a New York City charter school between 2000 and 2008. The authors underline that since 94% of charter school students have been admitted after having participated in a lottery, lotteries are not restricted to a subset of high performing charter schools and the external validity of the study is relatively high. Their results are also highly positive since they find that “by the time a charter school student has reached the end of eighth grade, he will be scoring about 30 points higher in math than he would have been scoring if he had been lotteried-out and remained in the regular public schools”, almost enough to close 86% of the average gap in test scores between one of the most affluent suburbs of New York and Harlem, where many charter schools are located. Attending a charter school closes 66% of the same gap in language. Abdulkadiroğlu et al. (2009) [1] study the effectiveness of Boston charter schools. They find that language and math scores increase by, respectively, 0.25 and 0.42 standard deviation for each year in a charter middle school. Results for attending a charter high school are similar for math but less robust for language. They are also able

to evaluate the effectiveness of pilot schools, which have less flexibility in term of teachers' contracts: attending a pilot school has not significant impact on test scores. This results seem externally valid since Angrist et al. (2010) [13] underline that they find results remarkably similar to Abdulkadiroğlu et al. (2009) [1] when studying a sample of 856 students who applied to a KIPP charter school in Massachusetts, with effects on the order of 0.35 standard deviation for math and 0.12 standard deviation for language. Positive evidence accumulate as Dobbie and Fryer (2011) [49] find that attending a charter school in Harlem from sixth through eighth grade increases test score by 0.687 standard deviation in math and 0.141 standard deviation in language, relatively to attending a public school. In Washington D.C., attending a charter school increases yearly achievement by 0.211 standard deviation in reading and 0.229 standard deviation in math (Curto and Fryer, 2014) [46]. Although these results are less informative as they do not instrument charter school enrollment with lottery status, Wong et al. (2014) [140] show that being offered to a high-performing charter high school in poor neighborhoods of Los Angeles significantly improves math and language test scores.

One study, though, contrasts with the rest of the literature. Gleason et al. (2010) [70] take advantage of a sample of lottery outcomes of 36 middle charter schools across 15 US states. They do not detect any significant impact of being enrolled in a charter school on test scores. However, attending a charter school seems to have heterogeneous effects according to the socio-economic status of students: the authors find that studying in a charter school “had positive effects in mathematics for more economically disadvantaged students and negative effects in both reading and mathematics for more economically advantaged students”. Still, the majority of results from studies on charter schools reviewed here are largely positive. Moreover, Angrist et al. (2016) [12] recently provided evidence that charter schools can have many long-lasting positive impacts: relatively to students attending a public school, students enrolled in a charter high school have higher scores at high school exit exams and university entrance exams (SAT), are more likely to qualify for a state-sponsored scholarship but also to attend a 4-year than a 2-year college. Overall, charter schools seem to have highly positive impacts on educational outcomes, at least in the USA.

Interestingly, one should note that the broad consensus achieved through randomization-based studies greatly differ from the results achieved using difference-in-differences. Such studies take advantage of students switching between a public and a charter school to create a treatment group, while controlling for constant differences with non-switchers over time. Estimates resulting from such approach might be biased if confounding factors affect both the switch and educational outcomes. In addition, as underlined by Epple et al. (2015) [56], these results provide

an answer to a very different question: do individual test scores increase when a student switches from a public to a voucher school?²¹ Evidence from such fixed effects strategies are often non significant, or even negative, for both language and math test scores (Bifulco and Ladd, 2006 [29]; Booker et al., 2007 [33]; Zimmer et al., 2012 [147]).

Finally, tying private operation and infrastructure provision also has the potential to improve educational outcomes, as demonstrated by the previously mentioned RCT in Pakistan. Barrera-Osorio (2013) [20] indeed estimate that test scores increase by 0.67 standard deviations in treatment villages, where new privately operated public schools have been established, relative to control villages. In addition, instrumenting student enrollment with treatment status, the authors conclude that children enrolled thanks to the program score 2 standard deviations higher than children in control villages.

Overall, the impact of PPPs on educational outcomes seems to differ according to their location, coverage and to the type of contracts used (voucher, charter, subsidies). For example, even if it is likely to depend on the institutional details, charter schools in the US seem effective at increasing test scores, as opposed to voucher schools. On the contrary, voucher schools in Colombia seem to have positive impacts on educational outcomes, which is not the case of the universal voucher system implemented in Chile. This latter case illustrates that the scale of the program and the incentives to students embodied in it matter. Additional evidence would be needed in order to reach more robust conclusions. Moreover, the mechanisms through which PPPs in education impact educational achievement should be studied more thoroughly in order to improve their design. It is also important to recognize that positive impacts from PPPs in education might come with downsides, in particular if these improvements stem from an increased segregation, as argued in the following section.

4.3 School Stratification

Previous studies provide evidence on the impact of attending a voucher/charter school on individual test scores. However, it is unclear if this impact, whether it is positive or negative, comes from a higher quality of such schools, positive externalities among students (peer effects) or a mixture of both. Disentangling these two channels would be highly valuable as it has important implications for the aggregate impact of the system on educational outcomes. For example, if most of the increase in test scores for charter schools is due to peer effects, the co-existence of public and charter schools might have detrimental consequences for students in the former: the fact that charter schools attract the best performing students, resulting in a high quality class composition, implies that public

²¹As opposed to the question of whether average test scores differ in public versus voucher school.

schools' student composition has deteriorated. Then, the coexistence between schools under a PPP contract and public schools might result in a dual system. This might contribute to maintain pre-existing inequalities in the long-run (Benabou, 1994) [38]. From this, understanding the impact of educational systems undertaking PPPs on socio-economic segregation is thus crucial before making any conclusion on their overall effectiveness, and should also inform the way these contracts are designed. We review here studies tackling such issues. We also include studies covering ethnic segregation across schools, as it is as likely to have long-term detrimental consequences on inequalities.

One of the most studied systems in term of socio-economic segregation is the Chilean universal voucher system. As underlined by Epple and al. (2017) [58], one of the feature of this system making it particularly subject to cream skinning²² is the fact that voucher schools screen applicants based on admission exams and/or interviews with students or parents. In addition, voucher schools have been allowed to charge tuition fees on top of the amount covered by the voucher, inducing a selection of students based on their socio-economic background. Actually, the authors note that this outcome was predicted by theoretical models including characteristics such as screening and tuition add-ons (Epple and Romano, 1998 [57]; MacLeod and Urquiola, 2009 [96]). For example, Epple and Romano (1998) [57] conclude that if schools are able to price discriminate on ability, in equilibrium sorting will be made along two dimensions: low-ability, high-income students will select in voucher schools by paying a higher tuition premium, while high-ability, low-income students will benefit from the voucher. In addition, MacLeod and Urquiola (2009) [96] advance that when schools are allowed to screen students on their ability, increased school competition will lead to socio-economic segregation.

Sorting coupled with peer effects may actually be one of the explanations for the absence of improvement regarding Chilean students' performance in international tests. As mentioned above, Hsieh and Urquiola (2006) [84] finding that an increased availability of private schools at the municipality level does not translate into higher test scores may mean that the gains in voucher schools are fully offset by the losses in public schools. A simple OLS regression shows that the educational attainment and income of parents in public schools are lower in municipalities with higher private enrollment: "a one standard deviation increase in the private enrollment rate is associated with 38–43% of a standard deviation decline in the relative income of public school parents". They also underline that it might have been possible that public schools improved thanks to the increased competition, while loosing their

²²The fact that the most privileged students and/or the ones with the highest ability migrate from public to schools under a PPP contract.

best performing students, two effects particularly difficult to disentangle. However, they show that relative grades of public school students worsened in communes with a larger increase in private enrollment, so that any positive impact of competition on public schools is more than offset by cream-skimming. Overall, they conclude that “the central effect of the school voucher program in Chile appears to have been to facilitate the exodus of the Chilean middle class from public schools, without much evidence that it has improved aggregate academic outcomes.”

Still, despite the lack of rigorous research, a broad consensus has emerged in Chile on the detrimental impact of the voucher system regarding socio-economic segregation. This has led to the 2008 reform, increasing vouchers by 50% for students from households in the bottom 40% of the income distribution. In addition, schools accepting these targeted vouchers were also forbidden to screen students - though this was apparently not implemented in practice - and to charge add-ons tuition to eligible students. The combination of a voucher independent of income with students screening and the possibility to charge additional fees were often cited as one of the main cause of stratification in Chile (Mizala and Torche, 2012) [105]. Thus, this reform was expected to decrease significantly the socio-economic stratification created by the voucher system, and to increase the pool of schools available and affordable for low-income children. Several studies have evaluated this reform.

For example, Navarro-Palau (2017) [112], uses two sources of variation - namely the timing of the reform and birth cutoff determining the timing of school entry - in order to identify the impact of the reform. She finds that “even though almost 50% of the population was eligible, the probability of enrolling in a public school only fell in around 1.6%”. Moreover, analyzing the characteristics of students who switched from public to private schools after the introduction of the reform, she shows that compliers are not the poorest students in the population, which might illustrate both cream-skimming by private schools and imperfect information, particularly for the poorest families. Overall, the socioeconomic composition of a student’s peers remained unchanged after the introduction of the reform. Although it does not decrease stratification, Navarro-Palau (2017) concludes that the reform may still have had some positive impact as she finds that test scores increased for students more likely to enroll in public schools. The author suggests that this comes from the public schools’ response to the increased competition faced after the reform. Aguirre (2017) [3] reaches similar results with respect to the impact on students’ choice by using a RDD, comparing students just above the income threshold to those just below. Being eligible for a targeted voucher does not impact the type of school chosen, as eligible students are not more likely to choose a private school, a school with higher test scores, a higher average socioeconomic status or a lower class size. The author suggests different

mechanisms that could explain these results. First, parents might not have been able to choose higher quality schools because only a subset of private schools agreed to participate in the targeted voucher program. In particular, among the schools that charged add-ons, only 50% joined the initiative. Still, the reform created important incentives for at least some private schools to accept lower-income students and it is surprising to observe that parents did not take advantage of their increased school choice. Demand side mechanisms could then be at play. Aguirre (2017) [3] advances that low-income households might suffer from a number of barriers preventing them to make the optimal school choice: a lack of information over school quality, the complexity to evaluate a large number of options, or the fact that parents simply choose the school in which their social group is over-represented. Then, even a promising reform seems unable to solve the socio-economic stratification of the Chilean voucher system. Estimating a structural model, Neilson (2017) [113] also produces consistent findings. Indeed, while the descriptive evidence he presents suggest that the gap in test scores between poor and rich students decreased after the implementation of the targeted voucher policy, he notes that this “increase in academic achievement was not the result of a massive reshuffling of students to more productive schools but driven by schools improving their quality”. Then, while the reform of the Chilean voucher might have implied a reduction of the inequalities in educational attainment, the evidence at hand indicate that this change was not due to weaker stratification.

Interestingly, Gazmuri (2017) [68] explains that the reform might have created room for more stratification. Using a student-level panel data (2005-2012), she finds that the probability of enrolling in a private school increased by 10 percentage points for low-SES students one year after the reform; leaving only the most deprived students in public schools. She suggests that it illustrates a certain decrease in cream-skimming based on socio-economic characteristics from voucher schools participating in the reform: these schools suddenly accepted students of lower SES thanks to the financial incentives they faced. However, she also shows that, following the increased probability that low-SES students enroll in private schools, more advantaged students were more likely to migrate to a private school that was not accepting targeted vouchers, because of high-SES households’ preferences toward high-SES peers for their children. As summed up by the author, “the program seems to have mainly caused a redistribution of the most vulnerable students between some private schools and public schools. Moreover, it kept higher-income students in the non-SEP²³ private subsidized schools and the most vulnerable students in public schools”.

