3rd International Conference on Public Policy (ICPP3)
June 28-30, 2017 – Singapore

Panel T05P01, Session 1- Performance Management
Public Management Reforms across the Globe: Results, Challenges and Issues

Title of the paper
Performance Appraisals: Their Patterns of International Prevalence

Author(s)
Naomi Aoki, Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, sppnma@nus.edu.sg

Stuti Rawat, Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, stuti.r@u.nus.edu

Date of presentation
June 29, 2017

Performance Appraisals: Their Patterns of International Prevalence

Naomi Aoki[a,b] and Stuti Rawat[a]

[a]Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore

[b]Corresponding author: sppnma@nus.edu.sg

ABSTRACT

Performance-based pay is arguably a popular technique, but its international prevalence remains poorly evidenced. 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, and examine whether these variations arise from country contexts. We found that performance-based pay tended to be used more among OECD countries, in less liberal economies, in cultures with a lower degree of uncertainty avoidance and a higher degree of masculinity, in those with more decentralized educational systems, and surprisingly, in places where teachers exert higher influence.

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

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, 20). Public organizations use it to control and motivate their employees, among other purposes (Behn 2003), and 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 be linked with rewards. So-called *performance-based pay* in particular links appraisals with monetary rewards. According to the Organization for Economic Co-operation and Development (OECD) (2005) performance-based pay was used, or at least considered, for public servants in two-third 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 local 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 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 do to run their public organizations (Hood 1991, 1995; McGuire 2001). In these debates, performance-based pay was often, albeit not always, regarded as a NPM-inspired technique (Butterfield, Edwards, and Woodal 2004; Park and Berry 2014; Aoki 2015). 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). This debate has arisen due to limited empirical evidence on NPM's prevalence.

Our research question regarding the prevalence of NPM by country conditions may be of importance to policy makers who are concerned with whether a certain technique abroad can be transferrable to their own country, given their national contexts. Dolowitz and Marsh (2002) held that policy makers increasingly rely on policy transfer from abroad, but 'inappropriate transfer' can occur when a policy is transferred without accounting for the 'economic, social, political and ideological contexts' of borrowing countries (p. 17). Pollitt (2000) argued that behind the apparent global spread of NPM was a normative assumption that some countries were well ahead, while others were lagging behind, and that the latter should learn from the former. However, others have argued that NPM-inspired practices in the West are not exportable to certain countries due to dissimilar political contexts (Sozen and Shaw 2002; Sulle 2010). Empirical evidence of the prevalence of NPM by context can clarify whether these concerns are justified.

In the past, studies have investigated the factors influencing the prevalence pattern of administrative reforms in government. Moon and deLeon (2001), for example, used the 'extended reinvention adoption index' (ERAI), which measures the degree to which 12 managerial innovations in line with the reinvention[i] themes were adopted by municipal governments in the United States (US) as of 1997. They found that the adoption of these innovations was determined by 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 cultures – 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 tends to be higher in the Anglo-Saxon administrative tradition.

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 and on education is reasonable, considering a finding by the OECD (2005) that the prevalence of performance-based pay depends on areas of public service, and because performance appraisal has been at the center of policy discourse in education. Our choice of the education sector was also prompted by the availability of the data: 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 education for15-year-olds in 65 countries. Like Moon and deLeon (2001), and Meyer and Hammerschmid (2010), this study pertains to a single year, and hence, does not yield causal inferences as to what drives or hampers the adoption of performance-based pay. Nevertheless, associational analyses are sufficient for addressing our questions regarding prevalence.

The rest of this article is structured as follows. The next section proceeds to present 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 hypothesize that performance-based pay is used more in countries which 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, 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 1996). 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).

However, we offer a caveat about generalizing from the above observations. This is because, first, NPM was endorsed as well by governments devoid of liberal ideology, such as labor administrations in Australia (Pollitt 2016) and Scandinavian welfare states (Green-Pedersen 2002). 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 it be the latter, it is unclear what may be considered NPM practices (Hood 1995; Pollitt 2016), and NPM-inspired reforms do not necessarily include performance-based pay. Fourth, performance-based pay might have been in use even before NPM emerged. 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 these considerations it is worthwhile to look at what the empirical evidence has to say regarding the hypothesis regarding performance-based pay

Political Regime

We consider whether the type of political regime makes a difference and hypothesize that performance-based pay is used more in liberal democracies than in less liberal regimes (**H2**). This hypothesis was inspired by the fact that in the views of a large number of scholars, the NPM-inspired managerial movement began in liberal democracies (e.g. Peters and Pierre 1998; Hartley 2003; Cheung 2005; Lee and Haque 2006; Sarker 2006). 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, although the confusion over this point, discussed above, also applies here. The second reason is that 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 the authorities (Aoki and Tay 2015).

Income Level

We postulate 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 such a global campaign to promote it 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 because low-income countries are inadequately prepared to implement performance appraisals (Smith 1995), which would require 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 program that takes up 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 and Ngwenya 2008; Zvavahera 2014).

Isomorphic Pressures

What Meyer and Rowan (1977) termed *isomorphic pressure* has been proven to drive business organizations to adopt certain policies (Montes and Jover 2004; Barreto and Baden-Fuller 2006; Nair and Prajogo 2009; Ando 2011, 2015; Cao et al. 2014; He et al. 2016), and Pollitt (2001) held that this applies to government organizations and argued that isomorphic pressures were behind the global spread of NPM-inspired practices. 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 a *normative* isomorphic pressure that promoted the spread of reforms. There was *mimetic* isomorphic pressure, too, accompanied by a powerful story that 'there was something new in the world of governance' and that some governments were ahead, while others were lagging behind (Pollitt 2000, 183), driving the lagging countries to mimic the leading countries for the sake of legitimacy. In line with these arguments, we examine whether the adoption status of performance appraisals depends on isomorphic pressures (**H4**).

Our hypothesis is also in accordance with empirical findings. Frumkin and Galaskiewicz (2004), for example, 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 organizations' 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 programs in the schools.

