

NBER WORKING PAPER SERIES

INCENTIVES IN MARKETS, FIRMS AND GOVERNMENTS

Daron Acemoglu
Michael Kremer
Atif Mian

Working Paper 9802
<http://www.nber.org/papers/w9802>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2003

We thank Robert Gibbons, Mathias Dewatripont, Bengt Holmstrom, Canice Prendergast, Raghuram Rajan, Jean Tirole, Jan Zabochnik and seminar participants at the NBER economics organization conference for useful comments and Jean Lee, Janina Matuszeski, and Courtney Umberger for research assistance. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

©2003 by Daron Acemoglu, Michael Kremer, and Atif Mian. All rights reserved. Short sections of text not to exceed two paragraphs, may be quoted without explicit permission provided that full credit including © notice, is given to the source.

Incentives in Markets, Firms and Governments
Daron Acemoglu, Michael Kremer, and Atif Mian
NBER Working Paper No. 9802
June 2003
JEL No.D23, L22, H10, H52

ABSTRACT

Most government expenditure is on goods that yield primarily private benefits, such as education, pensions, and healthcare. We argue that markets are most advantageous in areas where high-powered incentives are desirable, but in areas where high-powered incentives stimulate unproductive signalling effort, firms, or even government, may have a comparative advantage. Firms may be able to weaken incentives and improve efficiency by obscuring information about individual workers' contribution to output, and thus reducing their willingness to signal through a moral-hazard-in-teams reasoning. However, firms themselves may be unable to commit to not providing greater compensation to employees who distort their efforts to improve observed performance. Government organizations, on the other hand, often have to flatter wage schedules, thereby naturally weakening the power of incentives. We suggest that there are also endogenous reasons for why governments, even when they are run by self-interested politicians, may be able to commit to lower powered incentives than firms, because government operation makes yardstick comparisons, which increase the power of incentives, more difficult.

Daron Acemoglu
Department of Economics
MIT, E52-380B
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
daron@mit.edu

Michael Kremer
Department of Economics
Harvard University
Littauer Center 207
Cambridge, MA 02138
and NBER
mkremer@fas.harvard.edu

Atif Mian
Graduate School of Business
University of Chicago
atif@gsb.uchicago.edu

1 Introduction

While a range of transactions take place in markets, many important activities are organized within firms, and are partly shielded from market incentives. Still others are either directly or indirectly controlled by governments, with even more limited role for market incentives. The theory of the firm, starting with the seminal contribution of Coase [1937], has sought to identify why firms exist, what shapes their boundaries, and which activities are likely to be organized within firms. This theory does not directly address what determines the boundaries between firms and governments, however.

The theory of public finance, on the other hand, prescribes government intervention to supply public goods or more generally, to correct market failures. Whatever its normative merits, this theory falls short as an empirical description of government activity on two grounds. First, while some government expenditure is on public goods like interstate highways and scientific research, most public expenditure in developed countries is on goods that yield primarily private benefits, such as education, pensions, and health-care. For example, in the United States, more than half of the non-interest, non-military federal budget is spent by the Education Department, Social Security, and Health and Human Services.¹ Second, governments do not simply subsidize education, savings, and health, but actually *operate* schools, pension systems, and hospitals. Since in most cases the government can deal with market failures with Pigovian taxes and subsidies, and because existing evidence shows widespread inefficiencies in government provision (e.g., Barberis, et al. [1996], La Porta et al. [1999]), the theory of public finance does not provide an answer to the question of why governments are directly involved in the production of many services. An alternative view links the involvement of government in the economy to the empire-building tendencies of bureaucrats or rent-seeking behavior of politicians (e.g., Niskanen [1971], Bates [1981], Shleifer and Vishny [1994]). This view is consistent with widespread involvement by the government in the economy despite its inefficiency, but does not provide a theory for why governments operate schools, pensions, and health services rather than operating factories, for example.

In this paper, we develop the argument that high-powered incentives provided by markets may have costs as well as benefits, and that the relatively low-powered incentives

¹See U.S. Office of Management and Budget, historical tables.

in government operation may improve the allocation of resources (or at least carry lower costs) in sectors where high-powered incentives are most distortionary. In this view, the lack of incentives, often blamed as the source of government inefficiency, also emerges as the factor creating a comparative advantage for governments in certain areas. Furthermore, we use the same reasoning to show that firms may also be more useful as an alternative form of organization with stronger incentives than governments and weaker incentives than markets. It is useful to note from the outset that we do not mean to argue that ours is a complete theory of the division of economic activities between markets, firms and governments—many other factors influence the boundaries of firms and markets (see, among others, Williamson [1985], Grossman and Hart [1986], and Hart and Moore [1990]), and governments are clearly engaged in certain economic activities for public goods reasons or for rent-seeking reasons. Nevertheless, we hope that our approach will be complementary to existing theories of the firm, and will help in understanding in which activities government intervention may be less costly, even if it is primarily motivated by rent-seeking.²

The example of teaching, where government involvement is ubiquitous, can be used to illustrate our argument. Markets, and the associated price system, create high-powered incentives. In the teaching example, this corresponds to individual teachers selling their services to parents, on the basis of their past performance (e.g., test scores of their former pupils). This type of market organization would provide high-powered incentives to teachers, who would have a strong incentives to improve their pupils' performance. But it would also encourage unproductive effort to inflate test scores such as teaching to the test, cream-skimming of high-quality students, or even teacher cheating. More generally, high-powered incentives would be costly because producers could shift their effort to unproductive activities to manipulate the market's perception of their ability and the quality of their products.

