Performance-Based Pay: Investigating Its International Prevalence in Light of National Contexts

Aoki, N., & Rawat, S. (2020). Performance-based pay: Investigating its international prevalence in light of national contexts. *The American Review of Public Administration*, 50(8), 865-879.

ABSTRACT

Performance-based pay is a popular technique in public management, but its international prevalence remains poorly evidenced. This study asks to what extent countries are using performance-based pay and how local contexts matter to its adoption status. Focusing on the education sector, we report on the varying degrees to which teacher appraisals have been used and linked with monetary rewards across countries as of 2012. We found that performance-based pay tended to be used more in less liberal economies, in cultures with a lower degree of uncertainty avoidance and a higher degree of individualism, in those with more decentralized educational systems, and surprisingly, in places where teachers exert greater influence. Our results suggest that some country contexts matter to the prevalence pattern of performance-based pay, and they offer policy makers some insights into the debate on whether a managerial technique like this, used in other countries, is transferrable to their own national context.

Keywords: performance appraisal, isomorphism, education, post-NPM, PISA

Performance-Based Pay: Investigating Its International Prevalence in Light of National Contexts

Introduction

Performance appraisal is arguably a popular technique in public management. It is 'a methodology and set of procedures for rating the work performance of individuals according to objective standards and criteria applied uniformly across one or several organizations' (OECD, 2005, p. 20). Public organizations use it to control and motivate their employees, among other purposes (Behn, 2003). For appraisals to be effective, the principles of scientific management (Taylor, 1911), principal-agent theory (see Eisenhardt, 1989, for an overview), and the Expectancy Theory (Vroom, 1964) all suggest that appraisal results should be linked with rewards. So-called *performance-based pay* in particular links appraisals with monetary rewards. Based on the Organization for Economic Cooperation and Development (OECD) survey conducted in 2003, OECD (2005) concluded that performance-based pay was used, or at least considered, for public servants among two-thirds of OECD members, although the degree to which the technique was used varied considerably across nations, departments, and positions within governments.

Going beyond OECD membership, this study asks on a global scale to what extent countries are using performance-based pay and how national contexts matter to its adoption status. For some time, public management scholars have been asking one question or another regarding the international prevalence of managerial techniques. When so-called new public management (NPM) was still in fashion, arguably as a global paradigm, they debated whether or not countries were adopting NPM-inspired techniques, and hence, whether countries were converging or diverging in what they did to run their public organizations (Hood, 1991, 1995). In these debates, performance-based pay was often, albeit not always, regarded as an NPM-inspired technique (Aoki, 2015; Butterfield, Edwards, & Woodall, 2004; Park & Berry,

2014). In the post-NPM era, scholars have been debating over what has happened to NPM and whether it has been fully or partially replaced by other forms of management (Dunleavy, et al., 2005; Pollitt, 2016). The debate remains unsettled, due to limited empirical evidence on the prevalence of NPM-inspired practices.

Our question regarding the prevalence of performance-based pay in relation to country conditions may be of importance, particularly to some countries where the decision to promote or adopt the technique can be made nationally; see, for example, national initiatives to implement it in Malaysia (Goh, 2012), Qatar (Romanowski, 2014), and Singapore (Liew, 2012). In those countries, policy makers may be concerned with whether the technique, which has already been used abroad, is transferrable to their own country. This is a relevant concern today, as policy makers rely increasingly on policy transfer from abroad; however, 'inappropriate transfer' can occur when a certain policy is transferred without accounting for the 'economic, social, political and ideological contexts' of the borrowing country (Dolowitz & Marsh, 2002, p. 17) – one reason NPM-inspired practices in the West are seen to be not exportable to certain countries (Sozen & Shaw, 2002; Sulle, 2010). This study generates some insights into how transferable performance-based pay is from abroad, considering national contexts.

In the past, studies have investigated the factors influencing the prevalence patterns of administrative reforms in government. Moon and deLeon (2001), for example, examined the extended reinvention adoption index, which measures the degree to which 12 managerial innovations in line with reinvention themes were adopted by municipal governments in the United States (US) as of 1997. They found that the adoption of these innovations was associated with the reinvention values and political ideologies of chief administrators, the population size of municipalities, economic conditions, the presence of labour unions, and the existence of a council or city manager system within a municipality. Meyer and

Hammerschmid (2010) focused on 27 European countries and found that certain aspects of national culture – power distance and uncertainty avoidance – were negatively correlated with the degree of decentralization, while individualism was positively associated with it. The authors also found that the degree of decentralization tended to be higher in the Anglo-Saxon administrative tradition. Aoki (2019), using international data on public educations for 15-year-olds, revealed that performance-based human resource management tended to be used more extensively in less accountable and less wealthy economies, and that among 65 economies, the most intensive users included Tunisia, Indonesia, and Malaysia.

Instead of focusing on the whole of government, we focus on a specific sector – education – while aiming to draw implications for public management reforms at large. The focus on a single sector is reasonable, considering a finding by the OECD (2005) that the prevalence of performance-based pay depends on the area of public service. In education, a properly designed merit-based pay is seen as a means of aligning teachers' incentives to act with a school's mission (Ritter & Jensen, 2010; Ritter, Maranto, & Buck, 2009). Studies have found a positive impact of performance-based pay on students' test scores in India (Muralidharan & Sundararaman, 2011) and the US (Cooper & Cohn, 1997), although Dee and Keys (2004) found that a merit-based system in the US state of Tennessee had 'mixed success' with rewarding good teachers (p. 471). While empirical findings such as these still remain mixed and limited, the practice of paying teachers on the basis of performance has been adopted in many countries, as is empirically shown by the data in this study.

We make the most of the internationally comparable dataset produced by the Organization for Economic Cooperation and Development (OECD) in 2012, which offers information regarding the use of performance-based pay in the education of 15-year-olds. As in studies by Moon and deLeon (2001) and Meyer and Hammerschmid (2010), we perform associational analyses to address our questions regarding prevalence. Unlike Aoki (2019),

which used the same dataset but performed bivariate correlational analyses, we use multiple regression analyses and test whether a larger number of contextual variables matter to the prevalence pattern of performance-based pay across countries.

The rest of this article is structured as follows. The next section presents our hypotheses. The third and fourth sections explain our empirical strategy and discuss the results, respectively. The final section draws implications for research and policy discourse on public management reforms.

The Pattern of Prevalence by National Context – Hypotheses

Table 1 lists our hypotheses regarding the prevalence pattern of performance-based pay, each of which is elaborated below.

Economic Ideology

We hypothesized that performance-based pay is used more in countries that exhibit a higher commitment to market ideology (**H1** in Table 1). This hypothesis is in line with Christensen and Lægreid (2011a, 2011b), who proposed that economic ideology is one of the environmental factors that determine the adoption of administrative reforms in general. In particular, market ideology is deemed relevant, because NPM was originally proposed in some, albeit not all, places by those who were disposed toward *neoliberalism* or the *New Right* back in the 1980s (Tolofari, 2005) – an ideology that upholds a 'commitment to reduce the role of the state, compel individuals to become more self-sufficient and expose public services as far as possible to market forces' (Elcock, 1995, p. 34). Margaret Thatcher grieved over the fact that the British government had grown in size despite economic recessions, because it had been isolated from market pressures (Thatcher, 1993). In New Zealand, NPM reforms were inspired by the public choice school of thought (Hood, 1991), which essentially blamed public sector inefficiency on a lack of market mechanisms (Niskanen, 1971). In the US, performance-based pay in federal agencies was promoted under the Clinton

administration's National Performance Review initiative, aimed at reinventing the government (Dilulio, Garvey, & Kettl, 2001; Riccucci & Thompson, 2008) through neoliberal and entrepreneurial modes of governance (Steger & Roy, 2010).

However, we offer a caveat about generalizing from the above observations. This is because, first, NPM (or aspects of it) were endorsed as well by Labour administrations in Australia (Pollitt, 2016) and by Scandinavian welfare states (Green-Pedersen, 2002). As far as education is concerned, performance-based pay for teachers was promoted under the Obama administration (Maranto & McShane, 2012; Ravitch, 2014; Smarick, 2011), whose top priorities included social welfare policies (Mettler, 2010). Second, views regarding NPM do not make clear whether they are referring to NPM as an ideology or as a set of practices, and third, if the latter is the case, it is unclear what practices may be considered part of NPM (Hood, 1995; Pollitt, 2016); moreover, NPM-inspired reforms do not necessarily include performance-based pay. Fourth, performance-based pay had been in use even before NPM emerged (Podgursky & Springer, 2007). Finally, the economic ideology of a government that adopted performance-based pay in the first place may have been replaced by succeeding administrations disposed toward a different economic ideology. Given all of these considerations, it is worth looking at what the empirical evidence has to say regarding the association between market ideology and the status of performance-based pay (**H1-a** and **-b**).