National Culture

Building on Meyer and Hammerschmid (2010), we also posit that national culture matters to the adoption status of performance appraisals (**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 conductive to administrative reforms than others. Following Meyer and Hammerschmid (2010), we adopt Hofstede's definition of national culture **(**n.d.): 'the collective programming of the mind which distinguishes the members of one group or category of people from another.' Hofstede (n.d) 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, this culture would be more conducive to the use of performance appraisals in general, 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 power-distant culture, which the technique itself created. Furthermore, performance-based pay might fit more with national cultures with a higher degree of individualism; hence, we posit that it is used more in such cultures (**H6)** because they drive school staff (i.e. teachers and school principals) to monitor their individual accomplishments, and appraisals help them to know how much they have accomplished. Likewise, a 'masculine' society tends to be competitive, and hence, school staff are more willing to accept performance appraisals, which arguably induces competition among individuals for limited rewards (**H7**). Finally, we hypothesize that workers in some cultures like to work towards established performance targets, and hence, performance appraisals might be used more in national cultures with a higher degree of uncertainty avoidance (**H8**).

Administrative Structure

Christensen and Lægreid (2007, 2011a, 2011b) argued that administrative reform outcomes can be explained by the dynamic interplay between (ii) external environmental factors and (i) the cultural norms and internal structures of public organizations. We consider structure in terms of the degree to which organizations are decentralized, and consider 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 posit that performance-appraisal pay is used more in centralized system (**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 reform. Nevertheless, 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 and Patapan 2006, 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' (715). In education, too, decentralization and accountability through performance measures are often promoted in combination (Ladd 1996).

Teacher Group Influence

Our last hypothesis is 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 from Scotland (Reaves et al. 2002) to Japan (Aoki 2012) and Chile (Avalos 2004; Avalos and Assael 2006; Taut et al. 2010). In general, teachers oppose the use of performance measures in education because quantitative performance measures poorly reflect the complexity and breadth of teachers' work (Storey 2002; Ozga 2003; Levacic 2008; Ploom and Haldma 2013), because the use of rankings based on performance measures may hamper the design of the most appropriate contexts, curriculums, and approaches for students (Ozga 2003), because conducting appraisals can be an additional burden (Pretorius and Ngwenya 2008), and because appraisals can result in a 'tick-box mentality', damage employees' trust in their supervisors, and distort the focuses and priorities of education (Forrester 2011, 8). Teachers resist formally through unions, as well as informally by means of absenteeism, increased instrumentalism and dull compliance in their jobs (Ko 2001; Mather and 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 pressures as a control in their study, found them to be negatively associated with the degree of government decentralization.

<Insert Table 1 here.>

EMPIRICAL STRATEGY

Our dependent variables were derived from the school questionnaires in the 2012 PISA, administered by the OECD in 65 countries and subnational economies (for the sake of simplicity, both are hereafter referred to as 'countries').The PISA used both student and school questionnaires, and our dependent variables were derived from school principals' ordered responses to the school questionnaires which asked to what extent appraisals of and/or feedback to teachers (hereafter, simply 'appraisals') had directly led to changes in, respectively, (i) salary, and (ii) a financial bonus or other kind of monetary reward (hereafter simply referred to as a 'bonus').School principals were asked to select from among four items: no change (=0), a small change (=1), a moderate change (=2), or a large change (=3).In Table 2, these responses were labelled SALARY and BONUS. We performed ordered logistic regression analyses to test whether the hypothesized variables would predict a change in the log-odds of the school principals' responses being in a higher order (larger change) rather than in a lower order (hereafter, simply 'the utilization level of a performance-based pay'), according to: $\text{logit}[P(Y_i \le j)] = \alpha_j$ - βx_k, where $P(Y_i \le j)$ is the cumulative probability of a principal i's response category being in j or lower.

Some cautions are required to understand the units of analysis in this study. The PISA sampling method was designed to select representative 15-year-olds in each country, but not representative schools. Accordingly, the *PISA Data Analysis Manual* (OECD 2012) suggests that researchers merge school and student-level data and analyse the school information in terms of students' attributes – i.e., information about a 15-year-old student's school – and use weights to correct for differences in the sampling probabilities of both students and schools, and to equalize each country's contribution to the original international dataset. We followed these guidelines. As a result, the units of observation in our dataset are students, not schools.

In our logistic regressions, we used data for all types of schools, instead of just public schools. This is because organizations in general raise revenues from mixed sources, which blurs the line between public and private ownership (Moore 2000). In education, public schools raise some private fees, while private schools receive government funding. Our approach, therefore, is to control for both (i) the ownership category (i.e. public or private) and (ii) sources of school revenue, so that we can interpret the coefficients for key independent variables when ownership is public and when a school is fully funded by the government. This is done by coding two variables, as follows: PRIVATE in Table 2 was coded as one if a school was a private school – that is, 'managed directly or indirectly by a non-government organization; e.g. a church, trade union, business, or other private institution' – and coded as zero if it was 'managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise' (OECD 2011, 4). The other variable, NONGOV, represents the percentage of school funding from non-government sources, such that zero for this variable means that the school was 100% government-funded.

We make the most of existing international datasets for country-context variables. For economic ideology, we use the 2012 Economic Freedom of the World Index, published by the Fraser Institute (2017), labelled EFWI in Table 2. This index is based on the notion that in an economically free society, 'the primary role of government is to protect individuals and their property from aggression by others', and the index is 'designed to measure the extent to which the institutions and policies of a nation are consistent with this protective function' (Gwartney et al. 2014, 1). Based on this definition, the EFWI, using a scale from zero to ten, captures the degree of economic freedom according to: (i) the size of 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.[ii]

For a measure of the type of political regime, we use 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 et al. 2010).[iii] This indicator is used as a proxy for liberal democracy. To measure the income levels of countries, we utilize the log term of the Gross National Income (GNI) per capita in USD, labelled GNPLN in Table 2.

We use two variables as proxies for isomorphic pressure, one of which is OECD membership. Because the OECD extensively engages in knowledge sharing and dissemination in the area of governance and public administration, we assume that OECD members are under the normative isomorphic pressure to assimilate by adopting 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 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.[iv]We follow Meyer and Hammerschmid (2010) and use Hofstede's (n.d.) cultural dimensions of (i) power distance (POWER), (ii) individualism (INDIVI), (iii) masculinity (MUSCL), and (iv)uncertainty avoidance (AVOID).[v]

Asa measure of administrative structure, we use school principals' responses to the PISA regarding who bore considerable responsibility for the following twelve 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 contents, and (xii) deciding which courses are offered. School principals chose 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. We computed the percentage of students (units of observation), whose school principal, teachers, or governing board, had considerable responsibility for the 12 tasks. When a school principal's answer in regard to all of the stakeholders was zero, we counted the response as missing, because it indicated an unrealistic situation where no one was responsible for a task.