Firms could be useful in such cases, when high-powered incentives have more costs than benefits. Firms, here interpreted as teams of producers, may be able to weaken individual incentives by endogenously creating a moral-hazard-in-teams problem. In the teaching example, when individual teachers are organized in schools, parents have worse information about each teacher's contribution to students' learning. From the viewpoint

²In fact, in our model, government operation is controlled by a self-interested politician.

of a teacher, this implies that any effort that she exerts will lead the market to perceive not only her, but also all the other teachers in the school, as having greater ability. This moral-hazard-in-teams problem weakens a teacher's incentives, and may improve the allocation of resources when incentives provided by the market are too high-powered. Naturally, however, as the power of incentives declines so does productive effort. Firms are therefore costly because they also reduce the amount of productive effort relative to markets, and will only arise in equilibrium when the benefits of low-powered incentives outweigh the costs in terms of lower effort.

Firms may sometimes be unable to achieve low-powered incentives, however, because even though each employee may have low-powered incentives, the firm as an entity (or its owner) is the residual claimant of the firm's profits, and thus naturally has high-powered incentives. As a result, even when each teacher has weak incentives, the school will face a commitment problem vis-a-vis the market because it has reasons to manipulate individual teachers' compensation to strengthen their incentives. Consequently, it is often impossible to prevent the high-powered incentives of the firms themselves from *trickling down* to their employees. For example, a firm which advertises that its salespeople are not paid on a commission basis may nonetheless award raises or promotions to its top .

This is where governments come in. The standard view of government organizations is one in which individuals have low-powered incentives (e.g., Crozier [1967], Tirole [1986], Banerjee [1997]), and this is often viewed as a major source of inefficiency in government activities (e.g., Niskanen [1971], Hanushek [1996]). Our perspective suggests that the low-powered incentives might also be the reason why certain activities should be (and perhaps are) operated by governments. There are a number of theories for why incentives in government are low-powered. Yet, to assume that those controlling the government do not have profit-maximizing interests or do not care about the market's perception of their success would be too extreme—if parents care about teacher performance when schools are private, they will also do so when they are public, and reward politicians in charge of schools accordingly. Nevertheless, we show that even in this case, there is a reason for incentives in government organizations to be weaker: the absence of yardstick competition. In the teaching example, when a school (or individual teacher) improves its performance, it will be favorably compared to others that have not

done so.³ This comparison (yardstick) effect is missing for the government, weakening the incentives at the top of the government, even for self-interested politicians in charge of government activity. This lack of comparison creates a trickle-down of low-powered incentives in the whole government organization.⁴

Our analysis therefore suggests that industries for which high-powered incentives are desirable and there is little scope for unproductive signaling effort should operate as markets. An important determinant of whether there is a large scope for signaling is the degree of uncertainty regarding success in past performance and the quality of the services provided. In light of this, occupations where productive effort (or selection of the appropriate candidates) is important and where performance is easy to observe, such as in the entertainment and sports industries or in the production of relatively simple goods, are natural candidates for market organization. Activities where high-powered incentives could lead to detrimental signaling and where group reputation could dull incentives should be organized as firms. Examples may include most durable goods manufacturing, consulting services, or journalism. Because “Mom-and-pop” operations in these fields may have too much incentive to falsely advertise or exaggerate their past performance for quality, the lower-powered incentives prevalent in large corporations might be preferable. Perhaps for this reason, durable goods retailers and many financial service firms advertise that their employees do not work on commission.⁵

At the other extreme, government operation may be appropriate for tasks where it is difficult for customers to accurately separate true quality from efforts to signal quality, and where firms cannot commit to low-powered incentives to build a reputation against low-quality work.⁶ Nevertheless, government ownership may also be quite costly because

³On relative performance evaluation and yardstick competition, see, among others, Green and Stokey [1983] and Shleifer [1985].

⁴Note that the argument here is one for government control rather than direct government operation. Under certain circumstances, regulation by the government may achieve similar results, though there are natural reasons, related to incompleteness of contracts and ex post bargaining/renegotiation, for why government regulation of privately-owned firms may not work very well. On the other hand, if issues of incomplete contracts are unimportant and government regulation can produce *exactly* the same allocation as government operation, we might want to think of effective government regulation as equivalent to direct government operation.

⁵Large corporations may also be able to build up a reputation for quality and for not encouraging their employees to mislead customers because of their high-powered incentives.

⁶We abstract from reputation in our model, but it is worth noting that it may be particularly difficult to build reputation in areas where it takes a long time to determine quality, in part because incentives to deviate from the high-reputation strategy would be strong. For example, a pension fund that takes

of the usual inefficiencies and lack of incentives associated with government operations. This suggests that government operation/ownership should only arise in activities where overall effort is relatively unimportant, and the tendency of the private sector to signal quality is significantly greater than the ability of politicians to signal their competence by using government performance in these areas. The latter condition is more likely to be met when there are limited opportunities for politicians to signal their competencies, or equivalently, when their actions or abilities have only limited effects on outcomes (other than through the incentives they provide to teachers, for example). Possible activities where government operation might be preferable to markets and firms, for the reasons outlined here, include teaching, pensions, healthcare, and law enforcement.