Fortunately, this feature of the Chilean system does not seem to be universal. For example, the Swedish system appears to be less impacted by socio-economic segregation. Epple et al. (2017) [59] note that this might

²³Voucher schools that chose to do not accept targeted vouchers.

simply reflects adequacy of the rules of the program, designed to prevent such stratification. Indeed, all schools are prevented from screening or charging tuition add-ons. Böhlmark et al. (2015) [39] provide some evidence that the Swedish universal voucher system implemented in 1992 did increase segregation along both socio-economic and ethnic dimensions, even though they note that the magnitudes of their results is low. They implement a strategy similar to Hsieh and Urquiola (2006) [84], namely the comparison of municipalities with an important expansion of school choice with the ones where it was limited, coupled with instrumental variables. This approach allows to control for any time-invariant characteristics that might be correlated with the growth of private schools. They conclude that municipalities that benefited from an increased availability of schools are associated with a higher level of segregation between immigrants and natives, and between children of high/low educated parents. The authors warn that their results might not be interpretable as causal: for example, they show that the parallel trend assumption did not hold in the years preceding the introduction of the voucher with respect to income segregation, although it holds with respect to segregation between immigrants and natives.

A last possibility is that the *universality* of the program explains sorting. Then, targeted voucher schools might be more likely to circumvent this issue. In particular, since they are not allowed to screen students: they have to implement a lottery in case of over-subscription.²⁴ Chakrabarti (2013) [40] compare students eligible for a voucher as part of the already mentioned Milwaukee program, implemented from 1990. The author first provides a theoretical model concluding that the prohibition of both sorting and tuition adds-on might prevent sorting by income at the application stage, but not by ability. These predictions are backed up by the data. Analyzing applicants' characteristics from the first five years of the program, she finds that, out of the students eligible for a voucher, applicants do not come from wealthier households than non-applicants, but seem to have a higher ability, proxied by mother's education, time spent with the child in activities such as reading, math, writing, sports, educational expectations for the child, or proportion of parents participating in various parent-teacher activities. This sorting is based on parental preferences. However, further sorting based on ability does not appear when comparing successful applicants according to whether they choose to enroll in the voucher school or not.

However, the design of targeted program highly matters when one hopes to decrease stratification. Studying the India's Right to Education (RTE) Act of 2009, through which the government blocks 25% of the places in entry grades of all private schools for children from disadvantaged households while covering their tuition fees,

²⁴In such cases, two different sources of stratification still remain present. First, voucher schools might locate in very specific neighborhood. Second, students applying to voucher schools might have very different characteristics than those who do not.

Damera (2017) [47] shows that the absence of differences in test scores of lottery winning and losing children after 1.5 years can be explained by a mistargeting of the policy, implied by its poor design. The author shows that policy applicants were default private school goers as “99 percent of the treatments and 93 percent of the controls are enrolled in private schools”. He first explains that making eligibility contingent on income enabled ineligible households to participate in the program as income determination is particularly difficult in India. In addition, he advances that the important non-tuition costs, which are estimated to represent 1.3 times the tuition fees, simply deterred poor households to apply to the program since it did not cover these costs. Overall, this specific program failed to decrease segregation to the extent that “the poor [were] not induced to participate and the non-poor [were] not effectively prevented from participating”.

Regarding charter schools, they might resemble (well-designed) targeted voucher systems with respect to sorting. They are not allowed to screen students and do not charge tuition adds-on as they are (privately managed) public schools. However, the possibility of stratification from charter schools location choices and parental preferences remains. As underlined by Epple et al. (2015) [56], the main source of segregation comes from the fact that charter schools are concentrated in urban areas, where the minority and low-income households are disproportionately represented. They also note that households’ preferences reinforce such segregation. To disentangle the latter effect from the former, they track students who switched from a public school to a charter school thanks to longitudinal student-level data. This allows to capture the impact of individual selection as, conditional on the presence of a charter school, one can check whether switchers enroll in schools with a larger share of student of the same socio-economic background and/or ethnicity. Of course, estimates from such strategy are only valid for the population of switchers, so that their external validity is limited.

Booker et al. (2005) [34] implement the empirical approach described above using data on school switchers in Texas and California. In both states, they observe that black students are more likely to switch to charter schools with a larger share of black students and a higher degree of racial concentration relatively to their preceding public school. As noted by the authors, these results are consistent with the ones of Bilfulco and Ladd (2007) [30] who use the same technique in North Carolina. They also show that black students switch to schools with lower achieving peers than in their preceding school, while it is the contrary for white students. They conclude that “asymmetric preferences of black and white charter school students (and their families) for schools of different racial compositions help to explain why there are so few racially balanced charter schools.” Zimmer et al. (2009) [146], exploiting school

switchers data from 8 US States, also find that African American students transferring to charters schools are more likely to end up in schools with a higher share of black students. In addition to these evidence of ethnic stratification, Booker et al. (2005) [34] also show that Texan switchers have lower test scores than other students in the public school they leave. In fact, in both states, no evidence of cream-skimming by charter schools is found, while it is even the opposite in Texas. Zimmer et al. (2009) [146] similarly conclude against a systematic cream-skimming by charter schools. These results might reflect the *targeted* dimension of charter schools in the US, which might make some parents reluctant to send their children to such schools.

Overall, it seems that PPPs often increase the degree of socio-economic and ethnic segregation in the educational system, failing to solve the main issue created by the co-existence of public and independent private schools. Stratification actually seems to emerge when different types of schools coexist within a system, which might sound obvious as it creates more room from selection on both the supply and demand sides. Targeted systems might actually increase segregation even more as they institutionalize the fact that different schools will be directed at different types of students. For example, the Chilean reform increased stratification across voucher schools, as half of them refused to accept targeted vouchers, signaling they were targeting wealthier students. On an optimistic note, as noted by Kremer and Holla (2009) [90], socio-economic stratification might not be detrimental to educational achievement: Duflo et al. (2011) [51] show that tracking primary school students by initial achievement is beneficial as it allows teachers to adapt their teaching to the level of their students. Still, it is unclear whether such adaptation is systematically undertaken within PPPs educational systems nor whether it always overcome the absence of high ability peers for lower performing students. In addition, the income and ethnic inequalities perpetuated by such systems are anyway likely to result in a highly polarized society (Kremer and Sarychev, 2000) [91].

4.4 Cost-Effectiveness

Very few studies evaluate the impact of PPPs on the cost of education. To the best of our knowledge, only Angrist et al. (2006) [11], studying the targeted voucher system in Colombia, provide estimates of the impact of such a system on households and government's expenditures. They show that lottery winners faced a \$52 dollar increase in educational expenditure compared to lottery losers. This increase comes from expenditures associated to private school enrollment - as winners have a higher probability of enrolling in private schools - but also from the fact that some students who would have attended a private school anyway chose a more expensive school when winning

the voucher. In addition, voucher winners were less likely to work while studying: the authors estimate that they earned \$41 less than lottery losers. Then, the voucher actually increased households' expenditures by \$93. Since they received \$74 more through the voucher than lottery losers, households spend \$19 more than lottery losers. The authors also estimate that the voucher increased the government's expenditures by 24\$ per lottery winners, as the voucher they provide is larger than the public costs saved when students migrate from public schools. However, the authors argue that, to the extent that lottery winners completed more schooling and scored approximately 0.2 standard deviations higher, the benefits for participants outweigh the cost of the program. Still, they underline that the impact of the program on other students should be taken into account. If such voucher programs disproportionately affect ethnic and socio-economic segregation, its cost would be higher. Unfortunately, such impact is very difficult to quantify, so it is possible that vouchers are cost effective for lottery winners, but not for the universe of students.

5 PPPs effectiveness in Health

In this section, we review the empirical evidence over the impacts of PPPs on health services use, health-related outcomes, stratification, and cost-effectiveness. One should note that the literature covering PPPs in the health-care sector is very scarce and of relatively poor quality. We will sometimes refer to less rigorous studies but we will take care to underline their limitations.

The most developed contribution seems to cover voucher schemes and pay-for-performance programs, that have mainly been implemented in developing countries. Such policies often - but not always - target both public and private health providers. There are three main limitations with studies evaluating these schemes.

First, they are not always making clear whether the program is directed at public, private, or both types of health care providers. We thus have to exclude papers, which do not mention whether the policy in question involves the private sector. Second, it seems that many health care programs, such as vouchers, have been implemented through a foreign aid initiative, which would not fall under the definition of PPPs in health care. Again, we will exclude studies that do not explicitly state that the government was involved in implementing the project. Finally, the remaining evaluations, when they include both public and private facilities, often do not disaggregate results by type of providers. Still, as long as they are opened to private health care providers, these policies should be understood as PPPs as they imply the collaboration of the public sector with some private entities.

Ideally, we would like to disentangle the channels of a change in health-related outcomes, saying for example, whether an aggregate health improvement after the introduction of a voucher comes from the fact that more individuals can then afford private health care, which might be of greater quality, or because competition increases the overall quality of care. However, since the desirability of an health care system should be determined by its aggregated impact, we will also consider such aggregate evidence on health PPPs.

The change in scope of PPPs in the health sector comes from the fact that, similarly to the education sector, the low quality and unequal coverage of public health services has created room for private health-care providers in developing countries, making it increasingly possible for governments to take advantage of the private sector, with its potentially greater efficiency and flexibility.

Banerjee et al. (2004) [16], for example, describe the quality of public health services in a poor rural area of the state of Rajasthan in India as “abysmal”. Indeed, the public sector largely fails to provide the proper incentives, to limit issues such as moral hazard from health workers. The issue is even more striking than in the education

sector: averaging absenteeism rates across Bangladesh, Ecuador, India, Indonesia, Peru, and Uganda, Chaudhury et al. (2006) [41] find that 35% of government health workers were absent at the time they checked. These large absenteeism rates might in part be explained by the fact that a large fraction of public health workers also work in private facilities. In addition, absenteeism is higher in remote places, where the only health facilities are often primary health centers, which are less attractive and in where no doctors are present so that there is no threat of being monitored by a superior. Actually, Banerjee and Duflo (2006) [17] note that a single health worker is often running rural health centers, meaning that his absence is equivalent to a facility being closed. This is the reason why most rural inhabitants simply do not use government health facilities. Banerjee et al. (2004) [16] estimate that 57% of visits are made to private providers in rural Rajasthan. More generally, the poor quality of the public sector has created a large market for private health providers in developing countries. In India, for example, the private sector account for about 80% of outpatient care and 60% of inpatient care (Ministry of Health and Family Welfare, 2014) [117] while, in Sub-Saharan Africa, the private sector provides 50% of health care (World Bank, 2008) [142].

However, although private operators are more likely to provide the right incentives and monitoring, it does not systematically translate in high quality health care services. For example, in rural Rajasthan, 41% of private doctors do not have a medical degree and 17% have not even graduated from high school (Banerjee et al., 2004) [16]. Similarly, in Uganda, private health workers' knowledge over first-line management of malaria, pneumonia and diarrhoea is significantly lower than in public facilities (Buregyeya et al., 2017) [37]. In fact, the absence of proper regulation of private providers by the public sector creates important discrepancies in quality and access to health care (Health Research Institute, 2010) [79]. In addition, the multiplication of these private facilities might deteriorate even more the quality of public services as they are likely to attract the most skilled health workers. PPPs offer a way for the government to regulate and monitor the private provision of health services, by contracting over defined quality and coverage standards, while benefiting from its efficiency.²⁵

5.1 Health Services Utilization

There exists a large number of preventive health care services that would be effective at decreasing health risks, but that are not used by patients for various reasons: they might not know about it, their beliefs/customs might forbid the use of such services, or they simply cannot afford them. Such services encompass reproductive and sexual

²⁵Of course, the government might also try to provide the relevant incentives to increase the efficiency of the public sector. However, some evidence show that governments might not be able to commit to such an approach: for example, a promising experiment aiming at increasing government health workers attendance in India through financial penalties was made ineffective by the local health administration by creating many "exemptions" to the scheme (Banerjee et al. 2008) [18].

health care, maternal and child health care, or preventive goods such as insecticide-treated bed nets for preventing malaria. Demand-side interventions have been recommended in order to increase the use of these goods and services. In developing countries, vouchers have been one way for governments to try to remedy this insufficient utilization: they are distributed to a specific population, and give access to free or subsidized health goods or services, that are then reimbursed to participating providers by the relevant governmental agency. According to Brody et al. (2013) [36], vouchers both relax budget constraints that might prevent individuals from using a particular health service and give incentives to health providers to increase its supply: these two effects should result in an increased utilization of the targeted health care services. In addition, vouchers can lower stratification by allowing even the most deprived to choose their health care provider (Arur et al., 2009) [14].