Finally, the influence of teachers, labelled TEACHR in Table 2, was also derived from school principals' responses to the PISA: '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 this question was not asked in the 2012 PISA, and other international data on teacher union strength did not have sufficient country coverage. For each country, we created a variable called TEACHR. To do this we computed the share of students whose school principals said that teacher groups, such as staff associations, curriculum committees and trade unions, exerted a direct influence on decisions within the four areas mentioned in the question, generating four measures. TEACHR is the average of these four measures.

RESULTS

Before presenting the results from the ordered logistic regression, it is worth investigating the extent to which teacher appraisals are linked with salaries and bonuses across countries.

Figures 1 and 2 show the percentage distributions of students whose teacher appraisals led to different degrees of change in salaries and bonuses, respectively. 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 leading-edge – namely, the US, UK, and New Zealand – and only 1.4%and 1.1%of students had teachers whose appraisals were linked with 'a large change' in salaries and bonuses in Australia, whereas for salaries this number was highest at 43.1% in Hong Kong and 30.5% in Thailand, and for bonuses at 89.5% in Singapore. In Malaysia, Russia, Thailand and Tunisia, teacher appraisals were unequivocally perceived to be linked to both bonuses and salaries. On the other hand, in Singapore, Hungary and Poland, only a specific form of performance-based pay (i.e. appraisals linked to bonuses) was prevalent. In the middle of the spectrum lie countries like Peru and the Czech Republic, where appraisals were perceived to result in small to moderate changes in both salaries and bonuses.

As for the estimates from the logistic regressions, Table 3 shows the coefficients and standard errors clustered at the country level for Models I and II, in which the dependent variables were school principals' ordered responses regarding performance appraisals linked with salaries and with bonuses, respectively. For the sake of simplicity, we interpret the coefficient as the impact of a change in the predictor variable on the degree of utilization of performance-based salaries and bonuses. The results also show the odds ratios for alternative interpretations. Both models controlled for school enrolment size, school location, and student-teacher ratio.

Three variables are significant in both models. The voice and accountability indicator is one, but the sign is negative; we can conclude with confidence ($p<0.01$) that the utilization level of performance-based pay tends to be higher in less liberal political regimes. This maybe a surprising result, considering that performance appraisal can be a means of enhancing accountability in government, which is a virtue in liberal democracies. Instead, this result could be due to the fact that performance appraisal has been used in less liberal democracies even more as a means of controlling public servants, not necessarily of promoting 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.

While external intervention is not significant, OECD membership is; the utilization of performance-based pay tends to be higher among OECD members than among non-members. What is noteworthy is the large odds ratio; the odds for changes in salaries and bonuses are 3.12 and 4.06 times higher for OECD members than non-OECD members, respectively. The results for administrative structure are also consistent with our hypotheses; the utilization of performance-based pay tends to be greater in more decentralized countries ($p<0.1$), and as discussed, 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.

Unlike the aforementioned variables, teacher group influenceis significant only in Model II. While the relation between teacher group influence and performance-based salary is not statistically significant, we surprisingly find a positive and statistically significant ($p<0.1$) association between teacher group influence and performance-based bonuses. That is the greater teacher-group influence becomes, the more performance-based bonuses are used. This could have something to do with the ways in which teacher-group influenceis 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, which could have gone against the use of performance-based bonuses. Culture variables are also inconsistent between the two models. We find a statistically significant and negative association between uncertainty avoidance and performance-based salaries, and a statistically significant and positive association between masculinity and performance-based bonuses. Finally, income levels are significant only for bonuses linked to appraisals, and the negative sign indicates that this practice tended to be used by low-income countries. This could be due to the fact that low-income countries face tighter budget constraints than high-income countries in financing this practice, as the case of Zimbabwe, cited earlier, suggests.

DISCUSSION AND CONCLUSION

Before proceeding to draw implications from our findings, some caveats are worth noting. This study relied on information based on school principals' opinions, and 'small', 'moderate', and 'large' changes in salaries and bonuses can have different meanings for different principals. 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, but future researchers can revisit other measures. Furthermore, as we warned at the outset, our single-year data do not support causal inferences; the results do not suggest that the variables either facilitate or hamper the adoption of performance-based pay; rather, they simply suggest that the technique can be *used* in places under the conditions that these variables represent. This is an important caveat, especially for policy makers who are wondering whether performance-based pay can be used in their own countries.

With the above caveats in mind, our findings offer some insights into the debates discussed earlier over whether or not countries are converging or diverging in their practices, and whether performance-based pay still persists in the post-NPM era – questions that have not been addressed empirically. Although a single-year snapshot does not offer an answer in light of dynamic processes of conversion and diversion 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 one of them, is used extensively in some countries in their education sector. This finding offers partial support for Pollitt (2016), who argues that NPM is not dead.

Our results also suggest that some country contexts matter to the prevalence pattern of performance-based pay, and offer some insights for policy makers who may be considering 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 logistic regressions suggest that high-income countries and liberal democracies do not champion the use of performance-based pay, rather it tends to be used more in less liberal political regimes and low-income countries, and the perception that it is being used in those seemingly leading-edge countries may be outdated. To put it differently, the type of political regime and a country's income level are not good excuses for not being able to institutionalize this technique.

Our results also offer some general implications for existing theories concerning the adoption of administrative reforms. 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. As noted, we derived our hypothesis 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 the cultural hypothesis for masculinity, wherein we find a statistically significant and positive association between masculinity and performance-based bonuses. The finding for uncertainty avoidance is counter to our hypothesis. We find a negative association between uncertainty avoidance and performance-based salaries. This finding is somewhat consistent with those 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

Sarker, A.E. 2006. "New Public Management in Developing Countries: An Analysis of Success and Failure with Particular Reference to Singapore and Bangladesh." *International Journal of Public Sector Management* 19 (2): 180–203. doi:10.1108/09513550610650437.

