

# **The Economics and Politics of Teacher Merit Pay**

Andrew Leigh MP

**14 November 2011**

**Transcript**

---

The debate over merit pay can be summed up as follows: economists like it, voters love it, and teachers are divided. Can merit pay be made to work? Andrew Leigh MP discussed these issues with John Daley, Grattan's CEO.

Looking across the international evidence, Andrew Leigh surveyed three sets of data that are relevant to answering this question: impact studies of teacher merit pay schemes, evidence of teacher attitudes to merit pay, and surveys of attitudes in the general public to merit pay. Looking at the existing merit pay plans, one is struck by the fact that the incentive schemes are often very complicated, and most estimates are of short-run effects (so do not capture selection into the teaching profession).

Teacher attitudes are mixed, with new teachers more open to merit pay than their more experienced colleagues. US surveys find that voter support for merit pay is high and rising. Andrew Leigh will conclude with ten suggestions for future research on teacher merit pay.

**Speakers:**      **Andrew Leigh MP**  
                     **John Daley, CEO, Grattan Institute**

AUDIO: This is a podcast from Grattan Institute, [www.grattan.edu.au](http://www.grattan.edu.au).

JOHN: It's my very great pleasure to welcome you here this evening along with Andrew Lee. I'll introduce Andrew in a moment. But before I do that can I thank you all for coming. Can I thank Telstra for the opportunity to use this facility. Can I acknowledge the traditional owners of the land on which this event is taking place and pay our respects to them past and present. Can I ask those who have mobile phones to turn them off or at least turn them to silent. And with that, let me launch into tonight's seminar. Obviously, Grattan Institute runs five programs and two of them are about education, specifically higher education and school education, and so this is a topic near and dear to our hearts particularly because we also run a program in productivity growth and there's plenty of work to suggest that one of the biggest single levers to improving productivity and therefore wellbeing is to improve the quality of education.

So we're very lucky to have with us tonight Andrew Lee. Andrew Lee is an extremely rare beast in many ways. He is without doubt one of the country's most gifted economists and recently won the Australian Economic Society prize for the best young economist, that young economist being defined as under 40 I was interested that ... I don't know whether that means you're about to become old or ...

ANDREW: It says something about a profession doesn't it.

JOHN: But without doubt a terrific economist has interests that span across and cross an enormous variety of fields. Everything from his book on ... sorry, his work on whether certain names make you more likely to get a job offer or at least a job interview. It turns out that having a foreign sounding name makes it less likely that you'll get interviewed unless it's an Italian name, in which case you're more likely to get interviewed, at least for jobs in the food sector.

ANDREW: In Melbourne.

JOHN: And everything from that to doing work on the impact of recessions, on death rates and of course his book *Disconnected*, which looked at the way that our society is more or less involved in terms of people being involved with each other and connected to each other. He has a doctorate from Harvard in economics. Honours degrees in law and arts from University of Sydney? Yes, and I'm not sure which is more incongruous frankly given that Andrew is now the Member for Fraser, which is one of the two electorates in Canberra, whether it's more incongruous to see an economist with a doctorate from Harvard on the green benches, or whether it's to see a picture of Andrew wearing a hard hat on his website. So I'll leave Andrew

to answer which of those he finds more surprising. But without doubt, when Andrew moved into politics Australian economics lost a great deal but politics gained a lot in the process.

However we haven't lost Andrew entirely from the field of serious economics. He still manages to find time to do some work in the field and as one would expect from Andrew it continues to be terrific work and that's what we're here to hear about tonight. Andrew as well as working on the impact of recessions on death rates has always had a very keen interest in education, and a very, very keen interest in evidence based policy, in trying to work out how do you work out in a rigorous way whether policy is truly making a difference and in particular is making a difference in the direction you want it to have.

So tonight in many ways encompasses perhaps the two central interests of Andrew's work over the last couple of decades, education and getting the evidence. And his topic for tonight is around the economics and politics of teacher merit pay. So Andrew, starting as we always should do when we're talking about economics, ie not with economics at all but with philosophy and definitions. When you say merit based pay as my professors in Oxford used to say, what do you mean?

ANDREW: Well thanks John for that extraordinarily generous introduction. Can I also acknowledge we're meeting on the traditional lands of the indigenous peoples and pay my respects to their elders past and present, and also to thank Grattan for a somewhat unusual style of presentation. When I've given economics talks in the past they've been 80 PowerPoint slides after one another, so this is the first time I've done an Oprah Winfrey style presentation of a piece of economic research, so thank you for that.

So what I have in mind with teacher merit pay is schemes which pay more for doing a better job in the classroom. Now what does doing a better job mean? It can involve raising test scores, involving ... improving morale, boosting student performance, mentoring junior teachers. But I have in mind schemes that have a pay bonus for doing a better job as distinct from some of the other schemes that we see out there. So there are for example schemes which pay more for having a different academic qualification, for teaching in a remote school or for teaching a subject which is hard to staff. I don't think of that as merit pay.

JOHN: Those aren't bad things, it's just not what we're looking at.