Political Regime

For this paper, we considered whether the type of political regime makes a difference and hypothesized that performance-based pay would be used more in liberal democracies than in less liberal political regimes (**H2-a**). This hypothesis was inspired by the views of a large number of scholars that the NPM-inspired managerial movement began in liberal democracies (Hartley, 2003; Lee & Haque, 2006; Peters & Pierre, 1998). This perception is understandable for two reasons: first, the countries believed to be the originators of NPM

reforms – namely, the US, the UK, Australia, and New Zealand – are considered western liberal democracies. Second, performance-based pay might be used more in liberal democracies because the technique can be considered a means of making public administrators accountable to the public. However, there is also an argument that the technique fits well with illiberal political culture, because it can be considered a means of restricting individual autonomy and liberty, and of coercing employees to work towards standards set by authorities (Aoki & Tay, 2015). In addition, a bivariate correlational analysis in Aoki (2019) showed a negative correlation between liberal democracy and performance-based human resources management. Accordingly, we also proposed a competing hypothesis (**H2-b**).

Income Level

We postulated that the adoption status of performance-based pay might depend on the income level of a country (**H3**). A widely accepted narrative states that NPM-inspired managerial reforms in general were exported from higher- to lower-income countries via international organizations. Despite a global campaign to promote NPM in low-income countries, there is reason to believe that performance-based pay is more common in high-income countries than in low-income countries (**H3**), because low-income countries are inadequately prepared to implement performance appraisals (Smith, 1995), which call for resources to train supervisors to do appraisals, to install electronic systems, to produce all of the paperwork, and to fund financial rewards. After all, performance-based pay is an expensive programme that takes both financial and human resources to sustain. In fact, countries have been known to fail to institutionalize and sustain performance-based pay due to inadequate resources; see, for example, the case of the Zimbabwean Public Service Commission (Pretorius & Ngwenya, 2008; Zvavahera, 2014). Low-income countries are also frequently characterized by corruption and clientelism (see, for example, Moe & Wiborg, 2017), which may contribute to

a reluctance to accept the idea of rewarding employees based on performance and merit. Moreover, due to the lack of trust in government institutions generally found to be associated with corruption (Chang & Chu, 2006; Morris & Klesner, 2010), employees might doubt whether appraisals would be fair, resulting in a lack of support for performance-based pay.

Isomorphic Pressures

What Meyer and Rowan (1977) termed *isomorphic pressure* has been proven to drive business organizations to adopt certain policies (Ando, 2015; Cao, Li, & Wang, 2014; Montes & Jover, 2004) and has contributed to innovations in the public sector (Teodoro, 2011). Borrowing the classification of isomorphic pressures proposed by DiMaggio and Powell (1983), Pollitt (2001) observed that professionals with similar values and norms working across boundaries created *normative* isomorphic pressure, which promoted the spread of public management reforms. There is also *mimetic* isomorphic pressure, which drives countries under conditions of uncertainty to mimic seemingly leading countries. *Coercive* isomorphic pressure can occur when international organizations impose a certain reform on national governments as a condition for receiving assistance. In line with these arguments, we examined whether the adoption status of performance-based pay depends on isomorphic pressures (**H4**).

Our hypothesis is also in accord with empirical findings. Frumkin and Galaskiewicz (2004) found that the effects of normative isomorphic pressures on the degree of decentralization of hiring processes, the formalization of hiring documents, and functional departmentalization were higher for governments than for non-profit or for-profit organizations. Villadsen, Hansen, and Mols (2010) found that technological uncertainty – associated with an organization's inability to accurately forecast the technological requirements for future production – drove Danish municipality managers to engage in mimicking behaviours (i.e. comparing themselves to other municipalities and departments,

and gathering knowledge from them) when they made contracting-out decisions. Pitts et al. (2010) used the degree to which school superintendents in Texas schools interacted with other superintendents as a measure of normative isomorphic pressure, and their analysis revealed that this pressure had a positive impact on the status of diversity promotion and management programmes in the schools.

National Culture

Building on Meyer and Hammerschmid (2010), we also posited that national culture matters to the adoption status of performance-based pay (**H5-8**). This hypothesis is also in line with Alesina and Giuliano (2015), who believed in a close endogenous relationship between culture and institutions in general, and with Pollitt and Bouckaert (2011), who found that some cultures are more conducive to administrative reforms than others. Following Meyer and Hammerschmid (2010), we adopted Hofstede's definition of national culture (Hofstede, Hofstede, & Minkov, 2010, p. 6): 'the collective programming of the mind which distinguishes the members of one group or category of people from another.' Hofstede (1983) originally argued that national culture has four dimensions: (i) *power distance*, defined by 'the degree to which the less powerful members of a society accept and expect that power is distributed unequally', (ii) *individualism*, which is 'a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families', (iii) *masculinity*, defined as 'a preference in society for achievement, heroism, assertiveness and material rewards for success', and (iv) *uncertainty avoidance*, 'the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity.

Separate hypotheses can be conceived for these different dimensions of culture. In a higher power distance culture, individuals are less resistant to being controlled by their superiors through standard performance targets, and hence, it can be supposed this culture would be more conducive to the use of performance appraisals and performance-based pay

(**H5**). Our hypothesis that national culture matters was inspired by Aoki and Tay (2015), who argued that performance appraisals in Singapore fit with the society's relatively power-distant culture, which the technique itself helped to create. Furthermore, performance-based pay for individual teachers and school principals might fit more with national cultures with a higher degree of individualism; hence, we posit that it would be used more in such cultures (**H6**). This is because management in highly individualistic cultures might be expected to recognize the achievements of individuals, whereas in less individualistic cultures, compensation policies would be expected to focus on the accomplishments of social units, conformity to organizational values, or seniority (which reinforces a long-term relationship between employees and organizations) (Gomez-Mejia & Welbourne, 1991). Likewise, a society that Hofstede terms 'masculine' tends to be competitive, and hence, school staff would be more willing to accept performance appraisals, which induces competition among individuals for limited rewards. The idea of incentivizing workers with rewards also fits with the emphasis on material gain and achievements in masculine cultures. Accordingly, we posit that performance-based pay is used more in masculine cultures (**H7**). This hypothesis is in line with the empirical evidence for a congruence between masculine culture and merit-based pay in Newman and Nollen (1996). Finally, we hypothesized that workers in cultures high in uncertainty avoidance would prefer to work towards established performance targets to mitigate uncertainty; hence, performance appraisals might be used more in national cultures with a high degree of uncertainty avoidance (**H8-a**). However, it is also possible that uncertainty avoidance may hamper the adoption of performance-based pay (**H8-b**). because employees in such cultures may prefer fixed base pay rather than pay based on variable components (Gomez-Mejia and Welbourne, 1991). In the education sector, Bowen et al. (2015) found that teachers-in-training were more risk-averse than business and law students. Accordingly, Bowen et al. (2015) recommended that policy makers be aware of the fact that

teachers might oppose reform changes, such as the introduction of performance incentives. While this finding is from the Midwest region of the US, we posit that teachers' risk averseness and related uncertainty avoidance vary systematically across national cultures, which, in turn, has an impact on the degree to which teachers are likely to oppose reforms, including the adoption of performance-based pay.

Administrative Structure

Christensen and Lægreid (2011a, 2011b) argued that administrative reform outcomes can be explained by the dynamic interplay between (i) external environmental factors and (ii) the cultural norms and internal structures of public organizations. We considered structure in terms of the degree to which organizations are decentralized and proposed two competing hypotheses. On one hand, a centralized system may be more effective at ushering in radical reform or 'controversial practices' such as performance-based pay; hence, we posited that performance-appraisal pay would be used more with centralized systems (**H9-a**). This proposition is consistent with Pollitt and Dan (2013), who found that, compared to the incremental decision-making processes in, say, Germany and Norway, the centralized political system in the UK facilitated the implementation of large-scale NPM-type reforms. On the other hand, the opposite hypothesis is also possible (**H9-b**): decentralization and performance appraisals may be promoted in tandem because NPM-inspired reforms often involve a 'process of dismantling the logical structure of old-fashioned bureaucracy to give greater discretionary freedom to managers at all levels' (Kane & Patapan, 2006, p. 711), and decentralization creates a new challenge of preventing the abuse and misuse of new responsibility by managers who are released from 'strict obedience to rules and the commands of superiors' (p. 715).