Table 1 summarizes the predictions of our model in a stark way:

	No commitment problems in firms/ signaling by politicians important	Commitment problems in firms/ limited signaling by politicians
High-powered incentives beneficial/ product quality observable	Markets	Markets
High-powered incentives costly/ product quality not observable	Firms	Governments

At this point it is worth considering a few areas where governments play major roles in the provision of private goods. A large share of primary and secondary education provision is by the state in almost all countries, and in many countries this provision is highly centralized. (The United States, with its local school boards, is an exception). Even if one accepts the case for subsidizing education, it is unclear why governments operate schools, rather than simply subsidizing them. Consistent with the model, incentives are weaker in governments, and there is evidence that high-powered incentives in markets create major distortions. Harbison and Hanushek [1992] and Hanushek [1996] document that public school teachers have much flatter pay schedules than their private counterparts, and argue that this leads to low effort. However, an emerging body of work big risks may have high returns in the short-run and very bad outcomes only with low frequency.

shows that high-powered incentives in schools create adverse incentives and significant efforts by teachers to artificially inflate their pupils performance. For example, Glewwe, et al. [2003] find evidence that existing teacher incentives in Kenya are indeed weak, with teachers absent 20% of the time. They report on a randomized evaluation of a program that provided primary school teachers in rural Kenya with incentives based on students' test scores. During the time the program was in place, students in program schools had higher test scores. An examination of the channels through which this effect took place, however, provides little evidence of more teacher effort aimed at increasing long-run learning. Teacher attendance did not improve, homework assignment did not increase, and pedagogy did not change. There is, however, evidence that teachers increased effort to raise short-run test scores by conducting more test preparation sessions. While students in treatment schools scored higher than their counterparts in comparison schools during the life of the program, they did not retain these gains after the end of the program, consistent with the hypothesis that teachers focused on manipulating short-run scores.

Similar results are obtained in U. S. studies. Jacob [2002] investigates the effects of the No Child Left Behind education bill in Chicago Public Schools, which provided stronger incentives to teachers. He shows that this program led to a significant increase in math and reading achievement scores, but that these increases were influenced by teaching of test-specific skills, and that there were no comparable gains on state-administered exams. In a related study, Jacob and Levitt [2002] find substantial increases in teacher cheating in response to the introduction of high-powered incentives in Chicago. Similarly, Figlio and Winicki [2002] look at the link between nutrition and short-term cognitive functioning, and find that school districts in Virginia increase the number of calories in school lunches on days when high-stakes tests are administered, thus artificially inflating test scores.

Evidence from the three countries which have moved farthest in introducing markets into education, Chile, New Zealand, and the United Kingdom, is also consistent with the notion that moving to a more market-oriented system leads to high-powered incentives and carries significant costs. Hsieh and Urqiola [2002] argue that competition among private schools in Chile's voucher program in part takes the form of trying to recruit strong students who will raise average scores and making cosmetic changes to school

appearance. Ladd and Fiske [2000] find similar effects in New Zealand. Although Glennerster [2002] has a positive overall assessment of recent British efforts to establish a quasi-market in education and publish league tables of comparative school performance, he notes that test score gains on U.K. exams were not matched by comparable gains on international exams. This is consistent with the possibility that schools may have focused on preparing students for the exams used to prepare the league tables, rather than on broader measures of learning.

Similar issues arise in the administration of pensions. Pension systems are often run by governments, though they provide private goods. Diamond and Valdes-Prieto [1994] argue that in systems like the Chilean one, run by private firms, administrative costs are substantially higher than well-managed government-run systems.⁷ The bulk of the additional administrative costs comes from “advertising,” whereby individual funds try to raise their performance appearance, without any apparent direct benefits, which has a parallel to exerting bad effort in our model. This is also consistent with the evidence in Chevalier and Ellison [1999] that U.S. mutual fund managers have significant career concerns and consequently manipulate the composition of their investments in ways that may not be in the best interest of mutual fund investors.⁸

Finally, our mechanism also suggests possible reasons for why healthcare and law enforcement may be government provided. With private provision, healthcare providers may compete to improve their reputation by taking actions that make people feel better in the short-run but do not improve their long-run health. For example, U.S. hospitals provide more non-medical amenities than British hospitals, which face less competition. While it is certainly possible that British hospitals may be providing sub-optimal non-medical amenities, the evidence is also consistent with the notion that in the more market-based U.S. system, hospitals are trying to signal quality by providing easily observed non-medical amenities. This would suggest that the ratio of spending on these amenities to spending on medical care is too high in the United States, and perhaps

⁷For example, in Malaysia, where the government runs and manages the pension system, the Employees’ Provident Fund costs U.S.\$10 a year per active affiliate to administer or 0.32 percent of annual covered earnings. In Chile, the administrative costs average U.S.\$51.6 a year or 1.70 percent of annual covered earnings.

⁸The long time periods involved in pensions and the presence of many unsophisticated investors make pensions more prone to signaling and quality-boosting advertising and increases the potential costs of high-powered incentives for pensions relative to many other types of financial intermediation.

explains why the U.K. manages to achieve health outcomes nearly as good as the United States, while spending only 7.3% of GDP on health compared to the 13% the United States spends.⁹ Finally, law enforcement agents with too high-powered incentives may frame innocent people to appear more able to solve crimes, so regulation of incentives might again be necessary.¹⁰

Although a range of other factors undoubtedly affect the division of production activities between markets, firms and governments, we believe that the role of different organizations in regulating the power of incentives is an important, and largely ignored, element. In emphasizing these issues, we focus on firms as teams of producers, thus departing from the most popular theory of the firm, which identifies firms with ownership of assets (e.g., Williamson [1985], Grossman and Hart [1986], and Hart and Moore [1990]), and analyzes how asset ownership shapes (investment) incentives. Our approach is complementary as we emphasize how incentives differ in firms, markets, and governments, but we do not define firms with reference to asset ownership.¹¹ Instead here firms are defined by team production, which ensures that markets learn relatively little about the performance of each member. Another, perhaps more important, difference between this literature and our paper is the impact of firms on incentives. In current literature, the major problem is low-powered incentives because of *ex post* rent-sharing, and firms are formed in order to improve incentives and encourage investments (though see Rajan and Zingales [1998] on how high-powered incentives associated with ownership can lead to too much specialization). In contrast, the problem in our setting is “too high-powered” incentives, and firms are useful as they flatten these incentives. Naturally, in practice both roles are likely to be present.