A nice feature of voucher programs is that they can specifically be directed to the at-risk and under-served population. Nicaragua, facing high rates of unwanted pregnancies and illegal abortions, has used such a scheme in order to increase access and use of reproductive health care by poor adolescents in its capital city, Managua: between 2000 and 2001, 28,771 vouchers were distributed to male and female adolescents in poor neighborhoods of the city, giving access to “one consultation and one follow-up visit for counseling, family planning, pregnancy testing, antenatal care, sexually transmitted infections (STI) treatment, or a combination of these services” (Meuwissen et al., 2006) [103]. Participating clinics were selected through a competitive process, opened to both private, NGO, and public clinics. In order to evaluate the impact of the program on the use of reproductive health care services, Meuwissen et al. (2006) [103] exploit the fact that girls present at one of the distribution locations (markets, outside public schools, clinics) have both a high (but inferior to one) probability of having received a voucher and should have similar characteristics: they randomly administer individual questionnaires to female adolescents 3 to 15 months after the vouchers have been distributed, at the locations where most of the vouchers were given. Survey respondents are then considered as being girls who had the same probability of receiving a voucher. The authors advance that the fact that some had received a voucher while some did not was actually random. However, recipients and non-recipients display differences in observable characteristics, suggesting that they are not comparable. The authors argue that once they adjust for differences across survey sites (schools, markets, neighborhoods), the reception of a voucher was correlated with a significantly higher use of sexual and reproductive health care services. Sandiford et al. (2002) [130] use this data to conclude that, thanks to the program, 40% of the sex workers in the capital city have benefited from STI treatment. However, it is impossible to attribute this utilization rate to the voucher

program as we do not know whether, in the absence of this policy, the share of sex workers using STI treatment would have been lower.

In 2006, the government of Kenya, in collaboration with the German Development Bank created a voucher scheme in order to increase access to family planning services in three districts and two slums of Nairobi. The program offered the possibility for women below a given poverty threshold to buy a \$1.25 voucher that could be used in participating public and private health facilities in exchange of “long-acting and permanent methods” (LAPM) of contraception (sterilization, IUCD or implants). Surprisingly, the results of this program have been disappointing: according to Arur et al. (2009) [14], only 41% of the vouchers sold were actually used. This might come either from the fact that the low uptake of reproductive health care is not primarily due to budget constraints, or from other barriers such as difficulties in accessing participating health facilities. In the absence of randomization, Obare et al. (2013) [116] use a comparison group to evaluate whether exposure to the program increases the use of family planning services. They consider that women living within a 5-kilometer radius of a facility participating in the program since 2006 were exposed to the voucher program, compared to women living in non-vouchers site (*never exposed*) or within a 5-kilometer radius of a facility that entered the program the program in 2010 (*not exposed from 2006 to 2010*). A survey of 2,527 women living in those districts first reveals that there is no difference across groups in awareness of family planning. In addition, women living in neighborhood exposed to the program were not significantly more likely to have used LAPM compared to those never exposed to the voucher scheme. The authors advance that these findings may reflect a lack of demand for these specific contraceptive methods, or a failure in the promotion of the program since only 25% of the women exposed to the program since 2006 were aware of its existence.

Taken together, the Kenyan experience suggests several conditions for the success of a voucher scheme. First, budget constraints need to be the primary reason for the lack of use of the targeted health care services (Arur et al., 2009) [14]. Second, the schemes should encompass other potential barriers to access, such as transportation costs, and finally, voucher programs should be accompanied by an adequate communication campaign (Obare et al., 2013) [116]. The Kenyan program did not meet at least the first two conditions. However, results should be interpreted with caution, as there are significant differences in observable characteristics between the different groups compared. In addition, the definition of “exposure to the program” is arbitrary and could underestimate the impact of the program: it is possible that either some women not living within the 5-kilometer radius were actually using the

voucher or that only women living closer to the facility were actually exposed to the program.

Voucher schemes have also been extensively used in order to increase access to maternal health care services. Such programs are directly related to the target of “reducing the global maternal mortality ratio to less than 70 per 100,000 live births” by 2030, defined by the SDG. Antenatal care and institutional deliveries are key in reaching this goal, but their use is still far from being universal. Again, the issue seems often more related to demand-side factors (Bellows, 2012) [24] - such as lack of knowledge, inability to pay for these services, or cultural barriers - that vouchers have the potential to influence. In parallel to its reproductive health voucher program, Kenya has implemented a similar scheme for maternal health in the same three districts and two slums of Nairobi. It is especially relevant in a country where 56% of births are delivered at home (Kenya National Bureau of Statistics, 2010) [118]. Women under a defined poverty threshold were eligible to buy a 2.50\$ voucher giving access to a package of antenatal care, institutional delivery and post-natal care in both selected public and private facilities at a subsidized price.

Obare et al. (2013) [116] use the same sample of 2,527 women to compare the use of maternal health care and share of institutional deliveries between women considered as being exposed to the voucher program and women who were not.²⁶ The authors find no significant differences in the share of women who made four or more antenatal visits, nor in the share of women who visited a facility in the first semester of their pregnancy, with respect to their exposure status. However, the share of institutional delivery (as opposed to home delivery) were significantly higher in sub-locations exposed to the voucher program since 2006 compared to the other areas. The authors estimate that women exposed to the scheme were twice as likely to give birth in a health facility or under skilled care than the ones living in non-vouchers areas. Similarly, the voucher scheme also seems to have a positive impact on the use of postnatal care, as the share of women who use such services is significantly higher among women exposed to the program. In a subsequent analysis, exploiting two rounds of survey, Obare et al. (2014) [115] confirm these results by showing that the share of institutional deliveries significantly increased over time in voucher-exposed communities only. Interestingly, although this change occurred in both public and private facilities, the increase was more important in the latter, which might reflect the fact that women who could not afford health care in private facilities before the program used the voucher to do so. Unfortunately, Obare et al. (2013) [116] show that the program failed to close the gap between poor and non-poor women in use of maternal health care services: “poor women were, however, significantly less likely to have delivered at a health facility, have been assisted by a skilled provider during delivery, or have received postnatal care services compared with their non-poor counterparts

²⁶These results should be assessed keeping in mind the methodological caveats detailed above.

regardless of exposure to the programme”. This suggests that the voucher program fails to tackle all the dimensions constraining women access to such services.

As noted by Obare et al. (2013) [116], Bangladesh might have applied a more suitable maternal health care voucher program by taking care of transportation costs for voucher users, while making sure that selected facilities are close enough to potential beneficiaries. The scheme gives access to free health services for poor pregnant women, who, in addition, receive financial incentives for delivery with a skilled provider (either at home or at a health facility). The program also encompasses the supply-side as it offers financial incentives to providers that serve voucher patients. However, although being opened to both public and private health providers conditional on pre-defined quality standards, Hatt et al. (2010) [78] note that very few private providers actually participated in the program. This is due to both their absence from the poor areas targeted by the program, where they face weak incentives to locate, and to the possibly insufficient reimbursements rates offered by the scheme. Even though the program seems to be correlated with an increased access to maternal health services (Schmidt et al., 2010 [131]; Hatt et al., 2010 [78]; Nguyen et al., 2012 [114]), Schmidt et al. (2010) [131] underline that “the choice of providers in practice is extremely limited”. Ahmed and Mahmud Khan (2011) [4] conclude that the program might be more successful if combined with supply-side subsidies incentivizing private providers to relocate to under-served areas. Then, designing a voucher scheme allowing a government to harness the capacity of private health providers appears to be particularly tricky.

In 2006, the government of Gujarat, acknowledging the fact that Indian public facilities often “lack the capacity and reach to serve many rural areas” (Mohanan et al., 2014) [106], designed a public-private partnership specifically aimed at taking advantage of the large presence of the private sector in the state, the Chiranjeevi Scheme. The program delivered vouchers to pregnant women living under the poverty line, entitling them to free institutional delivery. In addition, the scheme was covering travel expenses of the beneficiary and of the accompanying person (Bhat et al., 2009) [27]. As noted by Mohanan et al. (2014) [106], many cross-sectional or before-after studies concluded that the program was highly effective at increasing the share of institutional deliveries (Bhat et al., 2009) [27]. However, implementing a difference-in-differences, the authors show that these results were mostly reflecting the bias from either self-selection of pregnant women into the scheme based on specific characteristics or time trends, since the rate of institutional deliveries might have increased even in the absence of the voucher program. They use the timing of implementation of the scheme across the different districts of Gujarat to identify whether changes in

the share of institutional deliveries were linked to the introduction of the voucher program across districts. They conclude that the program was not associated with a significant increase in the probability of institutional delivery or in the use of antenatal and postnatal services. On top of the presence of time-varying confounding factors that could bias the results, Mohanan et al. (2014) [106] advance two other possible explanations for these disappointing findings. First, it might be that the perceived quality of the private facilities participating in the program was low so that women did not want to use their services, even for free. It is indeed possible that the voucher scheme alone failed to provide the right incentives to health providers to improve their quality. The second possibility is that institutional deliveries in private hospitals were not actually free of charge, even for women benefiting from a voucher. Indeed, Mohanan et al. (2014) [106] find that the program has not been associated with a decrease in out-of-pocket expenditures related to institutional deliveries. This suggests that participating private providers might still charge fees or increase side-expenditures, not covered by the voucher program. For example, it has been reported that institutional deliveries were associated with larger expenditures on medicines for poor women benefiting from the program compared to poor women who did not receive a voucher (Bhat et al., 2009) [27]. Then, in the absence of specific regulations and monitoring, out-of-pocket expenditures charged by private providers might then offset the benefits of the voucher.

A last type of voucher program, giving access for free or at subsidized prices to insecticide-treated bed-nets (ITNs) that are an important tool in the prevention of malaria, has been particularly popular in developing countries. The public-private partnership implemented in Tanzania to launch such a scheme has received quite a lot of attention, although no rigorous evaluations has been conducted. Still, some studies offer useful information to assess the success of the program in increasing the use of ITNs. After some pilot projects and the use of social marketing interventions, two schemes were actually implemented in parallel from 2002: one was targeting pregnant women who, during their first antenatal care visit, received a voucher giving access to a ITN at a subsidized price after payment of a top-up to participating private retailers, while the second was directed at under-5 children (Donaldson and Thiede, 2011) [50]. Early studies, focusing on the initial PPP projects, find that the use of ITNs vouchers by pregnant women was rather low. For example, Marchant et al. (2002) [97], using a survey of pregnant women living in one of the areas of implementation of the program, find that only 33% of the 505 pregnant women who were interviewed were using an ITN. The main issue does not seem to be a misuse of the voucher, but rather a lack of awareness about the program, and implementation failures. Indeed, the large majority of individuals who received

a voucher used it to acquire an ITN (Marchant et al., 2002 [97]; Tami et al., 2006 [134]). In addition, Tami et al. (2006) [134] report that “94% of the ITNs bought with vouchers were used by those intended, women and children under 5 years”, demonstrating that the acquisition of a voucher translated into its effective utilization. However, very few people actually knew about the program (Mushi et al. 2003) [111]: two years after it started, less than half of the women with children under five surveyed had heard about the scheme and less than 12% had the chance to receive a voucher. Similarly, Marchant et al. (2002) [97] outline that while 28% of all the women interviewed had heard about the program, only 2% actually received a voucher through the scheme. The authors note that this could come from the fact that the voucher scheme was relatively new at the time of the interview, and that there had been problems with its implementation: some health workers appeared to be reluctant to distribute vouchers, while some retailers refused to exchange ITNs against the voucher. Overall, acknowledging that vouchers for ITNs have the potential to increase their use by vulnerable individuals, Tami et al. (2006) [134] recommend to implement major promotion campaigns in order to increase awareness of the scheme, as well as of its eligibility criteria. In addition, they underline that clinics and, to some extent, retailers should be increasingly monitored in order to make sure they correctly implement the scheme.