ANDREW: Exactly yeah, so let's define it down, it's being paid more for doing a better job in the classroom. And you know, you can tell stories as to why this would work well and stories as to why it would be a complete disaster. So typically people who want to tell a story about why merit pay is good will point to three things. They'll say that there is a big dispersion in the effectiveness of teachers, so for example if you look at test score value added, the top 90% seem to raise scores twice as fast as the bottom 10%. They'll say that the effectiveness of teachers doesn't track experience particularly well, so a salary schedule which is just based on experience isn't capturing much of the variation between teachers and they'll point to the fact that there's a lot of pay dispersion in other occupations.

So when you look at more effective teachers in the classroom, who then leave the teaching profession, they earn a higher wage, the less effective teachers in the classroom. So you can see the sort of fanning out in salaries for exiting teachers but a lot of pay compression for those who are in there. But then there's the critics. The critics will say it's all about family background, so really your merit pay scheme isn't doing anything other than just paying for family background or for luck. They'll say it's just going to lead to teaching to the test. It'll lead to favouritism if you run it through principals, it'll be too narrow if you run it through test scores, and if you pay the whole school then it's just going to lead to free riding, so teachers will say I'm just one of 60, why should I work harder. So there's a lot of theory out there and what I wanted to try and do in this paper was to say okay two sets of theories, what has the evidence got to say.

JOHN: Can we just take that back half a step, let's at least look at the theories that claim there's some sort of plausible argument like this might work. Are they arguing that this will essentially lead to more good people staying in the teaching profession? Or are they arguing that it will lead to more people who are in the teaching profession trying harder?

ANDREW: Great question. So most of the arguments that economists make around merit pay is that it will operate through selection. That is it will encourage great teachers to stay in the profession and it will attract great teachers into the profession.

JOHN: And not leave.

ANDREW: And not leave, exactly. So it'll work through what economists call selection effects, rather than through what economists would call incentive effects. So that partly explains why you see some degree of frustration among teachers at merit pay. I think teachers look at ... many teachers look at merit pay and say I can't work any harder than I am now, this is as good as I can do. And I think that's partly right, the incentive impacts tend to be fairly small, the selection impacts I think could be pretty big.

JOHN: Okay, so that's what we're trying to look at and that's at least the plausible arguments why it might work or might not work. What sort of evidence is there out there?

ANDREW: So there's less than I thought. I had this idea before going in that there would be a whole lot of really high quality studies. And it turns out that if you use the ... I hate the term gold standard, 'cause we went off the gold standard 40 years ago, but if you look at ...

JOHN: We could be going back on.

ANDREW: Yeah that's true, that's true. So let's look at gold standard randomised trials, then you've only got a handful. You've got studies in Nashville, New York, Andhra Pradesh and Kenya. And you see mixed results out of those, so the randomised trials, the Andhra Pradesh one is showing very promising results. New York and Nashville not so much. The Kenyan study is showing some results but fade out. What seems to matter here is the complexity of the scheme. So some of these schemes are really, really tough to understand. You put yourself in the shoes of a teacher who's faced with this merit pay scheme like the New York one, which is based on surveys from principals, from teachers, from students, it's based on some test scores, it's based on attendance records, it's all averaged up. You're compared against all the schools in New York, then there's a group of peer schools and your comparison versus your peer schools is worth triple the weight of your comparison compared to the rest of the schools in New York. Is anyone still following the way the New York scheme works? So it's not that surprising that ... to me at least, that the New York scheme isn't getting big results. The Andhra Pradesh one is super simple and seems to get better results. So I think part of it is not designing a scheme that teachers can't understand.

JOHN: If you can just give us some idea of the scale, 'cause I guess we can think about something say around class sizes where I think Ben's recent review of the evidence came up with something like 80 or 90 different studies. How many have we got here and how many of them at least pass a silver standard, and how many of them are frankly not a lot better than shooting darts at a dartboard?

ANDREW: Right, so I kind of stop at silver. My gold standard would be randomised trials, my silver standard would be natural experiments, so you've got a school that implements merit pay. You see that school before and after, you've got some other bunch of schools that don't implement merit pay, you see them before and after. Or some sort of a regression discontinuity, something like that. But I think I want what economists would call a natural experiment. So beyond the four randomised trials I mentioned, there's seven natural experiments, I think two of them are a bit rosey so that leaves me with five and those five find positive effects of varying sizes. So there's enough evidence there to suggest that it might work, but there's also enough evidence there to suggest that anyone who tells you merit pay always works or merit pay never works is either lying or they haven't looked at the evidence. It's sufficiently mixed to suggest that some kinds of merit pay work and some kinds probably don't.

JOHN: You've spoken a little bit about the ideal set up that you think it might have something to do with the non-complexity of the scheme. Does the evidence allow you to say anything else about what might be involved?

ANDREW: I think the size of the bonus shouldn't matter, so there's schemes which ... for example one of the schemes has prizes so doing ... when you're ... when you do well in the merit pay scheme that doesn't get you a pay bonus, it gets you a lottery ticket and then sometimes the lottery ticket pays off and sometimes it doesn't. That's not a scheme that seems to have any substantial impact. And there's other schemes which are based on group incentives, so if we ... if our entire school does well then we'll all get a bonus and there seems ... what's nice about Carthic Mollerin's Andhra Pradesh study is he actually has both an individual incentive and a group incentive, same randomised set up, looking at both of them, and the effects from the individual incentive are just streets ahead of the group incentive. For those who thought that providing a bonus to an entire school would maintain the sort of esprit décor of the staff room, that people would then share teaching notes and share resources and so on, that effect seems to be offset by the fact that in a big school, if it were all getting a bonus, there's an incentive for individual teachers just to say look I don't think it's worth me putting in effort on this one, so ...