Ailon, G. 2008. "Mirror, Mirror on the Wall: 'Culture's Consequences' in a Value Test of Its Own Design." *The Academy of Management Review* 33 (4): 885–904. doi:10.2307/20159451.

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

Ando, N. 2011. "Isomorphism and Foreign Subsidiary Staffing Policies." Edited by Jan Selmer. *Cross Cultural Management: An International Journal* 18 (2): 131–43. doi:10.1108/13527601111125987.

Ando, N. 2015. "Internal Mimetic Behavior of MNCs with Respect to Foreign Subsidiary Staffing." *Journal of Global Mobility: The Home of Expatriate Management Research* 3 (1): 46–61. doi:10.1108/JGM-08-2014-0036.

Aoki, N. 2012. “Determinants of Japanese Local Governments’ Decisions Concerning Performance-Based Reward Systems for Teachers.” *International Journal of Public Administration* 35 (11): 703–14. doi: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): 165–86. doi:10.1080/14719037.2013.792381.

Aoki, N., and M. Tay. 2015. “Managerialism Meets Regime: A Distinctively Singaporean Marriage?” *Asian Journal of Political Science* 23 (3): 346–65. doi:10.1080/02185377.2015.1078735.

Avalos, B. 2004. “Teacher Regulatory Forces and Accountability Policies in Chile: From Public Servants to Accountable Professionals.” *Research Papers in Education* 19 (1): 67–85. doi:10.1080/0267152032000179981.

Avalos, B., and J. Assael. 2006. “Moving from Resistance to Agreement: The Case of the Chilean Teacher Performance Evaluation.” *International Journal of Educational Research*, Education Reform and the Global Regulation of Teachers’ Education, Development and Work: A Cross-Cultural Analysis, 45 (4–5): 254–66. doi:10.1016/j.ijer.2007.02.004.

Barreto, I, and C. Baden-Fuller. 2006. “To Conform or To Perform? Mimetic Behaviour, Legitimacy-Based Groups and Performance Consequences.” *Journal of Management Studies* 43 (7): 1559–81. doi:10.1111/j.1467-6486.2006.00620.x.

Baskerville, R. F. 2003. “Hofstede Never Studied Culture.” *Accounting, Organizations and Society* 28 (1): 1–14. doi:10.1016/S0361-3682(01)00048-4.

Behn, R. D. 2003. “Why Measure Performance? Different Purposes Require Different Measures.” *Public Administration Review* 63 (5): 586–606. doi:10.1111/1540-6210.00322.

Beugelsdijk, S., R. Maseland, and A. van Hoorn. 2015. “Are Scores on Hofstede’s Dimensions of National Culture Stable over Time? A Cohort Analysis: Are Hofstede’s Culture Scores Stable over Time?” *Global Strategy Journal* 5 (3): 223–40. doi:10.1002/gsj.1098.

Butterfield, R., C. Edwards, and J. Woodall. 2004. “The New Public Management and the UK Police Service.” *Public Management Review* 6 (3): 395–415. doi:10.1080/1471903042000256556.

Cao, D., H. Li, and G. Wang. 2014. “Impacts of Isomorphic Pressures on BIM Adoption in Construction Projects.” *Journal of Construction Engineering and Management* 140 (12): 04014056. doi:10.1061/(ASCE)CO.1943-7862.0000903.

Cheung, A. B. L. 2005. “The Politics of Administrative Reforms in Asia: Paradigms and Legacies, Paths and Diversities.” *Governance* 18 (2): 257–82. doi:10.1111/j.1468-0491.2005.00275.x.

Christensen, T., and P. Lægreid. 2007. “The Whole-of-Government Approach to Public Sector Reform.” *Public Administration Review* 67 (6): 1059–66. doi:10.1111/j.1540-6210.2007.00797.x.

Christensen, T., and P. Laegreid. 2011a. “Democracy and Administrative Policy: Contrasting Elements of New Public Management (NPM) and Post-NPM.” *European Political Science Review* 3 (01): 125–46.

Christensen, T., and P. Lægreid. 2011b. “Complexity and Hybrid Public Administration—Theoretical and Empirical Challenges.” *Public Organization Review* 11 (4): 407–23. doi:10.1007/s11115-010-0141-4.

DiMaggio, P. J., and W. W. Powell. 1983. “The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields.” *American Sociological Review* 48 (2): 147–60. doi:10.2307/2095101.

Dolowitz, D. P. and D. Marsh. 2002. “Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making.” *Governance: An International Journal of Policy and Administration* 13(1): 5-24.

Eisenhardt, K. M. 1989. “Agency Theory: An Assessment and Review.” *The Academy of Management Review* 14 (1): 57–74. doi:10.2307/258191.

Elcock, H. 1995. “The Fallacies of Management.” *Public Policy and Administration* 10 (1): 34–48. doi:10.1177/095207679501000104.

Fang, T. 2003. “A Critique of Hofstede’s Fifth National Culture Dimension.” *International Journal of Cross Cultural Management* 3 (3): 347–68. doi:10.1177/1470595803003003006.

Forrester, G. 2011. “Performance Management in Education: Milestone or Millstone?” *Management in Education* 25 (1): 5–9. doi:10.1177/0892020610383902.

Fraser Institute. 2017. “Economic Freedom of the World.” *Economic Freedom: Datatset*. 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.” *Freedom in the World Data*. https://freedomhouse.org/report-types/freedom-world.

Frumkin, P., and J. Galaskiewicz. 2004. “Institutional Isomorphism and Public Sector Organizations.” *Journal of Public Administration Research and Theory: J-PART* 14 (3): 283–307.

Fund for Peace. 2017. “Fragile States Index Data | The Fund for Peace.” http://fsi.fundforpeace.org/indicators.

Gerhart, B., and M. Fang. 2005. “National Culture and Human Resource Management: Assumptions and Evidence.” *The International Journal of Human Resource Management* 16 (6): 971–86. doi:10.1080/09585190500120772.

Goldfinch, S., K. DeRouen, and P. Pospieszna. 2013. "Flying Blind? Evidence for Good Governance Public Management Reform Agendas, Implementation and Outcomes in Low Income Countries." *Public Administration and Development* 33 (1): 50–61. doi:10.1002/pad.1614.