Overall, due to the lack of robustness of the studies reviewed here, it seems difficult to reach clearcut conclusions on the impact of voucher programs on the use of health care services. The conclusion that can be drawn rather relates to a set of necessary conditions that should accompany voucher programs. First, financial constraints should be identified as an actual barrier to utilization for voucher schemes to be successful. Second, promotion campaigns should be implemented in order for people to be aware of the programs. Finally, monitoring seems required in order to make sure the scheme is properly implemented.

Another type of demand-side interventions, conditional cash transfers (CCT), actually require similar conditions to be effective. We review here CCT specifically directed at health, and not bundled with other type of conditions such as school attendance. In addition, as not all CCT programs fall under the scope of PPPs, we restrict our review to those schemes that explicitly involve private health care providers.²⁷ Some CCT programs of this type exist and have been evaluated. For example, in 2005, the government of India launched a CCT program, the Janani Suraksha Yojana (JSY), aimed at incentivizing women to give birth in a health facility. Eligible women were entitled to receive between 600 and 1000 Indian rupees in urban areas and between 700 and 1400 rupees in rural areas after

²⁷For example, we do not include studies covering the Nicaraguan Red de Protección Social program or the Indonesian PKH and Generasi CCT programs, to the extent that, although these schemes involved private health care providers, they encompassed conditions based on both education and health.

delivery in a government or accredited private health facility. Exploiting two rounds of a national health interview survey, Lim et al. (2010) [94] implement a difference-in-differences to evaluate the impact of this program on the use of institutional delivery and antenatal care. They find that receiving a payment from the JSY significantly increased the use of antenatal care and in-facility deliveries, although confidence intervals are large. Of course, the robustness of this study is limited by the fact that time-varying confounding factors are likely to be correlated with the selection of women in the scheme. In addition, even if these findings look promising, the authors note that the program failed to reach the most deprived part of the population, suggesting that an “improvement of the targeting of this programme is required”. They underline that communication strategies that do not require to be literate should be implemented, while program authorities should ensure that remote health facilities are accredited to participate in the scheme.

One can note that demand-side interventions give rise to many issues related to the fact that the supply-side is not monitored adequately. Another type of public-private partnerships, the “pay-for-performance” model, might solve this issue by involving a strict monitoring of the supply side: such contracts are aimed at providing health providers with financial incentives based on their success in achieving specific outcomes and/or a pre-determined level of service quality. This kind of policies can increase service utilization by motivating health providers to improve outreach interventions, reduce their user fees, adapt their clinic hours, or stimulate improvements in technical quality and responsiveness (Eichler, 2006) [53]. Unfortunately, Huillery and Seban (2014) [85] underline that most of the studies evaluating performance-based-financing policies in health care use simple before/after comparison or non-credible comparison groups, so that their estimates are likely to be biased by time trends or confounding factors. They acknowledge that the only scheme that has been studied rigorously is the Rwandan one, through evaluations combining randomization and a difference-in-differences (Basinga et al., 2011 [23]; Gertler and Vermeersh, 2013 [69]). However, it is highly unclear whether the government of Rwanda involved private health care centers in its pay-for-performance program, so we do not review this study.

Still, there exist some evidence of the impact of performance-based financing. Soeters et al. (2011) [133] study the effectiveness of a performance-based-financing program launched in two districts of the Democratic Republic of Congo, offering payments for specific health services to independent health facilities, while encouraging them to develop subcontracts with private providers.²⁸ Using two rounds of survey conducted in the treated districts

²⁸The health care services chosen to receive subsidies include “outpatient department consultancies, number of bed days, fully immunizing a child before twelve months of age, construction of a household pit latrine, and use by a woman of oral or injectable contraceptives.”

as well as two neighboring districts before and after the start of the program, and controlling for time-invariant confounding factors by implementing a difference-in-differences, the authors find that the change in the share of pregnant women delivering at a health facility (rather than at home) was 21 percentage points higher in control districts, and that the scheme did not significantly improve the use of other health care goods and services such as ITNs, antenatal care, use of modern family planning methods or visit to a health facility when being sick.²⁹ These surprising findings might be explained by the increased investments made by non-governmental organizations in control districts' health facilities. In addition, the scheme allowed treated facilities to negotiate their user fees with the community they served, while control facilities' user fees were regulated. This might have resulted in higher user fees in districts where the program was implemented, thus explaining the findings. It thus seems important that user fees regulations complement any performance-based financing schemes.

Zang et al. (2015) [145] implement a strategy similar to the one of the previous study in order to evaluate the impact of a performance-based-financing pre-pilot in Cameroon, launched in four health districts in 2011. Combining propensity-score matching and a difference-in-differences to control for observable and unobservable constant confounding factors, they report that the scheme did not significantly increase utilization of health care services such as outpatient consultation, vaccination of children, modern contraceptive methods, institutional delivery, antenatal care or antitetanic vaccination for pregnant women. However, the authors note that these findings might come from the high utilization of such services at baseline, implying that the margin for improvement was really small. In 2012, the government expanded the program to additional districts of the North West, South West, and East regions. Health facilities in selected districts were randomly assigned to one of four groups: performance-based financing; same monitoring, supervision and financial resources than performance-based financing but without financial incentives linked to performance; same supervision and monitoring; pure control group. Despite having been randomized, some baseline characteristics appear to be unbalanced.³⁰ In a study evaluating the scheme, De Walque et al. (2017) [137] attenuate this issue by combining a difference-in-differences to the ex-ante randomization, exploiting baseline and follow-up household and health facility surveys. However, they also note that, ideally, the randomization should have been conducted at the district level since the proximity between health facilities belonging to different treatment groups might allow them to influence each other. Another issue is that the household-level analysis

²⁹The latter outcome appears to be positively impacted by the program but this is only significant at the 10 percent level.

³⁰The performance-based financing package includes payments linked to both the quantity of certain services provided and the quality of care, measured by indicators such as the number of "new cases of STIs diagnosed and correctly treated in the month according to national protocols". The services encompassed by the performance payments include curative care (number of outpatient consultation, number of hospital bed days...), preventive care (vaccination, STI treatments, TB treatments,...) and reproductive health care. In addition, independent monitoring is performed.

suffers from an attenuation bias to the extent that the treatment variable is incorrectly measured: the authors arbitrarily assumed that household members use the closest health facility, while data show it is not the case for a substantial share of them. Fortunately, estimates from the household-level specifications are complemented by the ones from the facility survey, which do not suffer from measurement errors. Overall, although imperfect, the empirical strategy implemented in this paper offers one of the most rigorous evaluation of a performance-based-financing policy. Regarding the utilization of health care services, the findings from the household-level data are similar to the ones from the pre-pilot program. Performance-based-financing did not significantly increase the use of antenatal care, postnatal care, modern contraception methods, tetanus vaccination for pregnant women, ITNs, nor the share of delivery assisted by a skilled health worker. This lack of significance does not seem to fully come from the attenuation bias as some facility-level and household-level estimates are consistent. Still, they display significant differences since the analysis at the survey-level reveals that performance-based-financing did significantly increase the use of tetanus vaccination for pregnant women, postnatal care and modern contraception. Overall, estimates robust across the different specifications allow us to conclude that the program has non-negligible positive impacts on utilization of health care services since it significantly increases HIV testing and vaccination coverage for both children and pregnant women. Similarly to Zang et al. (2015) [145], the authors suggest that the absence of significant impact on antenatal care or institutional deliveries might be explained by the already high level of utilization of such services at baseline. Another possible explanation is the level of user fees, which might act as a “barrier on the demand-side”. They thus suggest that, conditional on the actual use of certain health care services being low, the government should combine such supply-side policies to demand-side ones - such as voucher schemes.

Finally, Huillery and Seban (2014) [85] provide another set of robust estimates of the impact of a performance-based-financing policy by studying a mechanism through which the Congolese government links payment to health facilities to the number of patients for some pre-determined health services.³¹ As part of a research project running from 2009 to 2013 in the Haut-Katanga district, the government randomly assigned its 96 health areas to either a treatment group - receiving performance-based payments - or a control group - receiving fixed payment of the same amount than the treated group to abstract from resource effects. The authors find that the policy did not have any significant impact on the utilization of any health care services, including the share of individuals who got sick but did not visit a health care facility, tuberculosis vaccination, institutional delivery, antenatal and postnatal

³¹The targeted services were outpatient first curative consultations, prenatal consultations, deliveries, obstetric referral, children complete vaccination, tetanus toxoid vaccination, and family planning consultations.

care or modern contraceptives. These results seem all the more surprising given that the program did significantly decrease user fees, facilities were incentivized to implement strategies aimed at attracting more patients, and health workers significantly augmented the number of preventive sessions organized at facility and the number of outreach activities in the community, aimed at making people aware of the services available. Contrary to previous studies, this cannot be explained by the fact that utilization rates were already high at baseline. The authors suggest that health care services do not fall in the category of normal goods. However, another possibility is that households' budget is so tight that they cannot afford even cheaper health services. In this case, a combination of vouchers with performance-based financing would make sense. The authors also note that one possibility is that the decrease in user fees has been understood as a signal of bad quality. However, this is not fully convincing since, if it was the case, we would observe a significant decrease in utilization of health care services in treatment relatively to control areas. Still, it is true that quality improvements might attract more patients. Performance-based financing should not only be based on the quantity of patients served but also on the quality of services offered, unlike what has been implemented in the Congolese context. Overall, the authors conclude that, while the health care providers proved unable to increase the utilization of health services, "specific interventions to stimulate demand for health may be combined with supply-side interventions like PBF," such as "interventions to improve awareness about the benefits of health products or to help people overcome behavioral issues like procrastination."

It seems difficult to infer general lessons on the potential for PPPs in health to improve service utilization. One reason is that it exists many forms of such interventions, which might differently impact our outcomes of interest. In addition, the literature is highly scattered, in the sense that very few papers exist on each possible type of PPPs, and that they are of unequal quality.³² Still, one general lesson that can be derived from the studies reviewed here is that, given the complexity PPPs in health can display, a proper diagnosis of the situation has to be made ex ante to understand the reasons of service under-utilization: is the issue a lack of demand for health services? If yes, can it be explained by a lack of awareness on the benefits of health care, by credit constraints, or by the poor quality of the services offered? The answers to these questions could help choose the adequate type of PPP to implement. In addition, an important consensus that seem to be reached through the different studies reviewed here is the necessity of combining demand- and supply-side policies in order to effectively increase health care utilization.

³²For example, while franchising can be understood as a type of health PPP, we were not able to find papers covering franchising policies directed at private health facilities implemented by a government, and not by the private sector.