JOHN: Does that suggest in fact it's not a selection effect, it is in fact an incentive effect?

ANDREW: There's clearly ... so Carthic can only be capturing incentive effects. We basically ... the only selection stories we can be getting are from a couple of studies which just compare either US schools or OECD countries and determined that those with merit pay schemes seem to have higher test scores. Those studies could be getting selection effects. The ones I've mentioned already are only operating off the incentive margin, and they're getting effects somewhere between a couple of months to a full year of learning out of the merit pay scheme. So if you were getting selection effects then I guess you'd be looking at something bigger than that.

JOHN: I mean I guess it does raise the question, 'cause there's an interaction between a performance reward scheme ie if you perform well then you will get some kind of bonus whatever it might be. And a scheme that is aimed at improving performance feedback, and inherently those things are linked. It's very difficult for me to provide you with a bonus unless I've got some view as to what your performance is. If I've got some view as to your performance at least the option, one hopes, to provide you with some kind of feedback about how you've done and what you could do better and indeed there's at least the possibility of something more sophisticated that provides you with interim feedback on the way through and then sums that up at the end of the year. Have any of the studies kind of looked at what sort of performance feedback was going on at the same time, given that those things are in fact quite closely linked?

ANDREW: Yeah so in the area of performance feedback is I think under-explored by economists. It would be nice to have Ben Jensen here who's sent apologies, he's working on early childhood investment at the moment ie kid minding. And ... but Ben tells me that the evidence around performance feedback is what I would regard as bronze standard evidence, so it's a before/after study, it finds some effects there, but I don't think we have much more in the area of performance feedback. I'd like to see a little bit more. You can tell ... so the plausible story you'd tell around why performance feedback works is that when you have say what we thought of as the old school inspector in the back of the classroom, they give teachers tips that they didn't realise beforehand and teachers are then able to kind of improve. The story that performance feedback wouldn't work is if the real challenge was the amount of time the teachers put into teaching, if it was something around the sort of selection into the profession, those sorts of stories it's harder to imagine that just providing more information is the answer and more likely that maybe we have to actually change something around the pay structure. But changing the pay structure's hard right, and I guess this is ... the whole message that I get out of this, is this is a thorny area and there's less evidence around than you think.

JOHN: I mean I think it is perhaps worth starting to look at and I don't know that anyone really has tried to apply this to an education experience looking at the more general corporate literature around performance feedback, 'cause there's no ... I don't think there's any particular reason to believe that education is a somehow sacredly different exercise from any other form of employment so that the kinds of performance feedback that are now common in many professions and many ... indeed in work places, are therefore necessarily inappropriate to



teaching, and I think there probably is a wider body of evidence around that although I'd agree with you much of it may not pass your silver standard let alone your gold standard.

ANDREW: Yeah it's a low evidence level on performance feedback. I'd like to see more.

JOHN: Fair enough. So we've talked about performance feedback and the evidence as to whether it does or doesn't work. What do teachers think about this?

ANDREW: So I've gone into this basically thinking that the surveys would show that 99% of teachers thought that performance pay was a complete disaster, and the other 1% didn't understand the question. It actually turns out that teachers are more positive towards performance pay in general. Most of the best surveys, as with the impact evaluations are coming out of US teachers, but if you phrase the question should teachers get additional compensation if their students show outstanding growth in test scores, two thirds of American teachers will say yes. If you use the word merit pay then it's a bit of a red rag to a bull and you will get typically less than 50%. You'll get more support among teachers for a scheme where the principal is making the decision than a scheme that is based purely on test scores. The only information we have out of Australia is a recent question that said would merit pay help retention and there you only get about a quarter to a third of teachers saying yes. So on average, teachers are pretty lukewarm towards merit pay, but then there's these interesting differences among teachers.

JOHN: Yeah, and so which teachers like it more?

ANDREW: So the teachers who are most positive about merit pay are teachers who are younger, male teachers are more supportive than female teachers and in the US, minority teachers, African American and Hispanic teachers tend to be more supportive of merit pay. Teachers who've gone through a merit pay scheme either seem to like it as much as they did before or sometimes a little better, so exposure to merit pay doesn't seem to turn you off it. The stuff we really want to know is whether more effective teachers are also more positive towards merit pay, no-one's been able to demonstrate that. And also learning something a little more about what it does to collegiality. So I think for me the most potent criticism of merit pay is it'll turn the staffroom from a terrifically warm, sharing, mentoring environment into a sort of law of the jungle environment in which everyone sits in their corners, shepherds their notes tightly and doesn't help the new teachers do better, and if we could get a bit more out of the surveys I think on that one.

JOHN: I think that's an interesting criticism, 'cause one I've heard and I think it presumably depends very much on how you implement it, and ... 'cause if you think again about corporate analogies, and there are plenty of corporate performance management systems where you're essentially one of the components, is your contribution or otherwise to the team, your individual contribution to the collective team and indeed workplaces that have implemented that well seem to be perfectly cooperative places, and indeed my own personal experience and I admit it doesn't count as evidence, but is that workplaces that actually were the most dysfunctional were often the ones that had the least effective performance management systems because you have people sitting around going well I'm working really hard, Andrew's slacking off, Andrew's getting paid what I'm getting paid, why should I lift a finger? And that kind of resentment actually is very, very corrosive or can be very corrosive for team spirit. As I said, you'd want to go and design a proper experiment to find out whether that was true, but at least at an anecdotal level you can tell a pretty plausible story about how that works in practice.