Teacher Group Influence

Our last hypothesis was education-specific and states that performance-based pay tends to be

used more in countries with weaker teacher influence (**H10**). This hypothesis was inspired by the fact that teachers have been opposing performance-based pay in a number of countries, ranging from Scotland (Reeves, et al., 2002) to Japan (Aoki, 2012), to the US (Moe, 2011; 2012) and Chile (Avalos & Assael, 2006). In general, teachers oppose the use of performance measures in education because quantitative performance measures poorly reflect the complexity and breadth of teachers' work (Ploom & Haldma, 2013; Storey, 2002), because conducting appraisals can be an additional burden (Pretorius & Ngwenya, 2008), and because appraisals can result in a 'tick-box mentality', damaging employees' trust in their supervisors, and distorting the focuses and priorities of education (Forrester, 2011, p. 8). Teachers resist formally through unions; Moe and Wiborg (2017) synthesized the studies of teacher unions in 11 countries and found that teacher unions were 'political adversaries' in the era of performance-based reform in all of these places (p. 17). Teachers also resist informally by means of absenteeism, increased instrumentalism, and weak compliance in their jobs (Mather & Seifert, 2011). Moon and deLeon (2001) found that the involvement of a labour union was significantly and negatively associated with initiatives to reinvent government practices by US municipal governments, although this study did not investigate performance appraisals in particular. Frumkin and Galaskiewicz (2004), while using union pressure as a control in their study, found it to be negatively associated with the degree of government decentralization.

<Insert Table 1 about here>

Empirical strategy

To explore the associations between national contexts and the degree to which performance-based pay is adopted, we performed regression analyses using a cross-country dataset. Our dependent variables were derived from the school questionnaires in the 2012 Programme for International Student Assessment (PISA), administered by the OECD in 65 countries and subnational economies (for the sake of simplicity, both are hereafter referred to as

'countries'), which were reduced to 51 countries in the final datasets after deleting observations with missing values. The PISA used both student questionnaires targeted at representative 15-year-olds in each economy and school questionnaires answered by their school principals, selected through two-stage sampling. Of these, we used the responses of the principals of public schools. Readers are reminded that all school information derived from principals' responses was weighted in the computation of country-level statistics (presented below) by merging school and student datasets and by applying final student sampling weights made available in the PISA dataset.

Our dependent variables represent the weighted proportion of public schools whose appraisals were linked with pay to different degrees in various countries. They were derived from school principals' responses to questionnaires that asked to what extent appraisals of, or feedback to, teachers (hereafter, simply referred to as 'appraisals') had directly led to (i) a change in salary or (ii) a financial bonus or other kind of monetary reward (hereafter, simply a 'bonus'). School principals were asked to select from among four items: no change, a small change, a moderate change, or a large change. We aggregated school principals' responses to this question at the country level to generate six dependent variables, listed in Table 2 with descriptive statistics. SALARY1 and BONUS1 are the within-country percentages of school principals who answered 'a small change' or more in terms of salary or a bonus, respectively. These two variables represent the degree to which performance-based pay had been adopted in the education system for 15-year-olds within a country, regardless of the extent of changes in pay – the information of our primary interest. The coding of these variables circumvents the problems associated with differing interpretations of the terms 'small,' 'moderate,' and 'large' changes, which is not the case for the other dependent variables that we tried, which are as follows: SALARY2 and BONUS2 are the country-level percentages of school principals who answered that appraisals were linked with 'a moderate change' or more.

SALARY3 and BONUS3 are the within-country percentages of school principals who answered, ‘a large change’. These two sets of dependent variables (SALARY2, BONUS2 and SALARY3, BONUS3) help to clarify whether the within-country percentages of performance appraisals linked with specific degrees of change depend on country contexts.

We made the most of existing international datasets for the country-context variables. For economic ideology, we used the 2012 Economic Freedom of the World Index, published by the Fraser Institute (2017), labelled EFWI in Table 2. This index measures ‘the degree to which the policies and institutions of countries are supportive of economic freedom’ (Gwartney, Lawson, & Hall, 2014, p. v), and it is based on the notion that economic freedom entails ‘personal choice’, ‘voluntary exchange coordinated by markets’, ‘freedom to enter and compete in markets’, and ‘protection of persons and their property from aggression by others’ (Gwartney, Lawson, & Hall, 2014, p. 1). Ranging from zero to ten, the index captures the degree of economic freedom according to: (i) the size of the government, (ii) the legal system and security of property rights, (iii) sound money (i.e. the consistency of monetary policy/ institutions, along with long-term price stability and the ease with which other currencies can be used), (iv) the freedom to trade internationally, and (v) regulation of credit, labour and business.

For a measure of the type of political regime, we used the World Bank’s Governance Indicator, ‘voice and accountability’, labelled VOICE in Table 2. This indicator measures ‘perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media’ (Kraay, Kaufmann, & Mastruzzi, 2010, p. 3). This indicator was used as a proxy for liberal democracy. To measure the income levels of countries, we utilized the log term of the gross national income (GNI) per capita in USD, labelled GNPLN in Table 2.

We used two variables as proxies for isomorphic pressure, one of which was OECD membership. Because the OECD extensively engages in knowledge sharing and dissemination in the area of governance and public administration reforms, we posited that OECD members are under normative isomorphic pressure to assimilate and adopt similar practices. Another measure is the Fragile States Indices, labelled EXTER in Table 2. This indicator captures the degree of external pressure on a country in terms of forms of foreign assistance, the presence of peacekeepers, the presence of UN missions, foreign military intervention, sanctions, and credit rating (The Fund for Peace, 2017). This indicator can serve as a proxy for coercive isomorphic pressure because these interventions happen when a state 'fails to meet its international or domestic obligations' (The Fund for Peace, 2017), regardless of its will. For national culture, we follow Meyer and Hammerschmid (2010) and use Hofstede's (1983) cultural dimensions of (i) power distance (POWER), (ii) individualism (INDIVI), (iii) masculinity (MASCU), and (iv) uncertainty avoidance (AVOID).

As a measure of administrative structure, we used school principals' responses to the PISA questionnaire identifying who bore considerable responsibility for the following 12 tasks at their schools: (i) selecting teachers for hire, (ii) firing teachers, (iii) establishing teachers' starting salaries, (iv) determining teachers' salary increases, (v) formulating the school budget, (vi) deciding on budget allocations within the school, (vii) establishing student disciplinary policies, (viii) establishing student assessment policies, (ix) approving students for admission to the school, (x) choosing which textbooks are used, (xi) determining course content, and (xii) deciding which courses are offered. School principals chose (ticked off) as many items as appropriate from among the following list: (a) principal, (b) teachers, (c) school governing board, (d) regional or local education authority, and (e) national education authority. To create a country-specific administrative structure variable, we first generated a new variable, referred to as (f) 'school responsibility', and coded this variable as one if

school principals ticked at least one of the school-level stakeholders (a-c) and otherwise as zero. We then created task-specific scores for each school by dividing this school responsibility variable (f) by the sum of (f), (d), and (e), where (d) and (e) were coded as one if school principals ticked, 'regional or local education authority' and 'national education authority', respectively, and otherwise as zero. We then generated each school-specific structure score by taking the means of the 12 task-specific scores and computing the within-country averages of the school-specific structure scores. When a school principal's answers with regard to all of the stakeholders amounted to zero, we treated the response as missing, because it indicated an unrealistic situation where no one was responsible for any task.

Finally, the influence of teachers, labelled TEACHR in Table 2, was derived from school principals' responses to the PISA question: 'Regarding your school, which of the following bodies exert a direct influence on decision making about staffing, budgeting, instructional content, and assessment practices?' We used this question from the 2009 PISA because it was not asked in the 2012 PISA, and other international data on teacher union strength for 2012 did not have sufficient country coverage. To create TEACHR, we first created school-specific measures by averaging four area-specific measures, each coded as one if 'teacher groups (e.g. Staff Association, curriculum committees, trade union)' exerted a direct influence, respectively, on staffingbudgeting, instructional content, and assessment practices. We then took the country average of these school-specific measures.