5.2 Quality of Services, Patient Satisfaction and Health-Related Outcomes

As happened in education, for which the large increase in enrollment rates has been followed by a learning crisis, increases in health care services utilization may be necessary but not sufficient to improve health-related outcomes in the population. Quality of health care is also an important determinant of the latter. And contrary to what happened in education, the switch of focus of health policymakers from coverage to quality has not yet fully happened, as many health care goods and services are still largely under-utilized.

One important limitation is that, unlike learning which can be assessed through standardized test scores, there is not a unique way to measure health. In addition, health data are difficult to collect, given medical secret and privacy regulations, so that fewer studies exist. The different papers reviewed here use a broad range of health-related outcomes, spanning from maternal morbidity to patient satisfaction and health workers' effort.

Voucher programs, when competitive, allow patients to choose their preferred health-care provider. The latter have thus an incentive to compete in order to attract more patients. This should translate in health care quality improvements and ultimately lead to better health outcomes in the population (Obare et al., 2013) [116]. In a study of the Nicaraguan reproductive health voucher program directed at adolescents, Meuwissen et al. (2006) [102] underline how doctors' knowledge, attitudes and practices related to sexual and reproductive health care services for adolescents translate into poor health care quality. Interviews conducted with health care workers before the implementation of the program reveal that doctors are "reluctant to prescribe contraceptives to younger adolescents", while they lack knowledge on contraceptive methods, STIs and syndromic treatment. Despite studying only a small sample, the authors argue that doctors working in public clinics scored significantly worse than private doctors in term of their attitude towards contraceptive methods accessibility and medical knowledge. Thus, it again underlines the importance of involving private clinics through health PPPs. However, although informative with respect to the factors of low health care quality, the before-after comparison implemented by the authors is too loose to be interpreted as the program's impact on doctors' attitudes and knowledge. A somewhat more rigorous study has been conducted to evaluate the impact of the voucher program for maternal health implemented in Bangladesh. However, as underlined above, private health providers did not face the right incentives to participate in the program. Then, the promising findings coming from the comparison of health-related outcomes in program intervention sub-districts to matched control ones cannot be interpreted as improvements from the involvement of private health providers in health care delivery.³³ Finally, studies evaluating the Chiranjeevi Yojana scheme have been able to

³³Hatt et al. (2010) [78] find that the incidence of stillbirths and of newborn deaths is significantly lower in the sub-districts where

retrieve its impact on maternal health outcomes. The most rigorous one, using a difference-in-differences, is the work conducted by Mohanan et al. (2014) [106], already discussed above. The authors report that the probability of delivery with complications - including “prolonged and obstructed labour, excessive bleeding, breech presentation, convulsions, hypertension, fever, incontinence or other birth-related problems” - did not significantly decrease after the introduction of the program. This might not be surprising given that we already underlined that the scheme did not increase the proportion of deliveries in health facilities. However, it is particularly alarming to the extent that, 54% of women interviewed by Mohanan et al. (2014) [106] suffered from delivery complications. Finally, we also have to note that, to the best of our knowledge, there exists no study evaluating the impact of ITNs voucher programs on health-related outcomes. However, again, the possibility to detect that such programs improve health-related outcomes relies on the fact that they actually increase ITNs utilization, which does not seem to have been the case in the previously mentioned ITNs studies. Overall, this review underlines the broad lack of research evaluating the impact of vouchers scheme on health-related outcomes. In addition, to be informative and help design health PPPs, future studies should particularly focus on disentangling the channels through which voucher programs improve (or not) health outcomes. In particular, the following questions should be addressed: is the improvement only mechanical, following from an increase in the use of health services? Is part of these improvements related to better quality of care? Or, on the contrary, does the poor quality of care limit the extent to which increases in health care utilization translate into better health outcomes?

The number of studies evaluating the health-related impacts of PPP for conditional-cash transfers program is also limited, but can be explained by the narrowness of the definition we use, implying that very few CCT programs can be considered as a PPP. Lim et al. (2010) [94], in their evaluation of the Janani Suraksha Yojana CCT program, are able to measure the program’s impacts on maternal health outcomes through a difference-in-differences analysis. Their findings are surprising as they report that, although the scheme significantly increased the share of in-facility deliveries and the use of antenatal care, it failed to decrease significantly the number of perinatal, neonatal and maternal deaths. However, the authors themselves recognize that their study was largely under-powered, as suggested by the very wide confidence intervals they report. Then, they speculate that “the reason a significant effect was not noted [...] is most likely due to inadequate statistical power rather than a lack of an effect.” Still, this supposition should be backed up by rigorous empirical evidence, as many other things might be at play and could explain a failure of the scheme to improve health outcomes. For example, it is possible the program has been implemented compared to the control ones.

that while increasing maternal health care and institutional delivery utilization, the scheme has participated in the deterioration of the quality of these services: the authors indeed mention that the JSY has led to “increased workloads and reduced quality of care in health facilities”, resulting in, for example, early discharge after delivery. Another possibility is that the ex-ante quality of care offered by health facilities was so poor that, even if the scheme did not actually deteriorate it, it prevented the observed increase in utilization to channel into better health outcomes. Overall, no conclusion can be made on the impact of this CCT program on health-related outcomes from this study only so that, as suggested by the authors themselves, “further investments in monitoring and evaluation — including both impact and process evaluation — are imperative to improve understanding of the association between JSY and health outcomes”. This advice seems relevant to the full set of ongoing and future health PPP programs.

Pay-for-performance schemes might be more directly associated to improvements in health-related outcomes and health care quality than demand-side initiatives. It is indeed the case that each of the performance-based-financing studies presented in the previous subsection evaluate the impact of the scheme on quality of health care and/or health-outcomes. Through a difference-in-differences, Soeters et al. (2011) [133] report that patient-perceived quality increased significantly more in the Congolese participating districts than in the control ones. However, patient perceived-satisfaction is an imperfect indicator of quality of care. For example, patients may interpret the increase in user fees in some participating health facilities as a signal of quality. A less subjective quality indicator, qualified staff in the participating health centers, increased by twenty-three percentage points in health facilities located in participating districts compared to only eight percentage points in control districts. This improvement reflects the fact that participating health facilities hired new qualified staff but also involved the private health providers through subcontracts. It would have been interesting to know whether, despite the lack of impact on health care utilization reported in the previous section, quality enhancements alone improved health outcomes. Unfortunately, no results are reported with respect to the latter, certainly due to data limitations.

Findings appear to be similar in Cameroon, where Zang et al. (2015) [145] combine a difference-in-differences with propensity-score matching in order to evaluate the performance-based-financing pre-pilot implemented in 2011. Indeed, while the program failed to increase health service utilization, it appears to have significantly improved several dimensions of health care quality, measured through a “health facility quality checklist” completed at the time of the health facility survey.³⁴ One difference with the previous study is that the program did not

³⁴The improvements happened along the following dimensions: institutional quality, outpatient care quality, maternity care quality,

significantly impact the number of staff per facility. Although there is again no assessment of the impact of the pilot program on health-related outcomes, the authors note that, at least in urban settings, “quality of care seems to be the most likely area for improvement as there may be little room to improve health service utilization in many settings.” Actually, the evaluation conducted after the expansion of the program to new districts, in which health facilities were randomly assigned to different treatment and control groups, shed light on the impact of performance-based-financing on a broader set of outcomes. Combining the ex-ante randomization with a difference-in-differences strategy, de Walque et al. (2017) [137] first show that the overall satisfaction score reported by women who attended antenatal care consultations increased in performance-based-financing facilities compared to control facilities, although this result is only marginally significant (p-value=0.086). Over 12 indicators of quality, their satisfaction increased significantly with respect to the cleanliness of the health facility, the level of privacy during the visit, the adequateness of the opening hours and the communication with the health worker, which seems to indicate that some efforts have been made by participating health facilities in order to attract more patients. The significant increase in overall satisfaction with respect to child health consultations confirms this further. However, these increases in overall patients’ satisfaction failed to channel into actual improvements in the quality of care, measured through standardized checklists completed by enumerators while observing antenatal care and children consultations.³⁵

Interestingly, to understand whether the program translated into quality changes at the facility level, an analysis of health workers satisfaction and availability of drugs and equipment is also performed. The authors find that the program significantly improved their satisfaction regarding the overall physical conditions of the facility building, and the quantity and quality of the equipment available in the health facility, which is confirmed by the analysis of the health facilities’ equipment which has significantly increased over time compared to control facilities. In addition, the number of nurses per facility also increased significantly over time in participating facilities compared to control facilities. However, the performance-based-financing program failed to impact significantly the quantity and quality of drugs (such as paracetamol, amoxicillin tabs or syrup, ORS, iron tabs, and cotrimoxazole, but also malaria treatment medicines and vaccines) available in the health facility. This might explain why no changes are reported by health workers with respect to their perceived ability to provide high quality care given the working conditions in the facility. Overall, the authors conclude that performance-based financing is an effective mechanism

family planning health care quality, vaccination and antenatal care quality, as well as supply availability.

³⁵The checklists comprised information on whether health workers were performing pre-defined routine activities, such as washing their hands, asking questions about the child’s health status, checking HIV status, testing for syphilis, and providing malaria prophylaxis.

in improving the structural quality of care, although they also note that monitored facilities in which payments were not linked to performance observed relatively similar improvements than in the performance-based-financing facilities. On the other hand, facilities which were both supervised and monitored without additional financing display few improvements relatively to control facilities. Hence, what matters most seems to be the allocation of additional resources to health facilities to allow them to upgrade the quality of their equipment and infrastructure. Still, other mechanisms should be explored in order to improve the overall quality of care.

Finally, Huillery and Seban (2014) [85] also study quality outcomes in the performance-based-financing launched by the Congolese government. They find no significant differences between treated and control health facilities with respect to consultation time, patients' understanding of health services, patients' satisfaction or, maybe more importantly, health workers' compliance rate with standard medical procedures. They note that this non-significant change in service quality was expected to the extent that the incentives were based on volume of services rather than on quality, similarly to the previous scheme.³⁶ In fact, given the decrease in user fees and the failure of the program to increase service utilization, participating health facilities reported significantly lower total financial resources than comparison health facilities. This led to a significant negative impact of the scheme on both the quantity and quality of equipment and infrastructure. Unsurprisingly given its failure to increase either utilization or quality of health services, the program did not improve population health status, measured by under-five children weight-for-height z-score, number of female death for perinatal reasons, and number of deaths of children under five. Overall, as already mentioned, complementary demand-side interventions might be needed in order to significantly increase service utilization under a performance-based-payment scheme, but Huillery and Seban (2014) [85] also underline the need to include financial incentives linked to service quality.

In addition to the lack of success of the program, the authors also underline how incentives payments could backfire by crowding-out workers' intrinsic motivation. They provide evidence that it actually is the case as the positive impact (significant at the 10% level) of the program on staff attendance was completely and significantly reversed once the scheme was withdrawn. They thus recommend that performance-based-financing programs should be used as a permanent policy in order to avoid such detrimental impacts, which are likely to decrease service quality.

Overall, it is again difficult to draw a general conclusion on the impact of PPPs on health-related outcomes from the studies reviewed in this section. For voucher and CCT programs, it is impossible to conclude whether an absence

³⁶Indeed, as already mentioned, payments were linked to the *number of patients* for some pre-determined health services such as outpatient first curative consultations, prenatal consultations, deliveries, obstetric referral, children complete vaccination, tetanus toxoid vaccination, and family planning consultations.

of health improvement comes from the failure of the schemes to increase utilization and to increase competition or if it would have failed anyway due to the poor quality of services offered. Findings are somewhat more informative for pay-for-performance programs. Such programs appear to be more likely to improve the “structural quality of care”. Interestingly, the allocation of additional resources to health facilities rather than incentives payments might be the main determinant of these improvements. However, an alternative that should be explored by future studies is to link payments to the quality of the service directly, rather than to its quantity. In any case, authors underline that such supply-side policies should be coupled with demand-side interventions, which would boost the financial resources available to health facilities. Finally, this literature is limited by the lack of data capturing health status.