ANDREW: Yeah I think that's right and it's worth sort of learning what we can from a corporate environment, while recognising that there's things that you have to do differently in teaching. It's ultimately a government run occupation so kind of rules around fairness I think have to meet a higher bar than they do if you were setting up a performance pay scheme for the local milk bar where you could be a little more ad-hoc and experimental.

JOHN: So we've talked about whether or not teachers like it, although I guess it would be worth just kind of picking that apart for a moment and saying this is something about which teachers are pretty lukewarm. By and large, these schemes are set up in such a way that they're not

going backwards. You know it's really only an opportunity for at least some people to go forwards. Why do you think that teachers are relatively speaking less excited about this?

ANDREW: So there's a lovely experiment by a Stanford grad student who gets teacher education, students and law students at Stanford in to play little economic games, and is able to learn from that how much weight they place on equality, and the teacher education students place much more weight on equal outcomes in the economic games than do the law students. So people who are selecting into teaching I think often have equality pretty high in their social welfare function, they've chosen teaching because they want to make a difference rather than making a buck, and I think that then encourages teachers to perhaps think more about wanting an environment in which people are paid the same rather than an environment in which the best earns twice what the worst earns.

JOHN: And what about school government departments? So what do they think about merit pay?

ANDREW: Hugely mixed views. So you ... I can see Nicole nodding there. You get very disparate views on merit pay. I'm just struck how much theory rules the roost in this area. Teacher merit pay is a little bit like indigenous affairs, there's so many participants in the public debate are driven by first principles rather than what you see out of on the ground evidence.

JOHN: Okay, so teachers a bit indifferent, school government departments very mixed. What about voters?

ANDREW: So voters ... finally some good news, voters like merit pay. There's ... it'll poll anything from two thirds to three quarters support in a US poll, largely unaffected by whether it's principles or test scores. I think because they're not particularly salient differences to the typical voter in the way that they're very salient to the typical teacher. But the sort of politics of merit pay means that they seem to be the first thing that's cut when the downturn comes. So for example one study of merit pay said well let's look at all the merit pay schemes in the US at a certain date, then let's come back a decade later and see how many are still in existence. And the answer is 25%, 75% of teacher merit schemes don't last the decade. And that's I think because they're cut, lost, they're the icing on the cake rather than the cake itself. Or else they're set up in such a way that they're merit pay in name only. My favourite is this Arizona merit pay scheme which required that ... said that teachers could get a bonus if students knew 10% more algebra at the end of the year than they did at the start of the year. Now at the start of the year they'd never studied any algebra. At the end of the year the typical student knows 90% more algebra than they did at the start of the year. So the test is can you do better than teaching one ninth as much algebra as the average teacher. So you can set up what is technically a merit pay bar that's really just ankle height to step over, and there's sort of myriad examples of watering down, changing ... destroying merit pay schemes, despite voters liking it. And you can't do that watering down in the same way with other reforms. So if you go to the election promising to cut class size, then pretty much voters can see whether you've cut class size and they can see if you try and put it back up the next year, that's not the case for merit pay.

JOHN: So you're saying there's a lack of transparency to voters and that transparency is maybe needed to really make it stick?

ANDREW: Yeah that's right, so the challenge of designing a scheme which is I don't know, I guess the simplicity thing matters as well because simplicity buys you a little bit of transparency, it's clean, people can see it, but you know in order to get simplicity you have to push back against the desire to get perfection. So for the kind of three econometricians in the room, if you're fitting an equation and you want to add some more variables you can almost always add some more variables that improve your goodness of fit. But a typical econometric equation will have what's called a penalty function. It'll say well if you want to add on another variable you've got to pay a price for putting that in. The typical merit pay scheme doesn't seem to make you pay a price for adding another variable. Why don't we add attendance to the equation. Why don't we also add a rating from what the school cleaner thinks of all the teachers. You know, you just keep on adding these things in until eventually it's not a particularly ... it's so complicated that teachers and possibly voters don't understand it.

JOHN: Now Andrew, one of the kind of obvious elephants in this particular room is we have a federal minister who's proposing a merit pay type scheme. Is that going to tick any of the boxes that have come out of the research that you've done?

ANDREW: So I certainly hope so, it's more than proposing. We went to the last election with a performance pay plan and that's at the moment in progress. We have the ... a lot of input from stakeholder bodies and process and I guess my feeling was that this is mostly a fun academic opportunity for me to delve into an interesting area of research but potentially it also plays some role in informing that. And making the point that I think is hard to make in this context, which is that the perfect complex system may actually be worse than the imperfect simple one.

JOHN: I guess the history might suggest that the most likely outcome is that voters will love it, teachers will be indifferent and education departments will gradually find reasons not to do it.