Green-Pedersen, C. (2002). "New Public Management Reforms of the Danish and Swedish Welfare State: The Role of Different Social Democratic Responses." *Governance* 15(2): 271-294. doi:

Gwartney, J., R. Lawson, and J. Hall. 2014. *Economic Freedom of the World: 2014 Annual Report*. Washington: Fraser Institute. 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): 439–50. doi:10.1080/0305006032000162011.

He, Q., S. Dong, T. Rose, H. Li, Q. Yin, and D. Cao. 2016. "Systematic Impact of Institutional Pressures on Safety Climate in the Construction Industry." *Accident Analysis & Prevention* 93 (August): 230–39. doi:10.1016/j.aap.2015.11.034.

Hood, C. 1991. "A Public Management for All Seasons?" *Public Administration* 69 (1): 3–19. doi:10.1111/j.1467-9299.1991.tb00779.x.

Hood, C. 1995. "Contemporary Public Management: A New Global Paradigm?" *Public Policy and Administration* 10 (2): 104–17. doi:10.1177/095207679501000208.

Hofstede, G. n.d. "National Culture." https://geert-hofstede.com/national-culture.html.

Javidan, M., R. J. House, P. W. Dorfman, P. J. Hanges, and M. S. de Luque. 2006. "Conceptualizing and Measuring Cultures and Their Consequences: A Comparative

Review of GLOBE's and Hofstede's Approaches." *Journal of International Business Studies* 37 (6): 897–914.

Jones, L. R., and F. Thompson. 1999. *Public Management: Institutional Renewal for the Twenty-First Century*. JAI Press.

Kane, J., and H. Patapan. 2006. "In Search of Prudence: The Hidden Problem of Managerial Reform." *Public Administration Review* 66 (5): 711–24. doi:10.1111/j.1540-6210.2006.00636.x.

Kaufmann, D., A. Kraay, and M. Mastruzzi. 2010. "The Worldwide Governance Indicators: Methodology and Analytical Issues." SSRN Scholarly Paper ID 1682130. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=1682130.

Ko, E. Y. L. 2001. "Performance Appraisal in Higher Education : The Impact of New Managerialism on the Motivation of Academic Staff." Ph.D., University of Bristol. http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.343353.

Kraay, Aart, Daniel Kaufmann, and Massimo Mastruzzi. 2010. *The Worldwide Governance Indicators : Methodology and Analytical Issues*. Policy Research Working Papers. The World Bank. doi:10.1596/1813-9450-5430.

Ladd, H., ed. 1996. *Holding Schools Accountable: Performance-Based Reform in Education*. Washington, D.C: Brookings Institution Press.

Lee, E. W. Y., and M. S. Haque. 2006. "The New Public Management Reform and Governance in Asian NICs: A Comparison of Hong Kong and Singapore." *Governance* 19 (4): 605–26. doi:10.1111/j.1468-0491.2006.00330.x.

Leiter, J. 2005. "Structural Isomorphism in Australian Nonprofit Organizations." *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations* 16 (1): 1–31. doi:10.1007/s11266-005-3230-1.

Levačić, R. 2008. “Financing Schools: Evolving Patterns of Autonomy and Control.” *Educational Management Administration & Leadership* 36 (2): 221–34. doi:10.1177/1741143207087774.

Magala, S. 2005. *Cross-Cultural Competence*. Psychology Press.

Mather, K., and R. Seifert. 2011. “Teacher, Lecturer or Labourer? Performance Management Issues in Education.” *Management in Education* 25 (1): 26–31. doi:10.1177/0892020610388060.

McSweeney, B. 2002. “Hofstede’s Model of National Cultural Differences and Their Consequences: A Triumph of Faith - a Failure of Analysis.” *Human Relations* 55 (1): 89–118.

Meyer, J. W., and B. Rowan. 1977. “Institutionalized Organizations: Formal Structure as Myth and Ceremony.” *American Journal of Sociology* 83 (2): 340–63.

Meyer, R. E., and G. Hammerschmid. 2010. “The Degree of Decentralization and Individual Decision Making in Central Government Human Resource Management: A European Comparative Perspective.” *Public Administration* 88 (2): 455–78. doi:10.1111/j.1467-9299.2009.01798.x.

Montes, F.J. L., and A.J. V. Jover. 2004. “Total Quality Management, Institutional Isomorphism and Performance: The Case of Financial Services.” *The Service Industries Journal* 24 (5): 103–19. doi:10.1080/0264206042000276865.

Moon, M. J., and P. deLeon. 2001. “Municipal Reinvention: Managerial Values and Diffusion among Municipalities.” *Journal of Public Administration Research and Theory* 11 (3): 327–52. doi:10.1093/oxfordjournals.jpart.a003505.

Moore, M. H. 2000. “Managing for Value: Organizational Strategy in for-Profit, Nonprofit, and Governmental Organizations.” *Nonprofit and Voluntary Sector Quarterly* 29 (1_suppl): 183–204. doi:10.1177/0899764000291S009.

Nair, A., and D. Prajogo. 2009. “Internalisation of ISO 9000 Standards: The Antecedent Role of Functionalist and Institutionalist Drivers and Performance Implications†.” *International Journal of Production Research* 47 (16): 4545–68. doi:10.1080/00207540701871069.

Niskanen, W. A. 1971. *Bureaucracy and Representative Government*. 1st Pbk. Print edition. Chicago: Aldine Transaction.

OECD. 2005. *Performance-Related Pay Policies for Government Employees*. OECD Publishing.

OECD. 2009. *PISA Data Analysis Manual. SPSS*. 2nd ed. Paris: OECD.

OECD. 2011. School Questionnaire for PISA 2012. http://www.oecd.org/pisa/pisaproducts/PISA12_ScQ_ENG.pdf

OECD. 2012a. *CBA PISA 2012*. Database PISA. http://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm.

OECD. 2012b. *PISA Data Analysis Manual*. 2nd edition. http://www.ipi.lu.lv/fileadmin/_migrated/content_uploads/Datu_analizes_rk.pdf.

Ozga, J. 2003. *Measuring & Managing Performance in Education*. CES Briefing 27.