5.3 Segregation

As already underlined, we are unaware of studies evaluating PPPs based on differentiating between participating private and public health facilities. Then, contrary to the section studying PPPs in education, we will be unable to discuss the potential impact of health PPPs on segregation with respect to health care, i.e. whether PPPs in health drive some individuals to use private or public health facilities based on their characteristics. Fortunately, segregation might not be a key determinant of health-related outcomes.

Indeed, as mentioned when studying education, segregation seems to emerge anyway as soon as different *types* of health facilities co-exist within a system. However, while for education peer-effects were appearing within a classroom, making segregation detrimental even when school quality was kept constant, within-facility health spillovers should actually be negligible. Then, stratification in health facilities based on patients’ characteristics would not matter as long as public authorities ensure that the different types of facilities follow the same standards with respect to both quality and accessibility. What matters is then to bring improvements directly to the poor and under-served individuals.

Overall, although PPPs might increase stratification by creating different health systems, they should not widen health gaps in the population to the extent that within-facility health spillovers should be negligible. In fact, if PPPs allow to increase the quantity and quality of the health services available to the more deprived population, such policies should participate in closing the existing health gaps.

5.4 Cost-Effectiveness

While the impact of PPPs in health on user fees and health facilities' revenues has briefly been mentioned in the previous sections, we also want to understand whether such interventions are cost-effective. Similarly to education, few studies seem to have implemented such cost-effectiveness analyses of PPPs in health care. Regarding voucher programs, Borghi et al. (2005) [35] try to assess the cost-effectiveness of the scheme implemented in the capital city of Nicaragua, Managua, which gave access to free STI testing and treatment, health education and condoms in both private and public health facilities for at-risk groups. They estimate the outcomes and costs of the ongoing project but also the ones that would have been incurred in the absence of the scheme, i.e. in the "status quo situation." They conclude that the voucher program was an "effective and efficient means of targeting and effectively curing STIs in high-risk groups." Indeed, according to their calculations, the average cost of the voucher scheme per STI cured went from US\$200 in the absence of the program to US\$118 under the voucher scheme. In addition, even though the authors estimate that the average cost per consultation was higher under the scheme (US\$41 compared to US\$12), they underline that their analysis does not capture the number of additional STI, and particularly HIV, cases averted through behavioral changes promoted by the program.

A range of studies have focused on evaluating the cost-effectiveness of the ITN voucher schemes implemented by the government of Tanzania through PPPs. Hanson et al. (2003) [74] estimate the cost-effectiveness of the earliest of these schemes, implemented in two rural districts from July 1996 to June 2000. They estimate that "the cost per death averted associated with treated nets was US\$ 1559 and the corresponding cost per DALY averted was US\$57". Focusing on the subsequent program, the Tanzania National Voucher Scheme, Mulligan et al. (2008) [109] estimate that, given the number of ITNs and re-treatment kits distributed through the scheme between 2004 and 2006 and their utilization rates, the intervention has resulted in "a total of 2,499,042 treated net years and a cost per treated net year of \$4.23". Using estimates of ITN effectiveness from the literature with respect to the number of children deaths avoided and to the incidence of malaria, the authors compute that the governmental program might have allowed to avoid "12,039 child deaths at an economic cost of \$873 per child death averted" and 794,995 malaria cases at a cost of \$13 per malaria case averted. This appraisal allows the authors to conclude that the Tanzanian PPP was a "a cost-effective strategy for delivering subsidized ITNs to targeted vulnerable groups".

6 Conclusion

This literature review is an attempt to summarize the state of the empirical analysis about the impact of PPPs in key sectors – infrastructure (defined to encompass energy, transport, water and sanitation, and telecommunications), education, and health - while shedding light on what remains to be studied. We have focused first on the impact of PPPs on efficiency, and examined whenever possible, the implications in terms of potential trade-offs between coverage and affordability.

Overall the findings are mixed. While there is some evidence of efficiency gains stemming from the implementation of PPPs, especially in infrastructure and education, it is hard to conclude unambiguously that PPPs are always the best solution. Regarding infrastructure, most of the literature struggles to convincingly address the fundamental selection issues and provide a satisfactory counterfactual to PPP projects. In education, while some programs appear to have been successful, this seems to depend crucially on their design and scale, and PPPs are also associated with an increase in school segregation. Finally, the state of knowledge is still quite limited when it comes to PPPs in health. We conclude that the jury is still very much out, and that more good quality studies with convincing identification strategies are needed.

References

- [1] Atila Abdulkadiroglu, Joshua Angrist, Susan Dynarski, Thomas J. Kane, and Parag Pathak. Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots. Technical Report 15549, November 2009.
- [2] Atila Abdulkadiroğlu, Parag A. Pathak, and Christopher R. Walters. Free to Choose: Can School Choice Reduce Student Achievement? *American Economic Journal: Applied Economics*, 10(1):175–206, January 2018.
- [3] Josefa Aguirre. Can Progressive Vouchers Help the Poor Benefit from School Choice? Evidence from the Chilean Voucher System. Technical Report ID 3123670, Rochester, NY, July 2017.
- [4] Shakil Ahmed and M. Mahmud Khan. A maternal health voucher scheme: what have we learned from the demand-side financing scheme in Bangladesh? *Health Policy and Planning*, 26(1):25–32, January 2011.
- [5] Priyanka Anand, Alejandra Mizala, and Andrea Repetto. Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile. *Economics of Education Review*, 28(3):370–381, June 2009.
- [6] L. Andres. The impact of Privatization on Firms in the Infrastructure Sector in Latin American Countries. *PhD Dissertation, University of Chicago*, 2004.
- [7] Luis A. Andres, J. Luis Guasch, Thomas Haven, and Vivien Foster. *The Impact of Private Sector Participation in Infrastructure: Lights, Shadows, and the Road Ahead*. World Bank Publications, July 2008. Google-Books-ID: I0IpzjC4U0QC.
- [8] Luis Alberto Andres, J. Luis Guasch, and Jordan Z. Schwartz. Main report. Technical Report 78679, June 2013.
- [9] Luis A. Andrés, Jordan Schwartz, and J. Luis Guasch. *Uncovering the Drivers of Utility Performance: Lessons from Latin America and the Caribbean on the Role of the Private Sector, Regulation, and Governance in the Power, Water, and Telecommunication Sectors*. The World Bank, June 2013.

- [10] Joshua Angrist, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review*, 92(5):1535–1558, December 2002.
- [11] Joshua Angrist, Eric Bettinger, and Michael Kremer. Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. *The American Economic Review*, 96(3):847–862, 2006.
- [12] Joshua D Angrist, Sarah R Cohodes, Susan M Dynarski, Parag A Pathak, and Christopher R Walters. Stand and Deliver: Effects of Boston’s Charter High Schools on College Preparation, Entry, and Choice. page 44, 2016.
- [13] Joshua D Angrist, Susan M Dynarski, Thomas J Kane, Parag A Pathak, and Christopher R Walters. Inputs and Impacts in Charter Schools: KIPP Lynn. *American Economic Review*, 100(2):239–243, May 2010.
- [14] Aneesa Arur, Nelson Gitonga, Barbara O’Hanlon, Francis Kundu, Maggie Senkaali, and Richard Ssemujju. Insights from Innovations: Lessons from Designing and Implementing Family Planning/Reproductive Health Voucher Programs in Kenya and Uganda. *Private Sector Partnerships-One project*, Abt Associates Inc., 2009.
- [15] Emmanuelle Auriol and Stéphane Straub. International Handbook On The Economics Of Corruption. volume 2. Edward Elgar Publishing, 2011.
- [16] Abhijit Banerjee, Angus Deaton, and Esther Duflo. Health, Health Care, and Economic Development. *The American economic review*, 94(2):326–330, May 2004.
- [17] Abhijit Banerjee and Esther Duflo. Addressing Absence. *The journal of economic perspectives : a journal of the American Economic Association*, 20(1):117–132, 2006.
- [18] Abhijit V. Banerjee, Rachel Glennerster, and Esther Duflo. Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System. *Journal of the European Economic Association*, 6(2-3):487–500, 2008.
- [19] Alexandro Barbosa and Isabel Brusca. Governance structures and their impact on tariff levels of Brazilian water and sanitation corporations. *Utilities Policy*, 34:94–105, June 2015.

- [20] Felipe Barrera-Osorio, David Blakeslee, Matthew Hoover, Leigh Linden, and Dhushyanth Raju. Expanding educational opportunities in remote parts of the world : evidence from a RCT of a public-private partnership in Pakistan. Technical Report 81322, August 2013.
- [21] Felipe Barrera-Osorio, De Galbert, Pierre Gaspard, James Paul Habyarimana, and Shwetlena Sabarwal. Impact of public-private partnerships on private school performance : evidence from a randomized controlled trial in Uganda. Technical Report WPS7905, December 2016.
- [22] Felipe Barrera-Osorio and Dhushyanth Raju. Evaluating public per-student subsidies to low-cost private schools : regression-discontinuity evidence from Pakistan. Technical Report 5638, April 2011.
- [23] Paulin Basinga, Paul J. Gertler, Agnes Binagwaho, Agnes L. B. Soucat, Jennifer Sturdy, and Christel M. J. Vermeersch. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet (London, England)*, 377(9775):1421–1428, April 2011.
- [24] Nicole M. Bellows. Vouchers for reproductive health care services in Kenya and Uganda. *Discussion Paper. KfW Bankengruppe*, 2012.
- [25] Sanford V. Berg and Rui Cunha Marques. Quantitative studies of water and sanitation utilities: a literature survey. Technical Report 32891, 2010.
- [26] Eric Bettinger, Michael Kremer, and Juan E. Saavedra. Are Educational Vouchers Only Redistributive? *Economic Journal*, 120(546):F204–F228, 2010.
- [27] Ramesh Bhat, Dileep V Mavalankar, Prabal V Singh, and Neelu Singh. Maternal Healthcare Financing: Gujarat’s Chiranjeevi Scheme and Its Beneficiaries. *Journal of Health, Population and Nutrition*, 27(2), September 2009.
- [28] Arunava Bhattacharyya, Thomas R. Harris, Rangesan Narayanan, and Kambiz Raffiee. Specification and estimation of the effect of ownership on the economic efficiency of the water utilities. *Regional Science and Urban Economics*, 25(6):759–784, December 1995.
- [29] Robert Bifulco and Helen F. Ladd. The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina. *Education Finance and Policy*, 1(1):50–90, March 2006.