In addition to the aforementioned variables for testing our hypotheses, we controlled for the percentage of schools located in cities with a population of 100,000 or above (CITY), the average student-teacher ratio (RATIO), and the funding schools receive from non-government sources (NONGOV) in a typical school year. After dropping 14 countries due to missing values, our data contained information for 51 countries. We checked for multicollinearity between independent variables based on the correlation coefficients shown

in Table 3 and the variance inflation factor (VIF) for each independent variable. While some of the correlation coefficients were high, at above 0.7, they were nonetheless below the tolerance threshold level of 0.8 (Rivenbark et al. 2019; Studenmund 2014). All the VIFs were below the tolerance threshold level of 10 (Chatterjee & Hadi 2006; O'Brien, 2007), with a mean value of 2.93, ruling out a severe degree of multicollinearity.**<Insert Table 2 and Table 3 about here.>**

Results

Before presenting the results from the regression analyses, it is worth investigating the extent to which teacher appraisals were linked with pay across the 51 countries in our final dataset. Figures 1 and 2 show the weighted percentage distributions of schools whose teacher appraisals led to different degrees of change in salaries and bonuses. Notice that a large number of countries were using performance-based pay, yet the extent of usage varied considerably. Interestingly, no teacher appraisals were linked with 'a large change' in salaries and bonuses in countries considered to be the 'leading-edge' – namely, the US, the UK, New Zealand, and Australia. In the US, 69.4% of the public schools had teachers whose appraisals were not linked to any change in salary or bonus. This is rather surprising, considering federal and state efforts to promote the technique (Goff, Goldring, & Canney, 2016; Holbein & Ladd, 2017; Muralidharan & Sundararaman, 2011) and the view that unlike the fields of medicine and law, the field of teaching in the US is under-professionalized, which makes it susceptible to the externally imposed "factory-style rationalization" (p. 40) of schooling inspired by scientific management (Mehta, 2013). This finding might rather reflect the fact that teacher unions in US are so powerful that they have been successfully blocking reformers' attempts to introduce performance-based pay (Moe, 2012; 2019); even faced with irresistible pressure from the popularity of accountability reforms throughout the 1990s, teachers there were able to sabotage the effectiveness of performance measures that had been

put in place by taking part in their design and trying to "water down any components they found threatening" (Moe, 2012, p. 20). In contrast, among the highest countries in terms of the weighted proportion of schools whose teacher appraisals were linked to 'a large change in salary' were Jordan (34.4%), Thailand (30.5%), and the Slovak Republic (29.4%). For bonuses, this figure was the highest for Singapore, at 89.5%. In Jordan, Thailand, the Slovak Republic, and Montenegro, teacher appraisals were unequivocally perceived to be linked to both bonuses and salaries, and they ranked among the top 10 in both rankings. On the other hand, in countries like Singapore and Poland, only a specific form of performance-based pay was prevalent, that is, appraisals linked to bonuses. In the middle of the spectrum were countries like Peru and the Czech Republic, where appraisals were perceived to result in small-to-moderate changes in both salaries and bonuses.

<Insert Figure 1 and Figure 2 about here>

As for the estimates from the regressions, Tables 4 and 5 show the results for salaries and bonuses, respectively, and for models with different dependent variables. The coefficients represent the change in the dependent variable associated with a one-unit change in the predictor variable, when other variables are held constant.

Starting from the results for salaries in Table 4, the voice and accountability indicator is statistically significant across all models, but the sign is negative; thus, we can conclude with confidence ($p < .01$) that the use of performance-based pay in education systems for 15-year-olds (SALARY1) tends to be higher in less liberal political regimes. The use of performance-based pay linked to moderate or higher changes in salary (SALARY2) is negatively associated with the indicator ($p < .01$), and so is performance-based pay that results in large changes (SALARY3), albeit at the 10% significance level. These results may be surprising, considering that performance appraisal can be a means of enhancing accountability in government, which is a virtue in liberal democracies. However, this result

could be due to the fact that performance appraisal has been used more in less liberal democracies to control public servants, not necessarily to promote democratic accountability. Another reason may be that less liberal democratic regimes tend to have more power and authority to usher in policy changes and introduce a contentious policy like performance-based pay. For bonuses in Table 5, the voice and accountability indicator is statistically significantly associated with the degree to which appraisals were linked to a moderate or larger change in bonus (BONUS2).

We did not find statistical significance for the OECD as a proxy for isomorphism; however, we did find that external intervention is negatively associated with the use of performance-based salary (Model I). This result is inconsistent with our hypothesis. In hindsight, the reason could be that the indicator reflects foreign interventions in forms unrelated to education, such as the presence of UN missions and foreign military intervention.

As for culture variables, we found a statistically significant and negative association between uncertainty avoidance and performance-based salaries ($p < .01$) and bonuses ($p < .1$), which is somewhat in line with the argument made by Bowen et al. (2015) mentioned earlier. We found a statistically significant and negative association between individualism and performance-based salary ($p < .05$). This is contrary to our hypothesis, and it could have something to do with the fact that performance appraisals essentially control individual behaviours and promote an audit culture in exchange for individual freedom, which might be unwelcome in a culture with a high degree of individualism. Finally, we did not establish any statistical significance for the GNI and the EFW index, and this could be because after all these years of promoting performance-based pay around the world, the prevalence pattern has penetrated to countries with diverse political and economic regimes.

While the association between teacher group influence and performance-based salary is not statistically significant, we surprisingly find a positive and statistically significant ($p < .05$) association between teacher group influence and BONUS1. That is, the greater the teacher-group influence becomes, the more performance-based the bonuses are. In hindsight, this could have something to do with the ways in which the variable teacher-group influence was operationalized; the variable was based on teachers' influence on staffing, budgeting, instructional content, and assessment practices at the school level, and this may not necessarily represent the *organized* influence of teacher unions at the higher levels of school administration or government, where decisions regarding the introduction of performance-based pay are usually made. It is also possible that the attitudes of the teachers working on the ground towards performance-based pay are not as negative as conventional wisdom suggests, as Ballou and Podgursky (1993) found in the US. The results for administrative structure are consistent with our hypotheses regarding both SALARY1 ($p < .1$) and BONUS1 ($p < .05$); the utilization of performance-based pay tends to be higher in more decentralized countries. As discussed above, this could be because decentralization necessitates the monitoring of decentralized agents, and in this regard performance-based pay can provide a means of managing them.

<Insert Table 4 and Table 5 about here>

Discussion and Conclusion

Before proceeding to draw implications from our findings, a few caveats are worth noting. This study relied on information based on school principals' opinions, which may suffer from individual biases. As noted earlier, we are aware of the limitations of the culture variables, which leaves an opening for future research to explore more measures. We used OECD membership and external interventions as measures of isomorphic pressure; future researchers could try other measures. Furthermore, readers are reminded that our purpose did

not extend to investigating dynamic causal impacts, that is, how country contexts hamper or facilitate the adoption of performance-based pay; it was limited to addressing the question of prevalence as far as the year 2012 was concerned, and our data and empirical model do not address causality. Given the fact that, to date, empirical investigations of the global prevalence of performance-based pay in all areas of public services have remained limited, the findings of this study in the education sector offer a starting point for discussions to revisit debates concerning the overall global prevalence of performance-based pay. Although a single-year snapshot does not offer an answer in terms of dynamic processes of convergence and divergence over time, the data have shown that countries differ considerably in the extent to which performance-based pay is used and linked with monetary rewards in their education sector. Scholars have also debated over whether NPM has been replaced by alternative governance models; our findings do not indicate whether or not an NPM *ideology* remains, but as far as NPM-type *practices* are concerned, our results show that performance-based pay, often considered to be such a practice, is used extensively in some countries in their education sector. Future research can investigate how these findings may be applicable to other areas of public services.

As noted earlier, while the impact of performance-based pay on educational outcomes is still under empirical scrutiny, the technique has been widely considered by national policy makers in different parts of the world. Our results offer some insights to policy makers who may be wondering whether a technique like this, used in other countries, is transferrable to their own national context. Besides the fact that this technique has been institutionalized in a variety of countries, the results from the regressions suggest that liberal democracies do not champion the use of performance-based pay; rather, it tends to be used more in less liberal political regimes. We did not find support for the hypothesis that income levels matter to the usage of the technique, and the perception that it is being used in seemingly leading-edge

countries may be outdated. To put it differently, the type of political regime and a country's income level are not convincing justifications for an inability to institutionalize this technique.