Park, S., and F. Berry. 2014. “Successful Diffusion of a Failed Policy: The Case of Pay-for-Performance in the US Federal Government.” *Public Management Review* 16 (6): 763–81. doi:10.1080/14719037.2012.750835.

Peters, B. G., and J. Pierre. 1998. “Governance without Government? Rethinking Public Administration.” *Journal of Public Administration Research and Theory: J-PART* 8 (2): 223–43.

Pitts, D. W., A. K. Hicklin, D. P. Hawes, and E. Melton. 2010. “What Drives the Implementation of Diversity Management Programs? Evidence from Public

Organizations." *Journal of Public Administration Research and Theory* 20 (4): 867–86. doi:10.1093/jopart/mup044.

Ploom, K., and T. Haldma. 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. doi:10.1108/17465261311310018.

Pollitt, C. 2000. "Is the Emperor In His Underwear?" *Public Management: An International Journal of Research and Theory* 2 (2): 181–200. doi:10.1080/14719030000000009.

Pollitt, C. 2001. "Convergence: The Useful Myth?" *Public Administration* 79 (4): 933–47. doi:10.1111/1467-9299.00287.

Pollitt, C. 2016. "Managerialism Redux?" *Financial Accountability & Management* 32 (4): 429–47. doi:10.1111/faam.12094.

Pollitt, C., and G. Bouckaert. 2011. *Public Management Reform: A Comparative Analysis - New Public Management, Governance, and the Neo-Weberian State*. 3 edition. Oxford ; New York: Oxford University Press.

Pollitt, C., and S. Dan. 2013. "Searching for Impacts in Performance-Oriented Management Reform." *Public Performance & Management Review* 37 (1): 7–32. doi:10.2753/PMR1530-9576370101.

Pretorius, S. G., and V. C. Ngwenya. 2008. "Teachers' Perceptions of and Attitudes towards Performance Appraisal in Zimbabwean Schools." *Africa Education Review* 5 (1): 144–64. doi:10.1080/18146620802144859.

Reeves, J., Christine F., J. O. Brien, P. Smith, and H. Tomlinson. 2002. "Chapter 2: The Concern with Performance and Continuing Professional Development." In *Performance Management in Education: Improving Practice*. London : SAGE Publications Ltd. doi:10.4135/9781446220931.

Shenkar, O. 2001. “Cultural Distance Revisited: Towards a More Rigorous Conceptualization and Measurement of Cultural Differences.” *Journal of International Business Studies* 32 (3): 519–35.

Smith, R. 1995. “Staff Appraisal in Higher Education — a Study of Performance Review at Nene College, Northampton.” *Higher Education* 30 (2): 189–205. doi:10.1007/BF01384096.

Søndergaard, M. 1994. “Research Note: Hofstede’s Consequences: A Study of Reviews, Citations and Replications.” *Organization Studies* 15 (3): 447–56. doi:10.1177/017084069401500307.

Storey, A. 2002. “Performance Management in Schools: Could the Balanced Scorecard Help?” *School Leadership & Management* 22 (3): 321–38. doi:10.1080/1363243022000020435.

Sozen, S., and I. Shaw. 2002. “The International Applicability of ‘new’ Public Management: Lessons from Turkey.” *International Journal of Public Sector Management* 15 (6): 475–86. doi:10.1108/09513550210439625.

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*: 345-354.

Taras, V., and P. Steel. 2009. “Beyond Hofstede: Challenging the Ten Commandments of Cross-Cultural Research.” In *Beyond Hofstede*, edited by Cheryl Nakata, 40–60. Palgrave Macmillan UK. doi:10.1057/9780230240834_3.

Taut, S., V. Santelices, C. Araya, and J. Manzi. 2010. “Theory Underlying a National Teacher Evaluation Program.” *Evaluation and Program Planning* 33 (4): 477–86. doi:10.1016/j.evalprogplan.2010.01.002.

Taylor, F. W. 1911. *The Principles of Scientific Management*. New York, London: Harper & Brothers.

Thatcher, M. 1993. *Downing Street Years*. New York: HarperCollins Publishers.

The Heritage Foundation. 2017. “Index of Economic Freedom.” http://www.heritage.org/index/explore.

Tolofari, S. 2005. “New Public Management and Education.” *Policy Futures in Education* 3 (1): 75–89. doi:10.2304/pfie.2005.3.1.11.

Touburg, G. 2016. “National Habitus: An Antidote to the Resilience of Hofstede’s ‘national Culture’?” *Journal of Organizational Change Management* 29 (1): 81–92. doi:10.1108/JOCM-11-2015-0219.

Verbruggen, S., J. Christiaens, and K. Milis. 2011. “Can Resource Dependence and Coercive Isomorphism Explain Nonprofit Organizations’ Compliance With Reporting Standards?” *Nonprofit and Voluntary Sector Quarterly* 40 (1): 5–32. doi:10.1177/0899764009355061.

Verschuere, B., E. Beddeleem, and D. Verlet. 2014. “Determinants of Innovative Behaviour in Flemish Nonprofit Organizations: An Empirical Research.” *Public Management Review* 16 (2): 173–98. doi:10.1080/14719037.2012.757347.

Villadsen, A. R., J. R. Hansen, and N. P. Mols. 2010. “When Do Public Managers Imitate Each Other? Mimetic Decision Making in Contracting Decisions of Danish Municipalities.” *Public Organization Review* 10 (4): 357–76. doi:10.1007/s11115-010-0111-x.

Vroom, V. H. 1964. *Work and Motivation*. 1 edition. New York: Wiley.

Williamson, D. 2002. “Forward from a Critique of Hofstede’s Model of National Culture.” *Human Relations* 55 (11): 1373–95. doi:10.1177/0018726702551100 6.

World Bank. 2001. World Development Report 2000/2001: Attacking Poverty. Oxford University Press: New York.

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.