- [30] Robert Bifulco and Helen F. Ladd. School choice, racial segregation, and test-score gaps: Evidence from North Carolina's charter school program*. *Journal of Policy Analysis and Management*, 26(1):31–56, December 2007.
- [31] Frédéric Blanc-Brude, Hugh Goldsmith, and Timo Väililä. A Comparison of Construction Contract Prices for Traditionally Procured Roads and Public–Private Partnerships. *Review of Industrial Organization*, 35(1-2):19, September 2009.
- [32] C. Bonnet, P. Dubois, D. Martimort, and S. Straub. Empirical Evidence on Satisfaction with Privatization in Latin America. *The World Bank Economic Review*, page 26(1): 133, 2012.
- [33] Kevin Booker, Scott M. Gilpatric, Timothy Gronberg, and Dennis Jansen. The impact of charter school attendance on student performance. *Journal of Public Economics*, 91(5):849–876, June 2007.
- [34] Kevin Booker, Ron Zimmer, and Richard Buddin. The Effects of Charter Schools on School Peer Composition. 2005.
- [35] Josephine Borghi, Anna Gorter, Peter Sandiford, and Zoyla Segura. The cost-effectiveness of a competitive voucher scheme to reduce sexually transmitted infections in high-risk groups in Nicaragua. *Health Policy and Planning*, 20(4):222–231, July 2005.
- [36] Carinne Meyer Brody, Nicole Bellows, Martha Campbell, and Malcom Potts. The impact of vouchers on the use and quality of health care in developing countries: a systematic review. *Global Public Health*, 8(4):363–388, 2013.
- [37] Esther Buregyeya, Elizeus Rutebemberwa, Phillip LaRussa, Sham Lal, Sian E. Clarke, Kristian S. Hansen, Pascal Magnussen, and Anthony K. Mbonye. Comparison of the capacity between public and private health facilities to manage under-five children with febrile illnesses in Uganda. *Malaria Journal*, 16, May 2017.
- [38] Roland Bénabou. Human capital, inequality, and growth: A local perspective. *European Economic Review*, 38(3):817–826, April 1994.
- [39] Anders Böhlmark, Helena Holmlund, and Mikael Lindahl. School choice and segregation: evidence from Sweden. Technical Report 2015:8, May 2015.
- [40] Rajashri Chakrabarti. Do vouchers lead to sorting under random private school selection? Evidence from the Milwaukee voucher program. *Economics of Education Review*, 34:191–218, June 2013.

- [41] Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 20(1):91–116, March 2006.
- [42] Zhenhua Chen, Nobuhiko Daito, and Jonathan L. Gifford. Socioeconomic impacts of transportation public-private partnerships: A dynamic CGE assessment. *Transport Policy*, 58(C):80–87, 2017.
- [43] SangHyun Cheon, David E. Dowall, and Dong-Wook Song. Evaluating impacts of institutional reforms on port efficiency changes: Ownership, corporate structure, and total factor productivity changes of world container ports. *Transportation Research Part E: Logistics and Transportation Review*, 46(4):546–561, 2010.
- [44] Andrew Yuen Chi-Lok and Anming Zhang. Effects of competition and policy changes on Chinese airport productivity: An empirical investigation. *Journal of Air Transport Management*, 15(4):166–174, July 2009.
- [45] Eshien Chong, Stéphane Saussier, and Brian S. Silverman. Water Under the Bridge: Determinants of Franchise Renewal in Water Provision. *The Journal of Law, Economics, and Organization*, 31(suppl_1):i3–i39, August 2015.
- [46] Vilsa E. Curto and Roland G. Fryer. The Potential of Urban Boarding Schools for the Poor: Evidence from SEED. *Journal of Labor Economics*, 32(1):65–93, January 2014.
- [47] Vijay Kumar Damera. Choice for the poor or poor choice? India’s school choice policy. *Blavatnik School of Government Working Paper, University of Oxford*, 2017.
- [48] Cours des Comptes. La politique immobilière du ministère de la justice. *Rapport public thématique. La documentation française.*, 2017.
- [49] Will Dobbie and Roland G Fryer. Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children’s Zone. *American Economic Journal: Applied Economics*, 3(3):158–187, 2011.
- [50] Dayl Donaldson and Michael Thiede. Tanzania National Voucher Scheme Evaluation: Public-Private Partnership to Distribute Insecticide-Treated Bed Nets to Pregnant Women and Infants. *USAID*, 2011.

- [51] Esther Duflo, Pascaline Dupas, and Michael Kremer. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5):1739–1774, August 2011.
- [52] Esther Duflo, Rema Hanna, and Stephen P. Ryan. Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4):1241–1278, June 2012.
- [53] Rena Eichler. Can “Pay for Performance” Increase Utilization by the Poor and Improve the Quality of Health Services? *Background papers for the Working Group on Performance Based Incentives*, January 2006.
- [54] Eduardo Engel, Ronald Fischer, and Alexander Galetovic. The Basic Public Finance Of Public–Private Partnerships. *Journal of the European Economic Association*, 11(1):83–111, 2013.
- [55] Eduardo Engel, Ronald D. Fischer, and Alexander Galetovic. *The Economics of Public-Private Partnerships: A Basic Guide*. Cambridge University Press, September 2014. Google-Books-ID: _cuKBAAAQBAJ.
- [56] Dennis Epple, Richard Romano, and Ron Zimmer. Charter Schools: A Survey of Research on Their Characteristics and Effectiveness. Technical Report 21256, June 2015.
- [57] Dennis Epple and Richard E. Romano. Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *The American Economic Review*, 88(1):33–62, 1998.
- [58] Dennis Epple, Richard E. Romano, and Miguel Urquiola. School Vouchers: A Survey of the Economics Literature. *Journal of Economic Literature*, 55(2):441–492, 2017.
- [59] Dennis Epple, Richard E. Romano, and Miguel Urquiola. School Vouchers: A Survey of the Economics Literature. *Journal of Economic Literature*, 55(2):441–492, 2017.
- [60] Antonio Estache, Beatriz Tovar de la Fé, and Lourdes Trujillo. Sources of efficiency gains in port reform: a DEA decomposition of a Malmquist TFP index for Mexico. *Utilities Policy*, 12(4):221–230, December 2004.
- [61] Antonio Estache, Marianela Gonzalez, and Lourdes Trujillo. Efficiency Gains from Port Reform and the Potential for Yardstick Competition: Lessons from Mexico. *World Development*, 30(4):545–560, 2002.
- [62] Antonio Estache and Martin Rossi. Have consumers benefited from the reforms in the electricity distribution sector in Latin America? Technical Report WPS3420, October 2004.

- [63] Antonio Estache and Stephane Saussier. Public Private Partnerships and Efficiency: A Short Assessment. *CESifo DICE Report*, September 2014.
- [64] Carsten Fink, Aaditya Mattoo, and Randeep Rathindran. An assessment of telecommunications reform in developing countries. *Information Economics and Policy*, 15(4):443–466, 2003.
- [65] Michael Ka Yiu Fung, Ken Kai Hong Wan, Yer Van Hui, and Japhet Sebastian Law. Productivity changes in Chinese airports 1995-2004. *Transportation Research Part E: Logistics and Transportation Review*, 44(3):521–542, 2008.
- [66] Sebastian Galiani, Paul Gertler, and Ernesto Schargrodsky. Water for Life: The Impact of the Privatization of Water Services on Child Mortality. *Journal of Political Economy*, 113(1):83–120, February 2005.
- [67] Katharina Gassner, Alexander A. Popov, and Nataliya Pushak. *Does Private Sector Participation Improve Performance in Electricity and Water Distribution?* World Bank Publications, 2009. Google-Books-ID: CyQFonsLqtMC.
- [68] Ana Gazmuri. School Segregation in the Presence of Student Sorting and Cream-Skimming. *TSE Working Paper*, 2017.
- [69] Paul Gertler and Christel Vermeersch. Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes. Technical Report w19046, Cambridge, MA, May 2013.
- [70] Philip Gleason, Melissa Clark, Christina Clark Tuttle, and Emily Dwoyer. The Evaluation of Charter School Impacts. Final Report. *US Department of Education*, page 259, 2010.
- [71] J. Luis Guasch. *Granting and Renegotiating Infrastructure Concessions*. World Bank Institute Development Studies. The World Bank, March 2004.
- [72] J. Luis Guasch, Jean-Jacques Laffont, and Stéphane Straub. Renegotiation of concession contracts in Latin America. *International Journal of Industrial Organization*, 26(2):421–442, March 2008.
- [73] Herath Gunatilake, Carangal-San Jose, and Mary Jane F. Privatization Revisited: Lessons from Private Sector Participation in Water Supply and Sanitation in Developing Countries. Technical Report 115, 2008.

- [74] Kara Hanson, Nassor Kikumbih, Joanna Armstrong Schellenberg, Haji Mponda, Rose Nathan, Sally Lake, Anne Mills, Marcel Tanner, and Christian Lengeler. Cost-effectiveness of social marketing of insecticide-treated nets for malaria control in the United Republic of Tanzania. *Bulletin of the World Health Organization*, 81(4):269–276, 2003.
- [75] Eric A. Hanushek. The Failure of Input-Based Schooling Policies. *Economic Journal*, 113:113–485, 2003.
- [76] Oliver Hart. Incomplete Contracts and Public Ownership: Remarks, and an Application to Public-Private Partnerships. *The Economic Journal*, 113(486):C69–C76, 2003.
- [77] Justine Hastings. Heterogeneous Preferences and the Efficacy of Public School Choice. page 62, 2009.
- [78] Laurel Hatt, Ha Nguyen, Nancy Sloan, Sara Miner, Obiko Magvanjav, Asha Sharma, Jamil Chowdhury, Paul Dipika, Mursaleena Islam, Rezwana Chowdhury, and Wang Hong. Economic Evaluation of Demand-Side Financing (DSF) for Maternal Health in Bangladesh. *Economic Evaluation of Demand-Side Financing (DSF) for Maternal Health in Bangladesh. Bethesda, MD: Review, Analysis and Assessment of Issues Related to Health Care Financing and Health Economics in Bangladesh, Abt Associates Inc.*, 2010.
- [79] Health Research Institute. Build and Beyond: The revolution of Healthcare PPPs (PwC) | Public private partnership. Technical report, 2010.
- [80] Eva I. Hoppe, David J. Kusterer, and Patrick W. Schmitz. Public-private partnerships versus traditional procurement: An experimental investigation. *Journal of Economic Behavior & Organization*, 89:145–166, May 2013.
- [81] C. M. Hoxby and J. E. Rockoff. The Impact of Charter Schools on Student Achievement. page 52.
- [82] Caroline Hoxby. Peer Effects in the Classroom: Learning from Gender and Race Variation. Technical Report 7867, August 2000.
- [83] Caroline M Hoxby, Sonali Murarka, and Jenny Kang. How New York City’s Charter Schools Affect Achievement. page 85, 2009.
- [84] Chang-Tai Hsieh and Miguel Urquiola. The effects of generalized school choice on achievement and stratification: Evidence from Chile’s voucher program. *Journal of Public Economics*, 90(8-9):1477–1503, 2006.

- [85] Elise Huillery and Juliette Seban. Performance-Based Financing, Motivation and Final Output in the Health Sector: Experimental Evidence from the Democratic Republic of Congo. Technical Report 2014-12, October 2014.
- [86] Elisabetta Iossa and David Martimort. The Simple Microeconomics of Public-Private Partnerships. *Journal of Public Economic Theory*, 17(1):4–48, 2015.
- [87] Yi Jiang and Xiaoting Zheng. Private Sector Participation and Performance of Urban Water Utilities in China. *Economic Development and Cultural Change*, 63(1):155–189, 2014.
- [88] Jooseop Kim, Harold Alderman, and Peter Orazem. Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program. *World Bank Economic Review*, 13(3):443–65, 1999.
- [89] Yin-Fang Colin Parker Kirkpatrick, David Zhang. An Empirical Analysis of State and Private-Sector Provision of Water Services in Africa. *World Bank Economic Review*, 20(1):143–163, January 2006.
- [90] Michael Kremer and Alaka Holla. Improving Education in the Developing World: What Have We Learned from Randomized Evaluations? *Annual Review of Economics*, One:513–542, 2009.
- [91] Michael Kremer and Andrei Sarychev. Why Do Governments Operate Schools? *Working Paper*, 2000.
- [92] Bernardo Lara, Alejandra Mizala, and Andrea Repetto. The Effectiveness of Private Voucher Education: Evidence From Structural School Switches. *Educational Evaluation and Policy Analysis*, 33(2):119–137, June 2011.
- [93] Moritz Liebe and Allyson Pollock. The experience of the private finance initiative in the UK’s National Health Service. *Centre for International Public Health Policy, Edinburgh*, page 20, 2009.
- [94] Stephen S. Lim, Lalit Dandona, Joseph A. Hoisington, Spencer L. James, Margaret C. Hogan, and Emmanuela Gakidou. India’s Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities: an impact evaluation. *The Lancet*, 375(9730):2009–2023, June 2010.
- [95] Thomas Peyton Lyon, A. Wren Montgomery, and Dan Zhao. A Change Would Do You Good: Privatization, Municipalization, and Drinking Water Quality. *Academy of Management Proceedings*, 2017(1):10499, January 2017.