Our results also offer some implications for existing theories concerning the adoption of administrative reforms in general. As noted earlier, Christensen and Lægreid postulated that administrative reform outcomes can be explained by the dynamic interplay between (i) external environmental factors and (ii) the cultural norms and internal structures of public organizations. We derived our hypotheses from their proposition that economic ideology is a part of the former. Although our analysis did not generate support for the influence of economic ideology, our results do support our cultural hypothesis for uncertainty avoidance, where we find a statistically significant negative association between uncertainty avoidance and performance-based bonuses. This finding is somewhat consistent with that of Meyer and Hammerschmid (2010), who found that uncertainty avoidance was negatively correlated with the degree of decentralization in 27 European states. Future studies can build on Meyer and Hammerschmid (2010) and on our study to investigate the nexus between country conditions and other types of management techniques in a variety of sectors.

References

Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, *53*(4), 898–944. https://doi.org/10.1257/jel.53.4.898

Ando, N. (2015). Internal mimetic behavior of MNCs with respect to foreign subsidiary staffing. *Journal of Global Mobility*, *3*(1), 61–46.

Aoki, N. (2012). Determinants of Japanese local governments' decisions concerning performance-based reward systems for teachers. *International Journal of Public Administration*, *35*(11), 703–714. https://doi.org/10.1080/01900692.2012.662778

Aoki, N. (2015). Institutionalization of New Public Management: The case of Singapore's education system. *Public Management Review*, *17*(2), 165–186. https://doi.org/10.1080/14719037.2013.792381

Aoki, N. (2019). After all these years, what has happened to the international prevalence of NPM-inspired managerial practices? *International Journal of Public Sector Management*, *32*(4), 403-417.

Aoki, N., & Tay, M. (2015). Managerialism meets regime: A distinctively singaporean Marriage? *Asian Journal of Political Science*, *23*(3), 346–365. https://doi.org/10.1080/02185377.2015.1078735

Avalos, B., & Assael, J. (2006). Moving from resistance to agreement: The case of the Chilean teacher performance evaluation. *International Journal of Educational Research*, *45*(4–5), 254–266. https://doi.org/10.1016/j.ijer.2007.02.004

Ballou, D., & Podgursky, M. (1993). Teachers' attitudes toward merit pay: Examining conventional wisdom. *Industrial and Labor Relations Review*, *47*(1), 50. https://doi.org/10.2307/2524231

Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, *63*(5), 586–606. https://doi.org/10.1111/1540-6210.00322

Bowen, D. H., Buck, S., Deck, C., Mills, J. N., & Shuls, J. V. (2015). Risky business: an analysis of teacher risk preferences. *Education Economics*, *23*(4), 470–480. https://doi.org/10.1080/09645292.2014.966062

Butterfield, R., Edwards, C., & Woodall, J. (2004). The new public management and the UK Police Service. *Public Management Review*, *6*(3), 395–415. https://doi.org/10.1080/1471903042000256556

Cao, D., Li, H., & Wang, G. (2014). Impacts of isomorphic pressures on BIM adoption in construction projects. *Journal of Construction Engineering and Management*, *140*(12), 04014056. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000903

Chang, E. C. C., & Chu, Y. (2006). Corruption and trust: Exceptionalism in Asian democracies? *The Journal of Politics*, *68*(2), 259–271. https://doi.org/10.1111/j.1468-2508.2006.00404.xChristensen, T., & Lægreid, P. (2011a). Complexity and Hybrid Public Administration—Theoretical and Empirical Challenges. *Public Organization Review*, *11*(4), 407–423. https://doi.org/10.1007/s11115-010-0141-4

Chatterjee, S., & Hadi, A. S. (2006). *Regression analysis by example*. John Wiley & Sons, Inc. https://doi.org/10.1002/0470055464Christensen, T., & Laegreid, P. (2011b). Democracy and administrative policy: contrasting elements of New Public Management (NPM) and post-NPM. *European Political Science Review*, *3*(01), 125–146Cooper, S. T., & Cohn, E. (1997). Estimation of a frontier production function for the South Carolina education process. *Economics of Education Review*, *16*(3), 313-327.

Dee, T. S., & Keys, B. J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management*, *23*(3), 471-488.

Dilulio, J. J., Garvey, G., & Kettl, D. F. (2001). *Improving government performance: An owner's manual*. Brookings Institution Press.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147–160. https://doi.org/10.2307/2095101

Dolowitz, D. P., & Marsh, D. (2000). Learning from abroad: The role of policy transfer in contemporary policy-making. *Governance*, *13*(1), 5–23. https://doi.org/10.1111/0952-1895.00121

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2005). New Public Management is dead--Long live digital-era governance. *Journal of Public Administration Research and Theory*, *16*(3), 467–494. https://doi.org/10.1093/jopart/mui057

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *The Academy of Management Review*, *14*(1), 57–74. https://doi.org/10.2307/258191

Elcock, H. (1995). The fallacies of management. *Public Policy and Administration*, *10*(1), 34–48. https://doi.org/10.1177/095207679501000104

Forrester, G. (2011). Performance management in education: milestone or millstone? *Management in Education*, *25*(1), 5–9. https://doi.org/10.1177/0892020610383902

Fraser Institute. (2017). Economic freedom of the world. Retrieved April 3, 2017, from Economic Freedom: Datatset website: http://bit.ly/2h5xBVI?__unam=b07cbf6-15ac93046cb-c09d3ec-4&has_js=1&_ga=GA1.2.231183659.1489434069

Freedom House. (2017). About freedom in the world | Freedom House. Retrieved April 3, 2017, from Freedom in the World Data website: https://freedomhouse.org/report-types/freedom-world

Frumkin, P., & Galaskiewicz, J. (2004). Institutional isomorphism and public sector organizations. *Journal of Public Administration Research and Theory: J-PART*, *14*(3), 283–307. https://doi.org/10.1093/jopart/muh028

Goh, P.S.C. (2012). The Malaysian teacher standards: a look at the challenges and implications for teacher educators. *Educational Research for Policy and Practice*, *11*(2), 73–87.

Goff, P., Goldring, E., & Canney, M. (2016). The best laid plans: Pay for performance incentive programs for school leaders. *Journal of Education Finance*, *42*(2), 127–152.

Gomez-Mejia, L. R., & Welbourne, T. (1991). Compensation strategies in a global context. *Human Resource Planning*, *14*(1), 29-41.

Green-pedersen, C. (2002). New public management reforms of the Danish and Swedish welfare states: The role of different social democratic responses. *Governance*, *15*(2), 271–294. https://doi.org/10.1111/1468-0491.00188

Gwartney, J., Lawson, R., & Hall, J. (2014). *Economic freedom of the world: 2014 Annual report*. Retrieved from https://www.fraserinstitute.org/sites/default/files/economic-freedom-of-the-world-2014-rev.pdf

Hartley, D. (2003). Education as a global positioning device: Some theoretical considerations. *Comparative Education*, *39*(4), 439–450. https://doi.org/10.1080/0305006032000162011

Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. *International Studies of Management & Organization*, *13*(1–2), 46–74. https://doi.org/10.1080/00208825.1983.11656358

Hofstede, G. H., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind: Intercultural cooperation and its importance for survival* (3rd ed). New York: McGraw-Hill.

Holbein, J. B., & Ladd, H. F. (2017). Accountability pressure: Regression discontinuity estimates of how No Child Left Behind influenced student behavior. *Economics of Education Review*, *58*, 55–67. https://doi.org/10.1016/j.econedurev.2017.03.005

Hood, C. (1991). A public management for all seasons? *Public Administration*, *69*(1), 3–19. https://doi.org/10.1111/j.1467-9299.1991.tb00779.x

Hood, C. (1995). Contemporary public management: A new global paradigm? *Public Policy and Administration*, *10*(2), 104–117. https://doi.org/10.1177/095207679501000208

Kane, J., & Patapan, H. (2006). In search of prudence: The hidden problem of managerial reform. *Public Administration Review*, *66*(5), 711–724. https://doi.org/10.1111/j.1540-6210.2006.00636.x

Kraay, A., Kaufmann, D., & Mastruzzi, M. (2010). *The worldwide governance indicators: Methodology and analytical issues*. https://doi.org/10.1596/1813-9450-5430

Lavy, V. (2007). Using performance-based pay to improve the quality of teachers. *The Future of Children*, *17*(1), 87–109.