Our paper is also closely related to the career concerns literature (e.g., Holmstrom [1999], Stein [1989], Meyer and Vickers [1997], Dewatripont, et al. [1999]), and to the multi-tasking literature (e.g., Holmstrom and Milgrom [1991] and [1994]), which also emphasizes the costs of high-powered incentives, though does not develop a theory of

⁹See OECD Health Data 2002, Table 10.

¹⁰In the case of law enforcement, another concern is the potential cost of delegating part of the legitimate use of coercion to non-state agents.

¹¹It is possible to extend our model so as to define firm boundaries by asset ownership. The interaction between incentives provided by asset ownership and the effect of firm size on the composition of effort might be an interesting area for future research.

organizational form from these ideas.¹² The role of firms as institutions for suppression of information has been discussed by other authors, in particular Gibbons [1998] and Gibbons and Murphy [1992], but not in a context where suppression of information is useful for weakening incentives and improving the composition of effort. Moreover, this work assumes that firms have no commitment problem, and therefore provides no role for the government. Acemoglu [1998] suggests an argument for separation of ownership and control in firms so that agents making financing decisions do not have too high-powered incentives, which might be detrimental in the presence of asymmetric information. Kremer [1997] argues that worker cooperatives typically have egalitarian wage schedules because the median voter in a cooperative will choose a wage schedule which redistributes from high- to low-ability workers. Redistribution in cooperatives will dull incentives for *ex post* effort, which is typically undesirable, but may be useful in sectors where workers would otherwise waste effort on unproductive signaling. A recent paper by Levin and Tadelis [2002] is also related—they emphasize the benefits of firms in manipulating incentives because of joint production, though their story is non-informational and static.

A number of other papers have sought to understand why governments run and control certain activities. Hart, et al. [1997] use the incomplete contracts approach to explain why governments run prisons, and attempt to define what the proper scope of governments should be. With private ownership, managers receive a greater share of the gains they create, but this also induces them to engage in too much cost-cutting at the expense of quality. We share with this paper the emphasis on the potential costs of high-powered incentives associated with private ownership, but in our setup, these incentives arise not because of bargaining between the government and managers, but from the career concerns of producers, and different ownership structures affect incentives by influencing information transmission and the degree of career concerns.¹³

The rest of the paper is organized as follows. Section 2 describes the environment and

¹²The literature on advertising with imperfect information about quality is also related in this context, though the focus is on the costs of advertising to reveal quality by high-quality suppliers (e.g., see Kihlstrom and Riordan [1984] for Milgrom and Roberts [1986]).

¹³Bowles and Gintis [1976], Lott [1999], Kremer and Sarychev [1997], and Pritchett [2002] suggest that governments may run schools in order to control what ideology is taught to students. See also Acemoglu and Verdier [2000] and Prendergast [2003] on how government intervention or bureaucratic decision-making may create inefficiencies even when they are potentially improving the allocation of resources.

characterizes optimal incentives in the context of a simple mechanism design problem. Sections 3, 4, and 5 then compare the incentive structure under markets, firms, and governments. Section 6 concludes with a discussion of the main findings.

2 Model

We now outline the basic model. For concreteness, we will focus on the teaching example, but the theoretical argument is more general.

2.1 The environment

Consider an infinite horizon economy with n infinitely lived teachers, and $n' > n$ parents in each period who would like to send their children to teachers for education. Each parent has one child, who can be matched with one teacher in every period. A teacher can only teach to one student at a time. Each teacher, i , is endowed with a teaching ability a_t^i at the beginning of period t . The exact level of a_t^i is unknown, but both teacher i and parents share the same belief about the distribution of a_t^i . The common belief about teacher i 's ability is:

$$a_t^i \sim N(m_t^i, v_t).$$

Ability evolves over time according to the stochastic process given by:

$$a_{t+1}^i = a_t^i + \varepsilon_t^i, \tag{1}$$

where ε_t^i is i.i.d. with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. The disturbance term ε could result from personal shocks, or it could reflect the ability of the teacher to adapt to changing education demands and technology.

We consider a multi-tasking environment where a teacher can exert two types of effort, “good” and “bad”, denoted by g_t^i and b_t^i respectively. The titles “good” and “bad” reflect the social value of these efforts. The total output of a teacher is calculated through the education or human capital, h_t^i , that she provides to her student. h_t^i depends on the ability, as well as the good effort of the teacher according to the equation:

$$h_t^i = a_t^i + f(g_t^i), \tag{2}$$

where $f(g)$ is increasing and strictly concave in g , with $f(0) = 0$, and $h_t^i = 0$ if the child is not taught by a teacher.

Parents only care about the level of human capital provided to their children. The expected utility of a parent at time t is given by:

$$U_t^P = E_t[h_t] - w_t,$$

where $E_t[\cdot]$ denotes expectations with respect to publicly available information at the beginning of time t and $\delta < 1$ is the discount rate, and w is the wage paid to the teacher.

There is perfect competition among parents for education, so each parent pays a teacher a wage w_t^i equal to the expected level of education provided to her child, i.e.:

$$w_t^i = E_t[h_t^i].^{14}$$

The expected utility of a teacher i at time t is given by the time separable utility function:

$$U_t^i = E_t \left[\sum_{\tau=0}^{\infty} \delta^\tau (w_{t+\tau}^i - g_{t+\tau}^i - b_{t+\tau}^i) \right],$$

where $w_{t+\tau}^i$ denotes the wage of the teacher at time $t + \tau$.