ANDREW: Right, right. And you know we've seen this story before. So Australia in the 19<sup>th</sup> century picks up the British merit pay model and that's kind of ... it's pretty much a red meat merit pay scheme. You get paid for having higher test scores, no adjustment for student background or anything like this, and that scheme dies because it's fundamentally pretty unfair, so teachers just want to teach affluent kids 'cause you get paid more the better the kids do. There's no adjustment for SES. And in fact I think it's David Day's biography of Curtin says that one of the early politicising events for Curtin is hearing his teacher talk about getting rid of this unjust merit pay scheme. So there's a Labor Party heritage in the merit pay story. I'm just not quite sure it's the right one from where I'm speaking.

JOHN: So Andrew, we know a little bit about all of this. If you really could get someone to do all of the natural experiments you could, what would you be trying to do, what would you like to know?

ANDREW: So first you'd want it to be randomised, so Victoria I'm not sure how well this is known but as I understand it Victoria is implementing the only randomised merit pay experiment which is looking at whether these sorts of incentives work.

JOHN: So by randomised you mean we've randomly selected which schools it will apply to and which schools not?

ANDREW: Yes exactly, so we test it in the same way as we test a new pharmaceutical, and the advantage of that is otherwise if you compare a school that wants merit pay and a school that doesn't want merit pay you might be learning something different about the kinds of teachers or principles at the school that wants it rather than about the general effect of merit pay. So you do a randomised evaluation because it lets you learn the effect of the policy rather than the effect of the kinds of people who want it. And then you'd want to try and think about a couple of different models. In my ideal world I think we'd have a few merit pay trials, but this is obviously time consuming as well, so there's always a trade off as to how much evidence you want to gather or how much you just want to go ahead and implement the policy.

JOHN: So if you could implement some randomised trials and somehow we ... there wasn't a political imperative to get this in place immediately, what would you want to do with those trials, what would you ... what kind of variables would you be wanting to be deliberately setting up so you could find out the ideal design of the scheme?

ANDREW: So you want to look at all the things the proponents and the critics say about merit pay. So the proponents say that it will raise test scores in the high stakes subjects but also that it'll raise test scores in the low stakes subjects, so you know, maybe you have a merit pay scheme in which your science score is irrelevant to the bonus, but then you want to track science scores, 'cause if your science scores are going down while your math scores are going up then your whole education outcome might not be going up. You probably want to look at teacher cheating, as occasionally you hear these stories that your merit pay will create incentives for teachers to cheat. A bunch of the studies I've talked about actually look quite carefully at this, none of them find any evidence of teacher cheating, but you should look at that.



We also probably want to know something about what it does for the camaraderie in the staffroom. Is this turning the school into a more competitive place or is it staying a cooperative environment, and then what happens when the scheme finishes? So did the scores just drop back down, were you just getting a sort of day of test blip or have you actually produced something real in the long run, so a few of the things I'd look at.

JOHN: So if there are any budding economists in the room who are looking for a quick research program there's one there.

ANDREW: Exactly and you know Melbourne is doing interesting randomised trials. You have a randomised child in early childhood intervention at the moment as well, so you're well on the way to becoming Australia's social laboratory.

JOHN: Terrific. Well Andrew that's been a great excursion through this area of teacher merit based pay and I wonder if we could now throw it open to the floor.

ANDREW: Absolutely.

JOHN: And see if there's any questions. So I think there's a mic on its way roving down quickly, terrific and we've got one just at the front here thank you.

Q: Andrew, I'm wondering how these randomised trials will work in Victoria, 'cause if you're saying that selection is the main effect here, won't this just encourage the talented teachers to seek out these particularly good schools rather than changing the behaviour of people within the school? The second element, if there's a long term selection effect here, ie affecting the people who enter the profession, isn't the key issue the credibility of a long term nature of the merit pay scheme? If you don't believe it's going to last then you'll still choose the alternative career.

ANDREW: Yeah so the answer to your second question must be yes, that ... so the political sustainability of the schemes has got to be ... is intimately tied to the selection effects, and we shouldn't expect to see them otherwise. In terms of movement across schools, that feels to me like something you'd actually want to measure, so this is I guess the point I was making before about the kinds of teachers who like and don't like merit pay schemes. We don't know very much as to whether more effective teachers like or loathe merit pay. So if we could observe that ... if we could observe not only the effect of a merit pays scheme on the school that had it but also on the applicant pool and the people who are putting up their hands to transfer out, that would be a second valuable piece of information to know. If it's making people work a bit harder but all the superstar teachers in the school are leaving, then this probably isn't something you want to roll out state wide. Thank you.

JOHN: Next? Yes thank you, one right here in the centre.

Q: I guess ... thank you for that talk, that was really interesting. I'm curious about how important you think it is to have longitudinal studies in terms of this, because I'm a maths teacher and one thing I find is that sometimes students have knowledge which is fragile, because they've been hurried to learn something very quickly, and so what they do is instead of having a deeper understanding of the topic they learn to feign performance on particular tests and stuff like that. And I just wonder how if there will be subtle effects where teachers get good at pushing students to feign knowledge and performance and whether you'll pick that up in short tests that are not longitudinal. And if this truly is Australia's social laboratory, there should be an ethics committee deciding on what you're doing ... what effect this will have on the children so should there be a precautionary principle here?