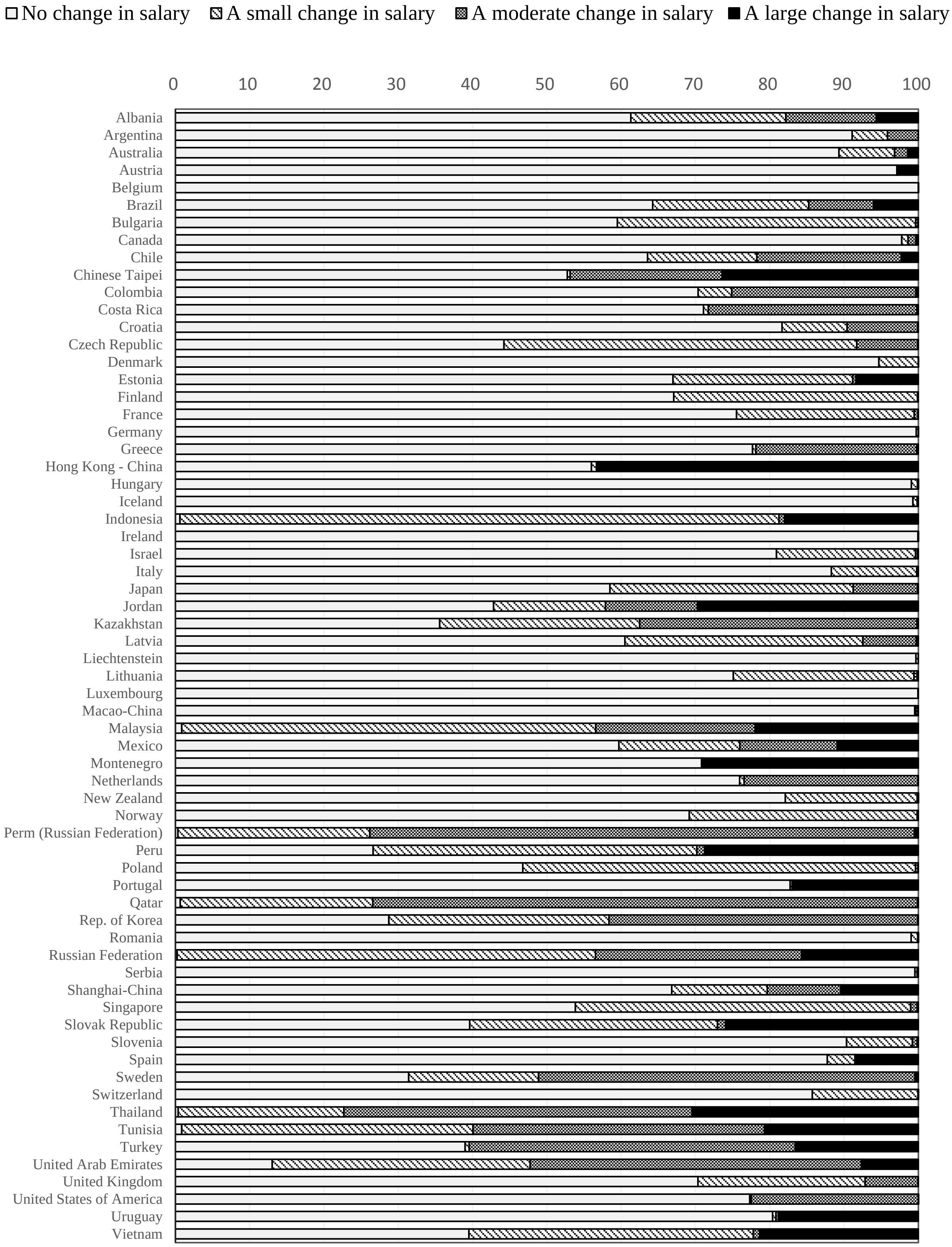

Figure 1. Percentage distribution of students whose teacher appraisals directly led to varying degrees of changes in salary, 2012.