The main assumption is that the true level of h_t^i provided by a teacher is not observable to parents. Instead, parents have to rely on an imperfect signal of h , given by the test score, s :

$$s_t^i = h_t^i + \gamma f(b_t^i) + \theta_t^i + \eta_t \tag{3}$$

where $\gamma \geq 0$, θ_t^i is an i.i.d. student-level shock distributed as $N(0, \sigma_\theta^2)$, for example, the ability of the students to learn or a summary measure of some other influences on the student's human capital. η_t is a common shock that *every* teacher receives in period t . For example, if all students are given the same test, η_t can be thought of as the overall difficulty of the test, or any other cohort-specific difference in ability or pre-schooling human capital. η_t is distributed i.i.d. and $N(0, \sigma_\eta^2)$. The variance σ_θ^2 measures the quality of signal s_t^i , while the variance of the common shock, σ_η^2 , also affects the informativeness of the signal. The lower these variances, the more precise the signal is in measuring the human capital contribution of teachers. Notice that the signal of

¹⁴The information contained in past student performance is already incorporated in $E_t[h_t]$, and we are assuming here that current and future student performances are not directly contractible.

human capital is imperfect in two ways. First, shocks θ and η make the test score a noisy signal for the student's human capital. Second, the signal can be inflated if the teacher exerts bad effort. The parameter γ measures the extent to which the signal can be manipulated by bad effort. It also captures the importance of output quality and composition of effort relative to the amount of effort (i.e., as γ declines, the importance of ensuring a high level of total effort increases).

The reason for calling the two types of efforts good and bad should be apparent now. Parents, as well as the society, only care about the good effort exerted by the teacher, since that determines the level of human capital h that the children receive. However, parents only observe the signal s , which can be manipulated by bad , as shown by equation (3). In real life, good effort can be thought of as real or honest teaching, where the teacher tries to ensure that the children understand conceptually what has been taught to them. Bad effort, on the other hand, may correspond to what is commonly referred to as “teaching to the test”. It involves rote learning, where the teacher just forces the students to cram certain essential facts or methods, without explaining the concepts behind them or the connection between the various facts and phenomena (see Hanaway [1992]). Such cramming is useless (or at least less useful than good effort) in terms of the human capital of the students, but it serves to inflate their test scores. Bad effort might also be interpreted as teacher cheating, which improves test scores, but clearly has no beneficial effect on pupils' human capital. In the context of the pension funds example as well as some other applications, we can think of bad effort as any activity that improves observed performance without affecting actual performance equally (e.g., advertising).

The timing of events in this world is as follows. In the beginning of every period t , parents form priors, m_t^i , on the abilities of teachers based on the historical test scores of the teachers. They then offer a wage w_t^i based on the expected ability of the teacher working with their child. The teacher then decides on the levels of good and bad effort, and h and s are realized at the end of period t . Ability a_t^i is then updated according to the stochastic process (1). The process then repeats itself in period $t + 1$.

We will limit the analysis to the rational expectations equilibrium path where all teachers choose $\{g_{t+\tau}^i, b_{t+\tau}^i\}_{\tau=0,1,\dots}$ optimally given their rewards, and the beliefs about teacher ability are given by Bayesian updating, given equilibrium behavior. We will

also focus on the long-run of the model so that the variance of each teacher's ability is constant, i.e. $v_t = v_{t+1} = v$. Finally, in the text, we focus on the case where n is very large, i.e., $n \rightarrow \infty$ (the Appendix gives equations that apply for $n < \infty$ as well). The $n \rightarrow \infty$ assumption allows us to ignore the common shock, η_t , since in a large population, the average ability of the entire population of teachers is known for sure. Thus the common shock can be backed out from the average of all test score signals in the population. The common shock will play an important role in Section 5 when we discuss the incentives with government-provided education.

2.2 Updating beliefs

We first determine how beliefs about teacher i 's ability will be updated at the end of period t based on new information revealed during period t . Parents' belief about teacher i at the beginning of period t can be summarized as, $a_t^i \sim N(m_t^i, v_t)$. Let $S_t = [s_t^1 \dots s_t^n]^T$ denote the vector of n test scores that the agents observe during period t .

Along the rational expectations equilibrium path, parents correctly infer effort levels g_t^i and b_t^i chosen by the teachers—or specifically by teacher i in this case. This means that parents can back out the part of S_t which only reflects the ability levels of the teachers, plus the noise. Let $Z_t = [z_t^1 \dots z_t^n]^T$ denote this backed out signal, where

$$\begin{aligned} z_t^i &= s_t^i - f(g_t^i) - \gamma f(b_t^i) \\ &= a_t^i + \theta_t^i + \eta_t \end{aligned}$$

Let a_{t+1}^i be the updated prior on teacher i 's ability conditional on observing Z_t . Then the normality of the error terms and the additive structure in equation (3) imply that $a_{t+1}^i \sim N(m_{t+1}^i, v_{t+1})$ where m_{t+1}^i and v_{t+1} denote the mean and the variance of the posterior distribution.

Using the normal updating formula for variance v_{t+1} , setting $v_{t+1} = v_t = v$ and focusing on the limit $n \rightarrow \infty$, we obtain the law of motion of the posterior of teacher i 's mean ability, m_{t+1}^i , as:

$$m_{t+1}^i = m_t^i + \beta(z_t^i - m_t^i) - \bar{\beta}(\bar{z}_t^{-i} - \bar{m}_t^{-i}), \quad (4)$$

where

$$\beta = \bar{\beta} = \frac{1 + \sqrt{1 + 4 \left(\frac{\sigma_\theta^2}{\sigma_\varepsilon^2} \right)}}{1 + 2 \left(\frac{\sigma_\theta^2}{\sigma_\varepsilon^2} \right) + \sqrt{1 + 4 \left(\frac{\sigma_\theta^2}{\sigma_\varepsilon^2} \right)}}, \quad (5)$$

z_t^i is the i th element of the vector Z_t , and refers to the signal from teacher i , while \bar{z}_t^{-i} is the average test score excluding teacher i . Since $n \rightarrow \infty$, we have $(\bar{z}_t^{-i} - \bar{m}_t^{-i}) \rightarrow \eta_t$, so the common shock is completely revealed and can be filtered out. The proof of (4) and (5) is given in the Appendix, where we also provide the expressions for the case of n finite.