ANDREW: So certainly all of these studies should go through ethics committees. There will either be ... if you've got university researchers involved they will have to go through an ethics committee per university. My understanding is that most government departments have similar processes, but I don't know the government departments as well. But your point about the fragility of test scores is a really important one. I've been struck by two findings. One is David Figlio's paper on eating to the test, in which he shows that the effect of ... one of the effects of high stakes testing in the US is to change the free ... the choice of foods in the free lunch

program, and the choice of high GI foods on the day of the test and schools doing strategic reallocation to make sure the menu on test day is optimal for test takers. The other one is Steve Levit's result which is that if you tell students just as they sit down in a test before they open the exam paper that they will receive a bonus ... I can't remember if it's like mobile phone credits or \$50 or something, you get this boost in learning that's equivalent to a couple of months of education. Now they can't possibly have learned anything more because they only get told about this incentive as they sit down in the exam room, but suddenly I guess they bump up their effort levels a little bit on the day, they focus a little bit more. But again that's got to be your kind of fragile learning. This can't be anything real. So we want to be aware of this stuff, but I don't think it should make us averse to anything that's ever measured numerically, I don't think it should turn us off kind of objective measures. We want to come back later, we want to measure the low stake subjects, the Andre Pradesh study which is ... so if you're an Indian policymaker I think this whole merit pay thing is much easier because the Andre Pradesh study is so well ... so rigorous and so convincing in its results. It gets no effect on cheating, big effect on the tested subjects, big effect on the non-tested subjects which have nothing to do with the merit bonus, and sustained for five years into the program. But the ... other studies have found fade out and that should worry us.

JOHN: One up the back there thank you.

Q: Hey Andrew, I've been teaching for about a year now and I came into the teaching profession very pro-merit pay, teacher performance, business incentives in the classroom etc. I've sort of changed my mind to some extent. At heart I believe in performance pay, but I think the leap between policy and actual ... the classroom is so vast, and I think that there's so many factors involved that I think there needs to be more thought put into. I think performance pay is very divisive, it's heavily premised on the culture and community of the school, small. So in effect I think it's insignificant. And the most important thing is how we measure it, and I want you to expand on that, because I get the feeling that when people start talking about results you get this whole 1984 effect, that it's all about results, and that's very important, but I think when you get into schools, teaching is a lot more than academic results, it's about relationships, it's about building human beings, and the factors involved in measuring this are so vast, I just think we're almost creating more problems than we ... you know, need to be creating. So could you just elaborate more on how you think we would measure performance pay to a point where it wouldn't be causing more harm than good.

ANDREW: Yeah, so look the teaching to the test thing is absolutely critical. If we had a perfect test, teaching to the test would be no issue. If our test could measure creativity and love of learning and all of those sort of expansive things that we know really matter, but we don't. We have narrower tests that measure literacy and numeracy and so you don't ... you want some recognition that what we're measuring is relatively narrow, but also a recognition that particularly in a small school, so when I talk to education ministers about this they say that they struggle to ... they're not confident that they have a sufficiently large base of principles who have the skills and the desire to take on managing salaries as well as all of the work that they take on as well, so it's that sort of challenge as to whether you want to put it in the hands of principles. One thing that seems to be relatively popular is the idea of school inspectors. Oddly enough both my libertarian friends and my union friends seem to quite like the idea of school inspectors, and maybe that's the idea, that a school inspector gets to see hundreds of teachers over the school year, whereas even a principal only gets to see dozens of teachers. Do you have views though on ... I mean if we put your feet to the fire and made you design a scheme as to how you'd design it?

Q: Yeah actually 'cause I'm a Teacher Australia associate, so I only joined the program because I was following the Gates initiative in America very closely, and I actually joined the program based on this idea of you know, bringing incentives into teaching and the like, 'cause my mum's a teacher and I thought I had some insight into it. If I was to design a scheme it would be based on the open door classroom policy, so you'd be filming teachers. It would all be independent, 'cause I know after now working in a school if you gave the principal \$50,000 to hand out, most public schools are strapped for cash, and that money wouldn't be going to teachers who are value adding in the classroom so to speak, it would be going to teachers who are pulling the extra yards and doing the extra leg work, but there just isn't money to pay staff

for. So running extracurricular activities and the likes, which are very valuable, but I think performance pay is all about doing the best for I guess students from low socioeconomic and those who are in schools that are struggling so to speak. So it would all be based on open door classroom, you'd be filming teachers, you'd be monitoring what students think what other staff think. But until you have a fairly rigorous system in place, that's consistent and independent, it's just not going to be effective and it's going to cause more problems than it should. Saying that I'm a strong believer in the system, but two points firstly, there just isn't enough money going around the system and I'm not all about teachers earning hundreds of thousands of dollars, but I do think the pay just is really inadequate. And second thing I think before you even get to merit pay there are too many issues that should take precedence, like there just isn't enough money for students how really need it, whether you talk about VCALs or whether you talk about supporting graduate teachers or other areas which I think desperately need the money more so than encouraging I guess incentives inside the classroom. So I just think there's a lot more issues that need to be considered before this even starts getting rolled out. And I'm a bit worried that it's starting to happen in Victoria now before all these big questions are being asked.

JOHN: Well I guess that's something that the national laboratory in Victoria might come to.

ANDREW: Yeah we should also be clear about the quantum of education funding, I mean we are nowhere near 1% of the education budget going towards merit pay, we're talking about small fractions of 1% being put out on trials. But I think that the general point that we should think about this in the context of performance management is important as well.

JOHN: Thank you. We had a question at the front here please, and then back.