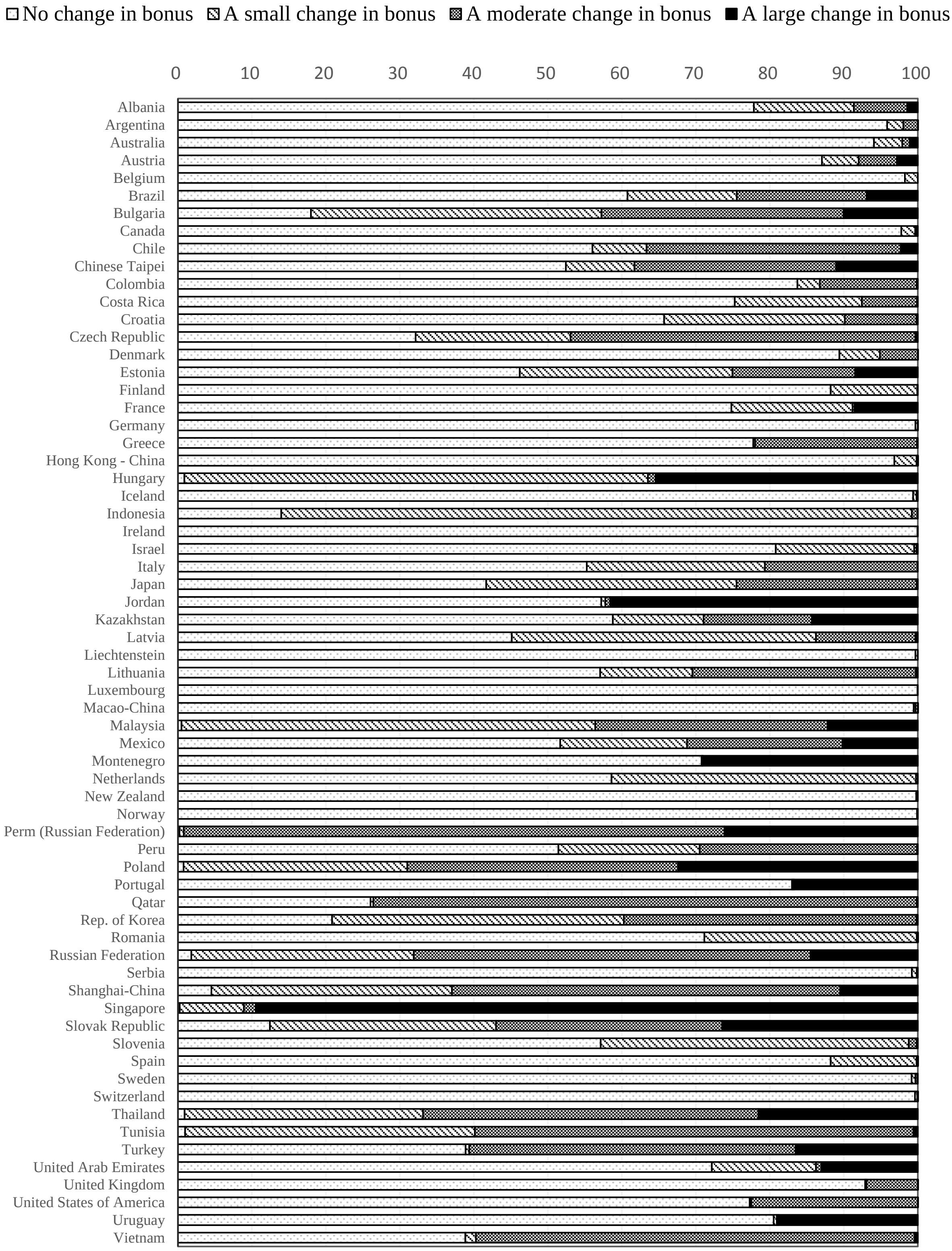

Figure 2. Percentage distribution of students whose teacher appraisals directly led to varying degrees of changes in bonuses, 2012.

Table 1. Hypothesized National Contexts and Causal Directions

Table 1. Hypothesized National Contexts and Causal Directions

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

Table 2. Descriptive Statistics

Variables		N	Min.	Max.	Mean	S.D
APPRA1	appraisals salary	467,900	0.00	3.00	0.59	0.92
APPRA2	appraisals bonus	467,306	0.00	3.00	0.64	0.96
PRIVATE	private school	478,853	0.00	1.00	0.22	0.41
NONGOV	non-government fund	435,550	0.00	100.00	78.87	32.10
TEACHR	teacher group influence	475,918	1.94	69.55	38.63	11.97
DECENT	structure	485,490	21.41	99.25	66.32	15.92
VOICE	voice and accountability	485,490	4.69	100.00	68.63	24.35
FREEDM	freedom in the world index	469,439	2.00	40.00	32.18	10.25
GNPLN	GNI (log)	472,213	7.35	11.50	10.02	0.90
EFWI	EFW index	478,101	4.92	8.98	7.32	0.59
IEFI	IEF index	472,213	48.00	89.90	68.45	8.17
POWER	power distance	462,255	11.00	100.00	57.30	20.53
INDIVI	individualism	462,255	13.00	91.00	49.69	24.28
MUSCL	masculinity	462,255	5.00	100.00	51.24	18.90
AVOID	uncertainty avoidance	462,255	8.00	100.00	68.17	21.04
OECD	OECD membership	485,490	0.00	1.00	0.63	0.48
EXTER	external intervention	467,385	0.80	7.60	3.52	1.91

Note. The units of observations are not schools, but students, because we merged school and student data, at the suggestion of the *PISA Data Analysis Manual* (OECD 2009).

Table 3. Estimation Results

	I. Salary			II. Bonuses		
	est.	s.e.	exp(β)	est.	s.e.	exp(β)
voice and accountability	-0.04	(0.01)	0.96***	-0.03	(0.02)	0.97**
EFW index	0.14	(0.28)	1.15	0.64	(0.45)	1.89
GNI (log)	-0.33	(0.24)	0.72	-0.92	(0.48)	0.40*
external intervention	-0.15	(0.14)	0.86	-0.20	(0.20)	0.82
OECD membership	1.14	(0.42)	3.12***	1.40	(0.49)	4.06***
teacher group influence	0.01	(0.01)	1.01	0.03	(0.02)	1.03**
structure	0.01	(0.01)	1.01*	0.02	(0.01)	1.02*
power distance	0.01	(0.01)	1.01	0.02	(0.01)	1.02
individualism	-0.02	(0.01)	0.98	-0.03	(0.02)	0.98
masculinity	-0.01	(0.01)	0.99	0.02	(0.01)	1.02**
uncertainty avoidance	-0.02	(0.01)	0.98**	-0.01	(0.01)	0.99
private school	-0.04	(0.41)	0.96	0.04	(0.58)	1.04
non-government fund	0.01	(<0.01)	1.01	-0.01	(0.01)	0.99
N	370,017			369,500		
PsuedoR^2	0.12			0.15		

Both models controlled for school enrolment size, location, and student-teacher ratio. Clustered robust standard errors in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, 10% level respectively.

[i]The term ‘reinventing’ government is considered to be a narrower version of NPM, which is used in the American context. It emphasizes market efficiency, public entrepreneurship, customer satisfaction, and competition (Moon and deLeon2001).

[ii]We also tried the Index of Economic Freedom (IEF) from the Heritage Foundation (2017),which led to similar conclusions about political regimes. This index measures economic freedom, but it does so on the basis of four broad pillars: (i) rule of law, (ii) limited government, (iii) regulatory efficiency, and (iv) open markets. Unlike the EFW index, each component in the IEF is measured on a scale of 100. The EFW index incorporates aspects which are unaccounted for in the IEF; these include judicial freedom, military interference, policing and crime, freedom of foreigners to visit, and black-market exchange rates. The EFW index does not explicitly account for price controls, which the IEF does.

[iii]We also tried Freedom House's (2017) 'Freedom in the World' indicator. The measure 'voice and accountability' uses this indicator, and it is based on a wider range of variables. Not surprisingly, the two indicators are highly correlated and yielded similar conclusions regarding economic ideology.

[iv]We also considered WTO membership, but all of the countries in the dataset were members, and hence, there was no variation in the membership status among countries.

[v]There are some critiques of these measures (see, for example, Søndergaard, 1994; McSweeney 2002; Williamson 2002; Baskerville 2003; Fang 2003; Beugelsdijk et al. 2015; Gerhart and Fang 2005; Magala 2005; Javidan et al. 2006; Touburg, 2016). Nevertheless, we used them because other measures of culture do not depart significantly from Hofstede's model (Taras and Steel 2009) and in the absence of more suitable alternatives, Hofstede's work continues to provide valuable insights (Williamson 2002), and was the best available for our study.