The equation (4) illustrates the relative performance evaluation (yardstick competition) in the presence of η_t . The coefficient $\bar{\beta}$ captures relative performance evaluation. It emphasizes that an improvement in the score of a teacher creates a negative effect on the market's assessment of other teachers.

The following lemma summarizes some important properties of the updating formula for future reference:

Lemma 1 Parents update their beliefs about teacher i 's ability level, according to equations (4) and (5), where $1 > \beta > 0$. β is increasing in σ_ε^2 , and decreasing in σ_θ^2 .

The intuition behind the last part of Lemma 1 is that any increase in the variance of θ , σ_θ^2 , increases the noise in the signal, and makes it less valuable, and hence reduces β . An increase in σ_ε^2 makes the signal more valuable due to a greater change in ability since last period. In other words, a greater σ_θ^2 relative to σ_ε^2 implies that a given variation in test scores is less likely to come from teacher ability, so parents put less weight on differences in test scores in updating their posterior about teacher ability.

2.3 Efficient Allocations

We define social welfare at time t , U_t^W , as the sum of the teachers' and parents' utilities. Since the ability of teacher i enters additively in a teacher's utility function, all teachers should choose the same effort level in a given period. Social welfare can then be written as:

$$U_t^W = \sum_{\tau=0}^{\infty} \delta^\tau (\bar{A} + f(g_{t+\tau}) - g_{t+\tau} - b_{t+\tau}) \quad (6)$$

where \bar{A} is the average ability of teachers in the population, and $g_{t+\tau}$ and $b_{t+\tau}$ are the good and bad effort levels chosen by all teachers. Since in the case $n \rightarrow \infty$, each individual teacher's ability is a random walk, the average ability level of the population does not vary over time.

First Best: Maximizing (6) gives us the first-best. In the first-best, there is no bad effort, $b_t = 0$, and the level of good effort, g^{FB} , is given by $f'(g^{FB}) = 1$.

Second-Best: The first-best is useful as a benchmark. Nevertheless, since teacher effort and the level of human capital are not directly observable, a more useful benchmark is solving for the optimal mechanism given these informational constraints.

Let $\Omega_t^i = [m_0^i \ s_0^i \ s_1^i \ s_2^i \ \dots \ s_{t-1}^i]$ be the information set containing the vector of test scores for teacher i at the beginning of period t . Ω_t^i summarizes *all* contractible information about teacher i up to t . Let $w_t^i(\Omega_t^i)$ be the wage paid to teacher i in period t . Then the constrained maximization problem to determine the second-best allocation can be written as:

$$\begin{aligned} & \max_{\{w_{t+\tau}^i(\Omega_{t+\tau}^i)\}_{\tau=0,1,\dots}} U_t^W \text{ subject to} & (7) \\ & \{g_{t+\tau}, b_{t+\tau}\}_{\tau=0,1,\dots} \in \arg \max_{\{g'_{t+\tau}, b'_{t+\tau}\}_{\tau=0,1,\dots}} E_t \left[\sum_{\tau=0}^{\infty} \delta^\tau (w_{t+\tau}^i(\Omega_{t+\tau}^i) - g'_{t+\tau} - b'_{t+\tau}) \right]. \end{aligned}$$

Solving (7) above gives us the following proposition, which is proved in the Appendix:

Proposition 1 The second-best solution is given by $g_{t+k} = g^{SB}$, and $b_{t+k} = b^{SB}$ for all k , with $g^{SB} < g^{FB}$. The optimal wage schedule is given by an affine transformation of the function $w_t^i = \alpha^{SB} m_t^i + \kappa$ for

$$\alpha^{SB} = \frac{1 - \delta(1 - \beta)}{\delta\beta f'(g^{SB})}, \quad (8)$$

and for any nonnegative κ . Both g^{SB} and α^{SB} are monotonically decreasing in γ , and we have $\alpha^{SB} < 1$, for $\gamma > \underline{\gamma}$.

Proposition 1 highlights the trade-off that the social planner faces given the informational constraints. The planner needs to provide incentives to teachers in order to induce effort. However, high-powered incentives lead to both good and bad effort. This association between good and bad effort increases the shadow cost of increasing good effort,

leading to a lower level of good effort in the second-best relative to the first-best. The parameter γ captures the cost of higher incentives in the form of bad effort. Hence, an increase in γ increases the scope for bad effort and reduces the optimal level of incentives for the teacher, α^{SB} , and the second-best level of good effort, g^{SB} .

Equivalently, Proposition 1 can be understood in terms of two different types of negative externalities created by bad effort. The first type of negative externality is driven from the fact that since the effort levels are directly unobservable, the market's expectation of any individual teacher's effort level (which they back out in equilibrium) is based on the expectation of the market as a whole: when a teacher (or a positive mass of teachers) is expected to exert bad effort, a *given* test score translates into a lower perception of ability for other teachers. More explicitly, we have $m_{t+1}^i = m_t^i + \beta(z_t^i - m_t^i) - \bar{\beta}(\bar{z}_t^{-i} - \bar{m}_t^{-i})$ and $z_t^i = s_t^i - f(g_t) - \gamma f(b_t)$ where b_t is the level of bad effort that teachers are expected to exert. Greater b_t reduces z_t^i , thus the perceived ability of other teachers. This negative externality is at the root of the inefficiency of various organizations. The second type of externality is driven by the presence of relative performance evaluation to back out η_t . Such evaluation creates a more direct negative externality from the actual level of bad effort by a teacher (as opposed to the expectation of bad effort): as a teacher exerts more bad effort, she increases parents' posterior about the common shock, η_t , and reduces their posterior about other teachers' abilities. However, as long as we assume that $n \rightarrow \infty$, this second type of externality is driven down to zero (since the common shock is completely revealed).¹⁵