Lee, E. W. Y., & Haque, M. S. (2006). The new public management reform and governance in Asian NICs: A comparison of Hong Kong and Singapore. *Governance*, *19*(4), 605–626. https://doi.org/10.1111/j.1468-0491.2006.00330.

Liew, W. M. (2012) Perform or else: The performative enhancement of teacher professionalism. *Asia Pacific Journal of Education*, *32*(3), 285-303.

Maranto, R., & McShane, M. Q. (2012). *President Obama and education reform: The personal and the political*. New York: Palgrave Macmillan.

Mather, K., & Seifert, R. (2011). Teacher, lecturer or labourer? Performance management issues in education. *Management in Education*, *25*(1), 26–31. https://doi.org/10.1177/0892020610388060.

Mehta, J. (2013). The allure of order: High hopes, dashed expectations, and the troubled quest to remake American schooling. Oxford: Oxford University Press.

Mettler, S. (2010). Reconstituting the submerged state: The challenges of social policy reform in the Obama era. *Perspectives on Politics*, *8*(3), 803–824.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340–363.

Meyer, R. E., & Hammerschmid, G. (2010). The degree of decentralization and individual decision making in central government human resource management: A European comparative perspective. *Public Administration*, *88*(2), 455–478. https://doi.org/10.1111/j.1467-9299.2009.01798.x

Moe, T. M. (2011). *Special interest: Teachers unions and America's public schools*. Washington, D.C.: Brookings Institution Press.

Moe, T. M. (2012). Teachers unions and American education reform: The politics of blocking. *The Forum*, *10*(1), Article 4.

Moe, T. M. (2019). *The politics of institutional reform: Katrina, education, and the second face of power*. Cambridge: Cambridge University Press.

Moe, T. M., & Wiborg, S. (2017). Introduction. In T. Moe & S. Wiborg (Eds.), *The comparative politics of education: Teachers unions and education systems around the world* (pp. 1–23). https://doi.org/10.1017/9781316717653.001

Montes, F. J. L., & Jover, A. J. V. (2004). Total quality management, institutional isomorphism and performance: the case of financial services. *The Service Industries Journal*, *24*(5), 103–119. https://doi.org/10.1080/0264206042000276865

Moon, M. J., & deLeon, P. (2001). Municipal reinvention: Managerial values and diffusion among municipalities. *Journal of Public Administration Research and Theory*, *11*(3), 327–352. https://doi.org/10.1093/oxfordjournals.jpart.a003505

Morris, S. D., & Klesner, J. L. (2010). Corruption and trust: Theoretical considerations and evidence from Mexico. *Comparative Political Studies*, *43*(10), 1258–1285. https://doi.org/10.1177/0010414010369072

Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy*, *119*(1), 39–77. https://doi.org/10.1086/659655

Newman, K. L., & Nollen, S. D. (1996). Culture and congruence: The fit between management practices and national culture. *Journal of International Business Studies*, 27(4), 753-779.

Niskanen, W. A. (1971). *Bureaucracy and representative government* (1st Pbk. Print edition). Chicago: Aldine Transaction.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673-690.

OECD. (2005). *Performance-related pay policies for government employees*. OECD Publishing.

Park, S., & Berry, F. (2014). Successful Diffusion of a Failed Policy: The case of pay-for-performance in the US federal government. *Public Management Review*, *16*(6), 763–781. https://doi.org/10.1080/14719037.2012.750835

Peters, B. G., & Pierre, J. (1998). Governance without Government? Rethinking public administration. *Journal of Public Administration Research and Theory: J-PART*, *8*(2), 223–243.

Pitts, D. W., Hicklin, A. K., Hawes, D. P., & Melton, E. (2010). What drives the implementation of diversity management programs? Evidence from public organizations. *Journal of Public Administration Research and Theory*, *20*(4), 867–886. https://doi.org/10.1093/jopart/mup044

Ploom, K., & Haldma, T. (2013). Balanced performance management in the public education system: An empirical study of Estonian general education schools. *Baltic Journal of Management*, *8*(2), 183–207. https://doi.org/10.1108/17465261311310018

Podgursky, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, *26*(4), 909–950.

Pollitt, C. (2001). Convergence: The useful myth? *Public Administration*, *79*(4), 933–947. https://doi.org/10.1111/1467-9299.00287

Pollitt, C. (2016). Managerialism redux? *Financial Accountability & Management*, *32*(4), 429–447. https://doi.org/10.1111/faam.12094

Pollitt, C., & Bouckaert, G. (2011). *Public management reform: A comparative analysis - New public management, governance, and the neo-Weberian state* (3 edition). Oxford: Oxford University Press.

Pollitt, C., & Dan, S. (2013). Searching for impacts in performance-oriented management reform. *Public Performance & Management Review*, *37*(1), 7–32. https://doi.org/10.2753/PMR1530-9576370101

Pretorius, S. G., & Ngwenya, V. C. (2008). Teachers' perceptions of and attitudes towards performance appraisal in Zimbabwean schools. *Africa Education Review*, *5*(1), 144–164. https://doi.org/10.1080/18146620802144859

Ravitch, D. (2014). *Reign of error: The hoax of the privatization movement and the danger to America's public schools* (Reprint edition). New York: Vintage.

Reeves, J., Forde, C., O'Brien, J., Smith, P., & Tomlinson, H. (2002). Chapter 2: The Concern with Performance and Continuing Professional Development. In *Performance Management in Education: Improving Practice*. https://doi.org/10.4135/9781446220931.

Riccucci, N. M., & Thompson, F. J. (2008). The new public management, Homeland Security, and the politics of civil service reform. *Public Administration Review*, *68*(5), 877–890.

Ritter, G. W., & Jensen, N. C. (2010). The delicate task of developing an attractive merit pay plan for teachers. *Phi Delta Kappan*, *91*(8), 32–37.Ritter, G. W., Maranto, R., & Buck, S. (2009). Harnessing private incentives in public education. *Review of Public Personnel Administration*, *29*(3), 249–269. https://doi.org/10.1177/0734371X09338890

Rivenbark, W. C., Fasiello, R., & Adamo, S. (2019). Exploring performance management in Italian local government: The necessity of outcome measures and citizen participation. *The American Review of Public Administration*, *49*(5), 545–553. https://doi.org/10.1177/0275074018775125

Romanowski, M.H. (2014). The Qatar national professional standards for school leaders: a critical discourse analysis using Habermas' theory of knowledge constitutive interests. *International Journal of Leadership in Education*, *17*(2), 174–199.

Smarick, A. 2011. Diplomatic mission: President Obama's path to performance pay. *Education Next*, *11*(1), 56–63.

Smith, R. (1995). Staff appraisal in higher education — A study of performance review at Nene College, Northampton. *Higher Education*, *30*(2), 189–205. https://doi.org/10.1007/BF01384096Sozen, S., & Shaw, I. (2002). The international applicability of "new" public management: lessons from Turkey. *International Journal of Public Sector Management*, *15*(6), 475–486. https://doi.org/10.1108/09513550210439625

Steger, M. B., & Roy, R. K. (2010). Neoliberalism: A very short introduction. Oxford and New York: Oxford University Press.

Storey, A. (2002). Performance management in schools: Could the balanced scorecard help? *School Leadership & Management*, *22*(3), 321–338. https://doi.org/10.1080/1363243022000020435

Studenmund, A. H. (2014). *Using econometrics: A practical guide* (Sixth edition, Pearson new international edition). Pearson Education.

Sulle, A. (2010). The application of new public management doctrines in the developing world: An exploratory study of the autonomy and control of executive agencies in

Tanzania. *Public Administration and Development*, *30*(5), 345–354. https://doi.org/10.1002/pad.580

Taylor, F. W. (1911). *The principles of scientific management*. New York, London: Harper & Brothers.

Teodoro, M. P. (2011). *Bureaucratic ambition: Careers, motives, and the innovative administrator* (1 edition). Johns Hopkins University Press.

Thatcher, M. (1993). *Downing street years*. New York: HarperCollins Publishers.