Q: I'm just wondering what it is possible to say about quantum, so assuming there is an extra cost, does the size of the merit pay differential have much influence on the outcomes, and also given the size of the cost, how does it compare with the potential size of the benefits?

ANDREW: So the quantum isn't ... doesn't seem to vary enough for me to say anything useful there, the bonuses are somewhere between two and 10% of salary in most of these schemes. I can't see a clear pattern coming out of the size of bonus and the effectiveness of the scheme. The size of effects you get, some were between a couple of months and a year of learning. That's bigger than you get for example ... so if we take Tennessee project star class size experiment as kind of being the upper bound for what you could hope to get from a class size reduction, they dropped class sizes by a sixth, they get a couple more months of learning. Instead of dropping class sizes by a sixth you could for example give a 12<sup>th</sup> of teachers 50% bonuses. My read of the literature is that would probably have a bigger impact on student test scores but it's all got to do with being able to design the program well.

JOHN: Thank you, yes there, please, and then we'll come down to the front here, please.

Q: I'd like to ask Andrew to go into a little bit more detail about this current proposal for a bonus pay scheme in Australia. I don't know how many times these schemes have to fail before people get the message, but the research as I understand it in the last couple of years has come up with in New York, in Chicago, in Tennessee, Nashville, Texas, all these schemes have failed to show any great benefit. But the scheme that's been proposed by the Labor Party for ... just before the last election had all the hallmarks of something on the back of an envelope. It really was a poorly thought out scheme. It listed I think four bases for assessing teachers. One was value added methods, another one was parental feedback, another one was qualifications. None of these have been tested or developed to any ... to the degree that would be necessary. The idea of assessing 200 ... is it 10% of teachers per year to get a bonus? That means 250,000 teachers in Australia have to be assessed every year to identify the 25,000 who are going to get a 10% bonus? This is an enormous ... when we do not have the methods for assessing teachers at present. Value added methods are un ... are not possible in Australia under current arrangements. Students are tested in Grade 3, 5, 7 and 9, it is impossible to use those even for ... even if you ... they only focus on literacy and numeracy, so we are promising something here which is just not implementable.

JOHN: Alright, I think we've got the point of the question.

Q: And it is a mistake to say that the Victorian scheme has been working, it has not.

JOHN: So I think we've got the point of your question, thank you.

ANDREW: So certainly the Victorian trial as I understand it, I'm not sure what its fate has been with the change of government, but I was lauding it as something that might teach us something about the effectiveness of merit pay in a context in which the evidence is really mixed. So we certainly have schemes that have worked well, the evidence out of all of the decent natural experiments and a couple of the randomised trials points towards effectiveness. And as you say there's schemes that haven't worked particularly well. So the Nashville study, the New York study don't seem to produce results. I think the challenge is not trying to do a simple up or down is merit pay a disaster or the solution to all problems in the world, but to get to a more nuanced why do some kind of merit pay schemes seem to work and why do some not, and you know, your point about value added methodology is important as well. Do you want to build test scores in and if so, what are you going to do about the fact that Australian kids are tested not only biannually as you say but also in the middle of the school year which makes attribution to teachers very, very tough.

JOHN: Thank you, I have a question here, please.

Q: Hello, I'm a union organiser so I work in the public sector, I said it very loudly so you can see my biases very clearly.

ANDREW: That's alright, I represent a party born out of the trade unions.

Q: That's great, but if you ... no, I'll tell you a couple of tales about things that I've witnessed in my schools, I work across a large number of schools in the north western suburbs of Melbourne, and recently I spoke to one of my non-teaching members in a school who is in a special school who works with students with a range of intellectual disabilities and physical disabilities. And she said to me that 10 years ago that school when the students would finish their Year 12, 70% of those students would go onto other sort of institutional settings and 30% would go onto further education or to jobs, and she said after a period of 10 years and the efforts that they've made in that school, now 30% of their students go onto other institutional settings and 70% go on to further education or jobs, that's an extraordinary outcome, and it's an effort that's been put in by both the teaching staff and the non-teaching staff in that school. And it's just wonderful to note, it's a great example. I also went out to another school very recently where a teacher pointed out a young girl to me and said she's had some very difficult times, this is the third or fourth school in her secondary school career. This is the first time in her secondary school career that she's been able to make friends, she's been able to talk to people. She's not a terrific academic performer but she's actually attending school, she's making enormous achievements. Now I guess what I'd like to know, Andrew, in a very round about way is can we please have a scheme that recognises those things, and I'm yet to see one that actually does that. And I'd also like to know that can we please have a scheme that truly actually rewards teachers above and beyond reasonable pay outcomes not in lieu of reasonable pay outcomes.

ANDREW: Yeah no I think that's absolutely right, you certainly see ... if you track teacher salaries relative to salaries of other university graduates, you see a decline since late 1970s which is about where I can take it back to, and that's true if you use other comparators. So there's an average salary challenge there before you get to the merit pay challenge. And the issue of special needs kids is a live one not just in merit pay but in other areas, so I've spoken to one of the principals in my electorate who said that the mother of a young boy with autism spectrum disorder said we think he's got to the stage where he's going to be able to sit the NAPLAN test next year. And she said she had this sort of simultaneous moment of yes and no, she was thinking yes that's wonderful for the boy, but that will also bring down the school's average NAPLAN results because it's very unlikely that he will score better than their average. So we need to do straightforward things like getting national definitions of special needs kids, which I know Barry McGore is working on, and then trying to look at the change in performance of individual kids and ... rather than trying to simply look at averages which don't tell you



anything like as much. And the reasons you've pointed to would weigh more in favour of I guess a principal rating scheme presumably or an inspector scheme that takes into account the sort of nuances of what's happening with special needs kids, and I guess what you're saying in this context is no test based scheme is going to do that properly.