It is also noteworthy that with the wage schedule $w_t = \alpha m_t + \kappa$, teachers are driven entirely by career concerns. There are no effort contingent contracts. Teachers are paid a wage every period depending on their perceived ability. This formulation of incentives is similar to the seminal career concerns paper by Holmstrom [1999]. The extra effort put in by the teacher in period t increases her test score in period t . There are no immediate rewards for this increase as the teacher has already been paid her wage. However, an increase in the test score at t increases the perceived ability of the teacher in period $t + 1$ due to the updating rule (4). Moreover, because of the recursive nature of (4), the increase in perceived ability in $t + 1$ has a (progressively

¹⁵This argument also shows that when $n < \infty$ and n declines, externality becomes stronger and the gap between the the second best and the first best widens.

dampened) ripple effect on all future expected abilities. Hence the present discounted value of the marginal benefit of higher test scores in period t can be summarized as: $\alpha\delta\beta[1+\delta(1-\beta)+\delta^2(1-\beta)^2+\dots] = \frac{\alpha\delta\beta}{1-\delta(1-\beta)}$. Notice that the marginal benefit of a higher test score is increasing in β , which is the coefficient on an individual teacher’s test score in the ability updating rule. We thus define β as the “career-concerns coefficient”. The marginal benefit is also increasing in α , which can be thought of as “the market-reward coefficient”—how much the market rewards a unit increase in the perceived ability of the teacher. Expression (8) can be understood as ensuring that the market-reward coefficient is at the right level to ensure an effort level of g^{SB} given the career-concerns coefficient implied by Bayesian updating, (5).¹⁶

The following corollary emphasizes that second-best effort can be achieved, alternatively, by manipulating β (if this were possible).

Corollary The second-best equilibrium can alternatively be described by fixing α , and setting the career-concerns coefficient on an individual teacher’s test score equal to

$$\beta_{SB} = \frac{1 - \delta}{\delta(\alpha f'(g^{SB}) - 1)}. \quad (9)$$

This discussion highlights two different channels via which the second-best allocation can be obtained. The first is by manipulating α , i.e., how the market rewards “success,” and the second is by manipulating β , i.e., the teachers’ career concerns. In the sections that follow, we discuss how successful different organizational forms are in manipulating the career-concerns coefficient to improve the allocation of resources.

3 Incentives in markets

The second-best solution highlights the trade-off in providing greater incentives to teachers. While high-powered incentives are good for increasing good effort, they also lead to an increase in bad effort, which, as emphasized above, creates a negative externality and is costly. In this and the next two sections, we consider three different organizational

¹⁶This discussion raises the possibility of beneficial government regulation directly manipulating α , for example, by tax policy. We do not consider this possibility since differential taxation of income from different occupations is rare in practice, and potentially very costly for a variety of reasons, including the distortions of such tax policies on the allocation of resources and talent across occupations.

structures—markets, firms, and governments—and compare the incentives they provide to teachers.

Consider first the textbook model of perfectly competitive markets. Every teacher works independently, and sells her teaching services in the market each period. There is perfect competition among parents for education, and as a result each teacher gets paid her full expected output, which is given by the human capital equation (2). Wage w_t^i is given by

$$w_t^i = m_t^i + E_t[f(g_t^i)]. \quad (10)$$

The market equilibrium is therefore identical to the second-best equilibrium, except that now α is fixed to be 1. This leads to the following result.

Proposition 2 The market equilibrium is characterized by good effort level g^M , where

$$f'(g^M) = \frac{1 - \delta(1 - \beta)}{\delta\beta}.$$

We have that $g^M < g^{SB}$ if $\gamma < \underline{\gamma}$, and $g^M > g^{SB}$ if $\gamma > \underline{\gamma}$.

The proof follows from Proposition 1. The result that $g^M < g^{SB}$ if $\gamma < \underline{\gamma}$ is similar to the result in Holmstrom [1999] that, with discounting, career concerns are typically insufficient to induce the optimal level of effort. So in this case, even markets do not provide strong enough incentives. There may be certain non-market institutions (e.g. tournaments) that may strengthen incentives even further, though we do not focus on those here.¹⁷ Therefore, when $\gamma < \underline{\gamma}$, markets are the preferred form of organization. This leads to the conclusion, mentioned above, that where quality concerns are unimportant relative to the total amount of effort/investments, services should be sold in markets.

The case where $\gamma > \underline{\gamma}$, on the other hand, leads to the opposite conclusion. Now the natural career concerns provided by the market equilibrium create too high-powered incentives relative to the second-best. The extent to which the market provides excessively

¹⁷One can also imagine organizations that reward teachers according to a wage function along the lines of $w_t^i = \alpha m_t^i + \kappa$ with $\alpha > 1$ to strengthen incentives beyond those provided by the market. Firms, modeled below as teams of teachers, are unable to do so, however, since the “balanced budget” requirement imposes that $\alpha \leq 1$ and $\kappa \geq 0$. See Holmstrom [1979].

high-powered incentives depends on the career-concerns coefficient, β , and via this, on σ_θ^2 and σ_ε^2 . When σ_θ^2 is small relative to σ_ε^2 , β is high, and teachers in the market care a lot about their pupils scores, giving them very high-powered incentives.

In the case with $\gamma > \underline{\gamma}$, since markets are encouraging too much bad effort, firms or governments may be useful by modifying the organization of production to dull incentives. We next turn to a discussion of the role of firms and governments in providing appropriate incentives when markets lead to too-high-powered incentives, i.e., when $\gamma > \underline{\gamma}$.