The Fund for Peace. (2017). Fragile states index data | The Fund for Peace. Retrieved April 3, 2017, from http://fsi.fundforpeace.org/indicators

The Heritage Foundation. (2017). Index of economic freedom. Retrieved April 3, 2017, from http://www.heritage.org/index/explore

Tolofari, S. (2005). New public management and education. *Policy Futures in Education*, *3*(1), 75–89. https://doi.org/10.2304/pfie.2005.3.1.11

Villadsen, A. R., Hansen, J. R., & Mols, N. P. (2010). When do public managers imitate each other? Mimetic decision making in contracting decisions of Danish municipalities. *Public Organization Review*, *10*(4), 357–376. https://doi.org/10.1007/s11115-010-0111-x

Vroom, V. H. (1964). *Work and motivation* (1 edition). New York: Wiley.

Zvavahera, P. (2014). An evaluation of the effectiveness of performance management systems on service delivery in the Zimbabwean civil service. *Journal of Management and Marketing Research*, *14*, 1–8.

Table 1. Hypothesized National Contexts and Directions of Association

H1.	market economy	(a) +; (b) -
H2.	liberal democracy	(a) +; (b) -
H3.	income levels	+
H4.	isomorphic pressure	+
	national culture:	
H5.	• power distance	+
H6.	• individualism	+
H7.	• masculinity	+
H8.	• uncertainty avoidance	(a) +; (b) -
H9.	decentralized structure	(a) +; (b) -
H10.	teacher group influence	-

Note: '+' denotes a positive relation between the use of performance-based pay and a specific aspect of national context. '-' denotes a negative association between the use of performance-based pay and a specific aspect of national context.

Table 2. Descriptive Statistics (*N* = 51)

		Mean	SD	Min	Max
Dependent Variables					
% of schools whose appraisals were linked with:					
SALARY1	a small change or more in salary	33.51	29.80	0.00	99.76
SALARY2	a moderate or large change in salary	13.43	18.76	0.00	77.48
SALARY3	a large change in salary	4.96	9.53	0.00	34.37
BONUS1	A small change or more in bonus	37.97	36.86	0.00	99.81
BONUS2	a moderate or large change in bonus	18.98	24.08	0.00	91.14
BONUS3	a large change in bonus	7.44	15.80	0.00	89.53
Independent Variables					
FREEDM	EFW index	68.86	7.41	50.50	87.50
VOICE	voice and accountability	72.67	21.84	18.31	100.00
GNPLN	GNI (log)	9.96	0.86	8.18	11.50
EXTER	external intervention	3.70	1.85	0.80	7.60
OECD	OECD membership	0.65	0.48	0.00	1.00
POWER	culture: power distance	56.08	22.07	13.00	100.00
INDIVI	culture: individualism	48.73	23.65	13.00	91.00
MASCU	culture: masculinity	45.61	21.55	5.00	100.00
AVOID	culture: uncertainty avoidance	69.12	22.19	8.00	100.00
TEACHER	teacher group influence	39.91	13.79	1.94	63.87
DECENT	structure	62.42	18.56	20.46	98.53
NGOV	average % of funding from non-government sources	10.96	15.63	0.00	69.00
CITY	% of schools located in city with population ≥ 100,000	32.23	23.72	0.00	100.00
SIZE	average school size	699.93	382.71	183.86	1721.41
RATIO	average student-teacher ratio	14.09	5.74	5.90	32.59

SD stands for standard deviation. Min and Max denote the minimum and maximum, respectively

Table 3. Correlation Coefficients between Independent Variables

VARIABLE		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	VOICE	1.00														
2	FREEDM	0.53	1.00													
3	GNPLN	0.70	0.58	1.00												
4	EXTER	-0.67	-0.55	-0.75	1.00											
5	OECD	0.70	0.34	0.66	-0.50	1.00										
6	TEACHER	0.09	0.00	0.05	-0.21	0.08	1.00									
7	DECENT	0.34	0.09	0.10	-0.20	0.30	0.44	1.00								
8	POWER	-0.74	-0.54	-0.62	0.47	-0.57	-0.13	-0.29	1.00							
9	INDIVI	0.69	0.44	0.69	-0.70	0.67	0.28	0.44	-0.67	1.00						
10	MASCU	-0.09	0.05	0.01	0.02	0.14	-0.22	-0.01	0.21	0.12	1.00					
11	AVOID	-0.30	-0.59	-0.43	0.43	-0.16	-0.14	-0.24	0.39	-0.43	0.09	1.00				
12	NGOV	-0.27	-0.12	-0.30	0.35	-0.06	-0.13	-0.16	0.05	-0.34	0.08	0.27	1.00			
13	CITY	-0.10	0.28	0.11	0.00	-0.04	0.13	-0.03	-0.05	-0.06	0.18	-0.08	0.24	1.00		
14	SIZE	-0.04	0.29	0.00	0.08	-0.10	-0.12	-0.10	-0.09	-0.14	0.15	-0.12	0.28	0.24	1.00	
15	RATIO	-0.27	0.04	-0.35	0.27	-0.38	0.02	-0.01	0.09	-0.27	0.08	-0.06	0.28	0.25	0.44	1.00

Table 4. Estimation Results for Salary (*N* = 51)

	SALARY1		SALARY2		SALARY3	
FREEDM	-0.51	(0.67)	0.21	(0.51)	0.30	(0.27)
VOICE	-0.90	(0.29)***	-0.81	(0.22)***	-0.21	(0.11)*
GNPLN	3.90	(7.46)	2.52	(5.62)	-3.76	(2.96)
EXTER	-5.17	(3.05)*	-2.37	(2.30)	-0.96	(1.21)
OECD	1.54	(11.11)	8.80	(8.38)	2.69	(4.42)
POWER	-0.03	(0.28)	-0.09	(0.21)	0.12	(0.11)
INDIVI	-0.61	(0.30)**	-0.12	(0.22)	0.04	(0.12)
MASCU	0.16	(0.19)	-0.01	(0.14)	-0.01	(0.07)
AVOID	-0.56	(0.19)***	-0.09	(0.14)	-0.04	(0.08)
TEACHER	0.29	(0.27)	0.21	(0.20)	-0.08	(0.11)
DECENT	0.39	(0.20)*	-0.08	(0.15)	0.01	(0.08)
NGOV	0.63	(0.25)**	0.12	(0.19)	0.12	(0.10)
CITY	-0.22	(0.15)	-0.34	(0.12)***	-0.15	(0.06)**
SIZE	-0.02	(0.01)**	-0.01	(0.01)	-0.0002	(0.004)
RATIO	0.31	(0.67)	0.58	(0.50)	-0.12	(0.27)
Constant	151.72	(107.65)	55.96	(81.21)	41.13	(42.78)
R-squared	0.665		0.519		0.483	

SALARY1 denotes the percentage of schools where appraisals have been linked with a small change or more in salary. SALARY2 denotes the percentage of schools where appraisals have been linked with a moderate or large change in salary. SALARY3 denotes the percentage of schools where appraisals have been linked to at least a large change in salary.

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels or better, respectively. Standard errors are in parentheses.

Table 5. Estimation Results for Bonus (*N* = 51)

	BONUS1		BONUS2		BONUS3	
FREEDM	-0.22	(0.93)	0.43	(0.72)	0.78	(0.46)
VOICE	-0.53	(0.40)	-0.68	(0.31)**	-0.27	(0.20)*
GNPLN	0.67	(10.29)	-2.04	(7.94)	-3.62	(5.14)
EXTER	0.50	(4.21)	-2.33	(3.25)	-1.17	(2.10)
OECD	11.37	(15.34)	7.57	(11.84)	0.52	(7.66)
POWER	0.41	(0.39)	0.21	(0.30)	0.23	(0.19)
INDIVI	-0.55	(0.41)	-0.15	(0.32)	-0.08	(0.20)
MASCU	0.43	(0.26)	0.17	(0.20)	0.0010	(0.13)
AVOID	-0.44	(0.26)*	-0.17	(0.20)	-0.19	(0.13)
TEACHER	0.86	(0.37)**	-0.03	(0.29)	-0.20	(0.19)
DECENT	0.73	(0.28)**	0.34	(0.21)	0.09	(0.14)
NGOV	0.31	(0.34)	-0.09	(0.26)	-0.01	(0.17)
CITY	-0.34	(0.21)	-0.05	(0.16)	0.04	(0.11)
SIZE	0.0046	(0.01)	-0.01	(0.01)	0.00	(0.01)
RATIO	-0.58	(0.92)	0.12	(0.71)	-0.61	(0.46)
Constant	26.69	(148.59)	49.57	(114.70)	25.24	(74.20)
R-squared	0.583		0.418		0.434	

BONUS1 denotes the percentage of schools where appraisals have been linked with at least a small change or more in bonus. BONUS2 denotes the percentage of schools where appraisals have been linked with a moderate or large change in bonus. BONUS3 denotes the percentage of schools where appraisals have been linked to at least a large change in bonus.