Q: I can't think of a scheme that would be effective in that.

ANDREW: Why not ... just for the sake of argument, why not an inspector based scheme? Couldn't an ... just because the inspector wouldn't see enough out of that child on the day that they were visiting, they wouldn't see enough out of the interaction there?

Q: Well aside from some of the issues that many people I know would have with inspectors, but it's not just that, it's also about sort of understanding the effort that's made of an entire school community to make change in students, not just teachers and not just teachers in particular subject areas. And just recognising that a positive and excellent educational experience is not limited to those very tiny things. I just ... if an inspector was in there every day I don't know, over a period of a year, I don't know. But I can't imagine a scheme that could actually credit those situations, and I see them every day.

ANDREW: Yeah, yeah.

Q: Can I just add, because I've gone through the scheme very early in my career with inspectors, they see a very, very small part of teaching. They don't know the context of what's happening, they don't actually have time for you to explain what is actually going on and what part ... what your teaching is. They certainly don't know the dynamics of the classrooms and what's gone on with each of the students. So I think that it would be a disaster to use such a scheme and introduce inspectors, and I'm sure that a lot of the teaching profession will be horrified with that suggestion.

ANDREW: It's a challenge, this is not a world in which there is ever going to be some kind of a perfect merit pays scheme out there.

JOHN: Thank you. Can we take a question here and then I'm afraid that's going to need to be the last one so that we can finish up on time, thank you.

Q: Thanks Andrew. The first part of my question I guess is about if behind this entire move is about retention and attraction to teaching, which I'm guessing that is one of its motivations, what are your perspectives then on how education is actually valued within different cultures, because I think that has a lot to do with it. And in many respects when you look at other corporate studies of pay and why people leave and move between jobs, how much people are paid often has very little to do with why you quit jobs and take other jobs. It might be about that over the water cooler but it turns out that a lot of studies show that money has very little to do with all sorts of other jobs around why we leave and why we go and take others. And so interested in your perspective on that. And then the second part is what's your understanding then of other structures that have been tried within education, so for example using a band scheme rather than a scale, and when you create things like a leading teacher which we also have within education here in Victoria. So ... and that seems to, from what I can see, creating leading teachers actually tends to not only create a sense of value and status and recognition, but that teacher becomes a mentor and can help other teachers, and there is a certain level of pay associated with it, but that's not the reason why teachers want to be leading teachers. They want to be leading teachers because they want to help other teachers and they want to help more students.

ANDREW: I think the point about the status is there, but I'm not sure how much more governments can do to recognise great teachers, I'm not sure that ... we for example have teacher awards schemes and invariably the stories of the teachers who are recognised are extraordinary. One of the things I managed to do at ANU was to get a scheme through where the graduating class nominated the high school teacher that they thought had made the greatest difference in their lives and then that high school teacher was brought onto stage at the ANU graduation and recognised. So those schemes are I think helpful but I'm not sure that



they're going to bring about the sort of tectonic status change that you'd like to see. There's a pretty good mapping I think between occupational status and pay in not only in Australia, but in most developed countries, and you see that pay matters right, so I have this paper where I look at the test scores of kids entering teacher education classes and map that against the average salary of new teachers. When you get a good union pay deal, the starting salary of new teachers go up, you see the TERs in teaching go up as well. Kids are responding fast to the salary incentives, it's not all pay but on the margin salaries do seem to make a difference. The leading teacher models I think are interesting. The challenge is how sustainable they'll be. I guess Andrew's point before, that a scheme which doesn't endure won't have much impact on selection into the profession, and you see this sort of ... my favourite example of whiteanting is a scheme in New South Wales for exemplary teachers in which with in a few years 95% of teachers were rated as exemplary and it was then quickly in the next round just rolled back into base pay. So you need to work out some way in which these sort of schemes will endure.

JOHN: Thank you very much. Well I'm afraid ladies and gentlemen we will have to call it ... call an end to proceedings at that point. Andrew, can I thank you on behalf of all of us for a fascinating discussion, a terrific paper, a real attempt to bring some hard evidence to an area which is clearly vexed, it's clearly difficult. From the sound of things the evidence doesn't have all of the answers, but it's good to see that underneath the politician there is still an academic bursting with the academics right answer to every question which is here's some directions for some promising new research.

ANDREW: And can I also extend an invitation to the audience, if the magic of technology worked right then halfway through my talk this paper popped up on the website above, it's got a blog, you can leave comments there and I'd love to hear more comments on the paper itself.

JOHN: And as with all Grattan events ... or most Grattan events, this event has been recorded and will be available in both full blown listen to it as you go form and a transcript shortly. So thank you very much all for coming. Thank you, Andrew, very much for your contribution and we look forward to welcoming you back to another Grattan event within not too long a time. So thank you.

ANDREW: Thanks John.

AUDIO: This has been a podcast from Grattan Institute. Want to hear more? Check out our website, [www.grattan.edu.au](http://www.grattan.edu.au).

End of recording