- [96] Bentley MacLeod. Anti-Lemons: School Reputation and Educational Quality. Technical Report qt3rc708kd, July 2009.
- [97] Tanya Marchant, Joanna Armstrong Schellenberg, Tabitha Edgar, Rose Nathan, Salim Abdulla, Oscar Mukasa, Hadji Mponda, and Christian Lengeler. Socially marketed insecticide-treated nets improve malaria and anaemia in pregnancy in southern Tanzania. *Tropical medicine & international health: TM & IH*, 7(2):149–158, February 2002.
- [98] Tonia Marek, Catherine C. O’Farrell, Chiaki Yamamoto, and Ilyse Zable. Trends and opportunities in public-private partnerships to improve health service delivery in Africa. Technical Report 33646, August 2005.
- [99] David Martimort and Jerome Pouyet. To build or not to build: Normative and positive theories of public-private partnerships. *International Journal of Industrial Organization*, 26(2):393–411, 2008.
- [100] David Martimort and Stéphane Straub. Infrastructure Privatization and Changes in Corruption Patterns: The Roots of Public Discontent. *Journal of Development Economic*, pages 69–84, 2009.
- [101] David McKenzie and Dilip Mookherjee. The Distributive Impact of Privatization in Latin America: Evidence from Four Countries. *Economía*, 3(2):161–218, 2003.
- [102] Liesbeth E. Meuwissen, Anna C. Gorter, Arnold D. M. Kester, and J. A. Knottnerus. Can a comprehensive voucher programme prompt changes in doctors’ knowledge, attitudes and practices related to sexual and reproductive health care for adolescents? A case study from Latin America. *Tropical medicine & international health: TM & IH*, 11(6):889–898, June 2006.
- [103] Liesbeth E. Meuwissen, Anna C. Gorter, and André J. A. Knottnerus. Impact of accessible sexual and reproductive health care on poor and underserved adolescents in Managua, Nicaragua: a quasi-experimental intervention study. *Journal of Adolescent Health*, 38(1):56.e1–56.e9, January 2006.
- [104] Jonathan N. Mills and Patrick J. Wolf. Vouchers in the Bayou: The Effects of the Louisiana Scholarship Program on Student Achievement After 2 Years. *Educational Evaluation and Policy Analysis*, 39(3):464–484, September 2017.

- [105] Alejandra Mizala and Florencia Torche. Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, 32(1):132–144, January 2012.
- [106] Manoj Mohanan, Sebastian Bauhoff, Gerard La Forgia, Kimberly Singer Babiarz, Kultar Singh, and Grant Miller. Effect of *Chiranjeevi Yojana* on institutional deliveries and neonatal and maternal outcomes in Gujarat, India: a difference-in-differences analysis. *Bulletin of the World Health Organization*, 92(3):187–194, March 2014.
- [107] Vincenzo Mollisi. The Impact of Public-Private Partnerships on productivity of the Italian District Heating industry. *Working Paper*, 2016.
- [108] A. Wren Montgomery, Thomas Peyton Lyon, and Dan Zhao. Not a Drop to Drink?: Drinking Water Quality, System Ownership, and Stakeholder Attention. *Research in the Sociology of Organizations*, 2018.
- [109] Jo-Ann Mulligan, Joshua Yukich, and Kara Hanson. Costs and effects of the Tanzanian national voucher scheme for insecticide-treated nets. *Malaria Journal*, 7:32, February 2008.
- [110] Karthik Muralidharan and Michael Kremer. Public and Private Schools in Rural India. in *Rajashri Chakrabarti and Paul Peterson, eds., School Choice International: Exploring Public-Private Partnerships*, 2009.
- [111] Adiel Mushi, Joanna R M Armstrong Schellenberg, Haji Mponda, and Christian Lengeler. Targeted subsidy for malaria control with treated nets using discount voucher system in Tanzania. *Health policy and planning*, 18:163–71, July 2003.
- [112] Patricia Navarro-Palau. Effects of differentiated school vouchers: Evidence from a policy change and date of birth cutoffs. *Economics of Education Review*, 58:86–107, June 2017.
- [113] Christopher Neilson. Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students. page 62, 2017.
- [114] Ha T. H. Nguyen, Laurel Hatt, Mursaleena Islam, Nancy L. Sloan, Jamil Chowdhury, Jean-Olivier Schmidt, Atia Hossain, and Hong Wang. Encouraging maternal health service utilization: an evaluation of the Bangladesh voucher program. *Social Science & Medicine (1982)*, 74(7):989–996, April 2012.

- [115] Francis Obare, Charlotte Warren, Timothy Abuya, Ian Askew, and Ben Bellows. Assessing the population-level impact of vouchers on access to health facility delivery for women in Kenya. *Social Science & Medicine*, 102:183–189, February 2014.
- [116] Francis Obare, Charlotte Warren, Rebecca Njuki, Timothy Abuya, Joseph Sunday, Ian Askew, and Ben Bellows. Community-level impact of the reproductive health vouchers programme on service utilization in Kenya. *Health Policy and Planning*, 28(2):165–175, March 2013.
- [117] Ministry of Health & Family Welfare. National Health Policy 2015 Draft. December 2014.
- [118] Kenya National Bureau of Statistics. Kenya Demographic and Health Survey 2008-2009. 2010.
- [119] Tae H. Oum, Nicole Adler, and Chunyan Yu. Privatization, corporatization, ownership forms and their effects on the performance of the world’s major airports. *Journal of Air Transport Management*, 12(3):109–121, May 2006.
- [120] Tae H. Oum, Jia Yan, and Chunyan Yu. Ownership forms matter for airport efficiency: A stochastic frontier investigation of worldwide airports. *Journal of Urban Economics*, 64(2):422–435, September 2008.
- [121] Harry Anthony Patrinos and Felipe Barrera-Osorio. The role and impact of public-private partnerships in education. Technical Report 47949, March 2009.
- [122] Sergio Perelman and Tomas Serebrisky. Measuring the Technical Efficiency of Airports in Latin America. *Utilities Policy*, 22:1–7, February 2012.
- [123] Carlos Pestana Barros and Peter U. C. Dieke. Performance evaluation of Italian airports: A data envelopment analysis. *Journal of Air Transport Management*, 13(4):184–191, July 2007.
- [124] Carlos Pestana Barros and Peter U. C. Dieke. Measuring the economic efficiency of airports: A Simar–Wilson methodology analysis. *Transportation Research Part E: Logistics and Transportation Review*, 44(6):1039–1051, November 2008.
- [125] Simon Porcher. Do Markets Reduce Prices? *EPPP Chair Working paper*, 2012.
- [126] Simon Porcher and Stephane Saussier. Public Versus Private Management in Water Public Services: Taking Stock, Looking Ahead. *Sorbonne Business School Working Paper*, 2018.

- [127] Jens K. Roehrich, Michael A. Lewis, and Gerard George. Are public–private partnerships a healthy option? A systematic literature review. *Social Science & Medicine*, 113:110–119, July 2014.
- [128] Cecilia Elena Rouse. Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program. *The Quarterly Journal of Economics*, 113(2):553–602, May 1998.
- [129] David S. Saal and David Parker. The impact of privatization and regulation on the water and sewerage industry in England and Wales: a translog cost function model. *Managerial and Decision Economics*, 21(6):253–268, April 2001.
- [130] Peter Sandiford, Anna C. Gorter, and M Salvetto. Vouchers for Health, Using Voucher Schemes for Output-Based Aid. *Public Policy for the Private Sector*, Viewpoint. 1-4., 2002.
- [131] Jean-Olivier Schmidt, Tim Ensor, Atia Hossain, and Salam Khan. Vouchers as demand side financing instruments for health care: a review of the Bangladesh maternal voucher scheme. *Health Policy (Amsterdam, Netherlands)*, 96(2):98–107, July 2010.
- [132] Jhih-Shyang Shih, Winston Harrington, William A. Pizer, and Kenneth Gillingham. Economies of Scale and Technical Efficiency in Community Water Systems. *Working Paper*, 2004.
- [133] Robert Soeters, Peter Bob Peerenboom, Pacifique Mushagalusa, and Célestin Kimanuka. Performance-based financing experiment improved health care in the Democratic Republic of Congo. *Health Affairs (Project Hope)*, 30(8):1518–1527, August 2011.
- [134] Adriana Tami, Juliet Mbatia, Rose Nathan, Haji Mponda, Christian Lengeler, and Joanna R. M. Armstrong Schellenberg. Use and misuse of a discount voucher scheme as a subsidy for insecticide-treated nets for malaria control in southern Tanzania. *Health Policy and Planning*, 21(1):1–9, January 2006.
- [135] Khushbu B. Thadani. Public Private Partnership in the Health Sector: Boon or Bane. *Procedia - Social and Behavioral Sciences*, 157:307–316, November 2014.
- [136] Lourdes Trujillo, María Manuela González, and Juan Luis Jiménez. An overview on the reform process of African ports. *Utilities Policy*, 25(C):12–22, 2013.

- [137] De Walque, Damien B. C. M, Paul Jacob Robyn, Hamadou Saidou, Gaston Sorgho, and Maria Ward Steenland. Looking into the performance-based financing black box: evidence from an impact evaluation in the health sector in Cameroon. Technical Report WPS8162, August 2017.
- [138] Peter F. Wanke and Carlos Pestana Barros. Public-private partnerships and scale efficiency in Brazilian ports: Evidence from two-stage DEA analysis. *Socio-Economic Planning Sciences*, Complete(51):13–22, 2015.
- [139] Patrick Wolf, Babette Gutmann, Michael Puma, Brian Kisida, Lou Rizzo, Nada Eissa, and Matthew Carr. Evaluation of the DC Opportunity Scholarship Program: Final Report. NCEE 2010-4018. *National Center for Education Evaluation and Regional Assistance*, January 2010.
- [140] Mitchell D. Wong, Karen M. Collier, Rebecca N. Dudovitz, David P. Kennedy, Richard Buddin, Martin F. Shapiro, Sheryl H. Kataoka, Arleen F. Brown, Chi-Hong Tseng, Peter Bergman, and Paul J. Chung. Successful schools and risky behaviors among low-income adolescents. *Pediatrics*, 134(2):e389–396, August 2014.
- [141] Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. 2012.
- [142] The World Bank. The business of health in Africa : partnering with the private sector to improve people’s lives. Technical Report 44143, January 2008.
- [143] The World Bank. Public-private partnerships : reference guide version 2.0. Technical Report 90384, January 2014.
- [144] The World Bank. World Development Report 2018: Learning to Realize Education’s Promise. Technical report, 2017.
- [145] Omer Zang, Sebastien Djienouassi, Gaston Sorgho, and Jean Claude Taptue. Impact of performance- based financing on health-care quality and utilization in urban areas of Cameroon. *African Health Monitor*, (20):5, October 2015.
- [146] Ron Zimmer, Brian Gill, Kevin Booker, Stephane Lavertu, Tim R. Sass, and John Witte. *Charter Schools in Eight States: Effects on Achievement, Attainment, Integration, and Competition*. RAND Corporation, 2009.
- [147] Ron Zimmer, Brian Gill, Kevin Booker, Stéphane Lavertu, and John Witte. Examining charter student achievement effects across seven states. *Economics of Education Review*, 31(2):213–224, April 2012.