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels or better, respectively. Standard errors are in parentheses.

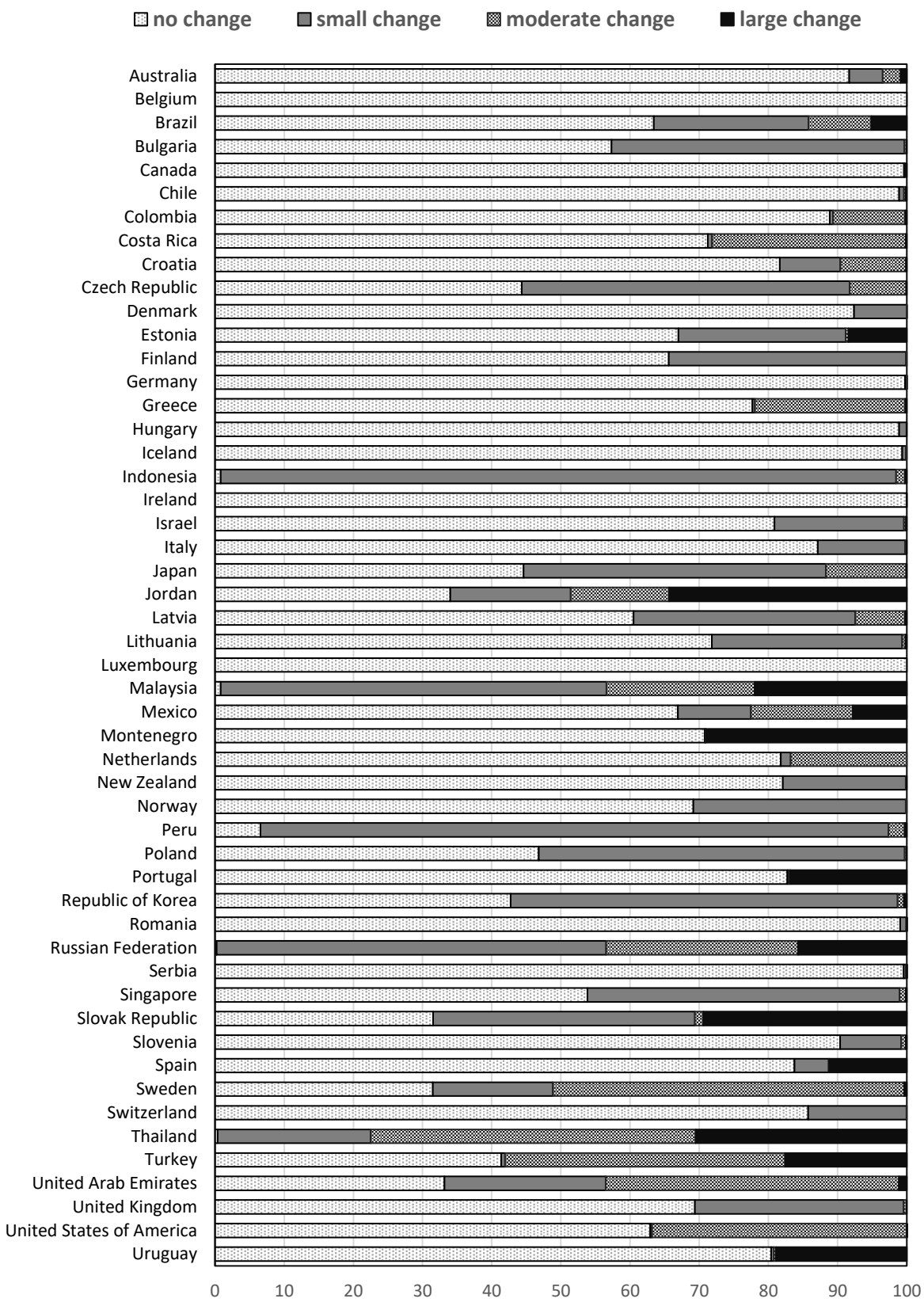

Figure 1. Weighted percentage distribution of schools whose teacher appraisals led directly to varying degrees of changes in salary, 2012.

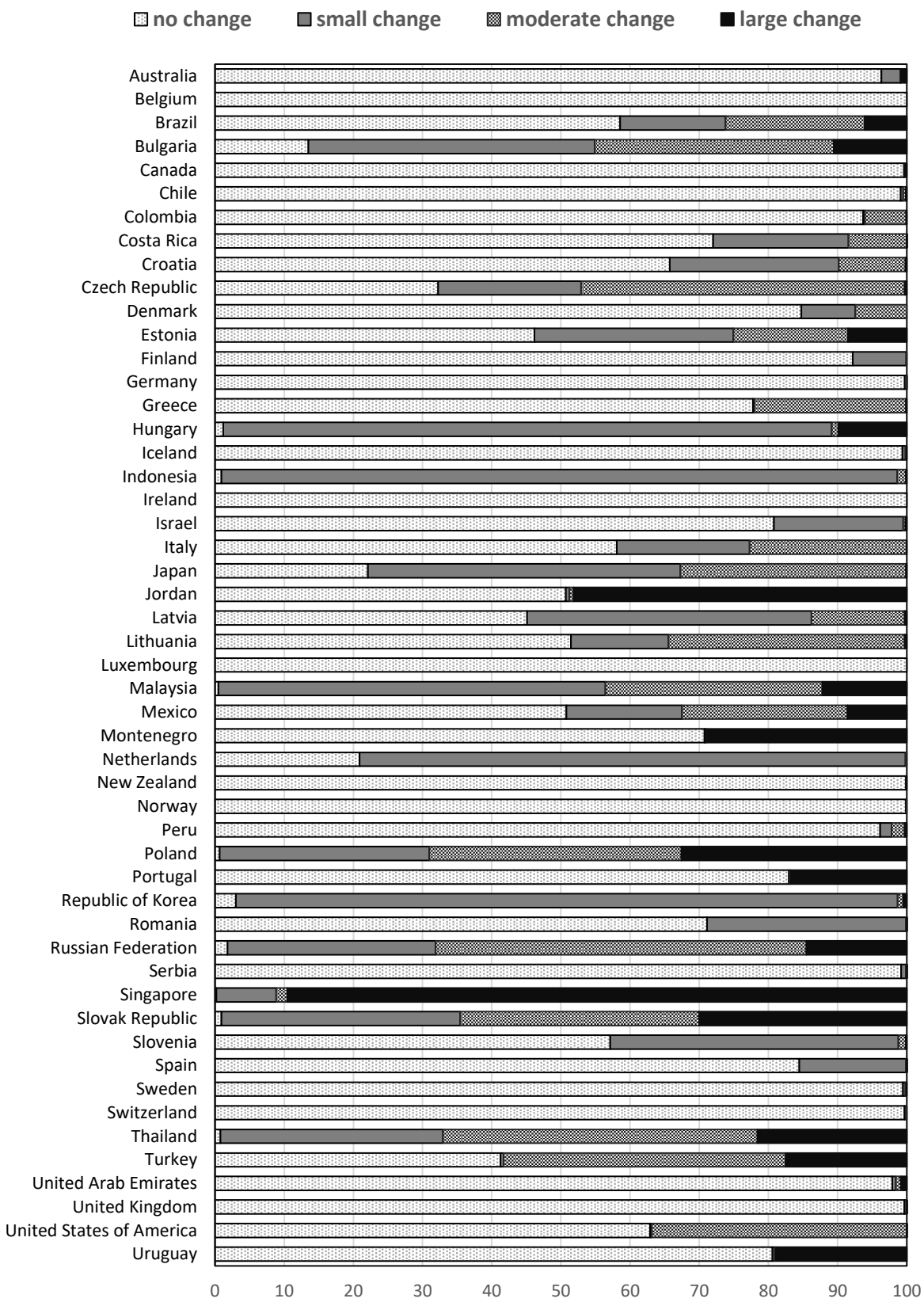

Figure 2. Weighted percentage distribution of schools whose teacher appraisals led directly to varying degrees of changes in bonuses, 2012.