4 Incentives in firms

The previous section showed that the presence of competitive forces can lead to powerful incentives, even when it is socially optimal to have lower-powered incentives. We now consider how firms can attempt to overcome these problems by creating teams of teachers to weaken the signaling ability of individual teachers. In doing so, we provide a theory of the firm based upon the advantages that firms provide over pure markets in terms of providing duller incentives.

We model the firm as a partnership of K teachers working together.¹⁸ The teachers are now organized as a team that is engaged in joint production of teaching. As a result, a firm of K teachers can teach K children, and each child will acquire human capital equal to:

$$h_t^{jk} = a_t^{jk} + f(g_t^{jk}),$$

for some teacher k in firm j .

What makes teachers inside a firm different from teachers in a market environment is the fact that the parents only observe the aggregate or average test score of all the teachers (or pupils) in the firm. Therefore, an important function of firms in one economy is to shut down the individual signals (test scores) of teachers.

More specifically, consider an allocation where there are J firms in the economy with the j th firm made up of K^j teachers, so that $\sum_{j=1}^J K^j = n$. Moreover, we think of $K^j < \infty$ for all j , so that as $n \rightarrow \infty$, we also have $J \rightarrow \infty$. Denote the set of teachers constituting firm j by \mathbb{K}^j and index teachers within each firm by k . The market then

¹⁸The results are similar if we model the firm as a corporation with a single owner.

observes J signals every period, where each signal is the average test score of all the teachers in the firm. Let \bar{s}_t^j denote the signal from firm j at time t , then:

$$\bar{s}_t^j = \bar{a}_t^j + \overline{f(g_t)}^j + \gamma \overline{f(b_t)}^j + \bar{\theta}_t^j + \eta_t \quad (11)$$

where $\bar{a}_t^j = \frac{1}{K^j} \sum_{k \in \mathbb{K}^j} a_t^{jk}$, $\bar{\theta}_t^j = \frac{1}{K^j} \sum_{k \in \mathbb{K}^j} \theta_t^{jk}$, $\overline{f(g_t)}^j = \frac{1}{K^j} \sum_{k \in \mathbb{K}^j} f(g_t^k)$ and $\overline{f(b_t)}^j = \frac{1}{K^j} \sum_{k \in \mathbb{K}^j} f(b_t^k)$

As before, along the equilibrium path parents can back out the signal $\bar{z}_t^j = \bar{a}_t^j + \bar{\theta}_t^j + \eta_t$ from \bar{s}_t^j (again, with $J \rightarrow \infty$, η_t is backed out perfectly). Let

$$\bar{m}_t^j = \frac{1}{K^j} \sum_{k \in \mathbb{K}^j} m_t^{jk} \quad (12)$$

be the expected ability of the firm j at time t . Then parents update their belief about teacher k 's ability according to an updating formula similar to (4) above:

$$m_{t+1}^{jk} = m_t^{jk} + \beta_F (\bar{z}_t^j - \bar{m}_t^j) - \overline{\beta}_F (\bar{z}_t^{-j} - \bar{m}_t^{-j}). \quad (13)$$

Although parents can only observe the average test score of all the teachers in the firm, it is in theory possible for those inside the firm to have more information about each individual teacher's performance. We assume that, in addition to the average test score of all the teachers in the firm, insiders also observe the following signal of each teacher's performance (test score):

$$s_t^{jk} = a_t^{jk} + f(g_t^{jk}) + \gamma f(b_t^{jk}) + \theta_t^{jk} + \tilde{\theta}_t^{jk} + \eta_t, \quad (14)$$

where $\tilde{\theta}_t^{jk}$ is a normal error term, distributed as $N(0, \sigma_\theta^2)$. When $\sigma_\theta^2 \rightarrow \infty$, so that $\tilde{\theta}_t^{jk}$ has a very large variance, we obtain the case where insiders observe exactly the same information as outsiders — i.e. there is no “asymmetry of information”. We will start with this case of no asymmetry of information between insiders and outsiders, and then analyze the case where insiders have better information than the market.

Bertrand competition between parents again ensures that a group of teachers will be paid their expected contribution to human capital. Thus the *average* earnings of a teacher in firm j , \bar{w}_t^j , is:

$$\bar{w}_t^j = \bar{m}_t^j + \overline{f(g_t)}^j, \quad (15)$$

where \bar{m}_t^j . The average perception of the teachers in firm j is defined above in (12).

4.1 Symmetric information— $\sigma_{\theta}^2 \rightarrow \infty$

In this case, the firm can only make payments to teachers conditional on the past history of \bar{s}_t^j , the average signal from all the teachers (as well as the initial prior about the individual teacher). We assume that the objective of the firm is to maximize total net returns of all the employees in the firm. We also assume that the set of teachers in firm j , \mathbb{K}^j , is chosen at time t and is not changed thereafter. In other words, teachers do not switch “teams” after initial assignment.¹⁹ Then, the maximization problem is:

$$\max_{\mathbb{K}^j} \max_{\{g_{t+\tau}^{jk}, b_{t+\tau}^{jk}\}_{k \in \mathbb{K}^j}} E_t \left[\sum_{\tau=0}^{\infty} \sum_{k \in \mathbb{K}^j} \delta^{\tau} \left(\bar{w}_{t+\tau}^j(K^j) - g_{t+\tau}^{jk}(K^j) - b_{t+\tau}^{jk}(K^j) - m_t^{jk} \right) \right],$$

where $\bar{w}_{t+\tau}^j$ is given by (15) and K^j denotes the total number of teachers in that firm. \bar{w}_t^j , $g_{t+\tau}$, and $b_{t+\tau}$ are written as functions of K^j to emphasize that the size of the firm will influence incentives and payments. In addition, the maximization is over the choice of a set \mathbb{K}^j , and finally, the term $\sum_{k \in \mathbb{K}^j} m_t^{jk}$ is the sum of the outside options of the teachers and also acts as a convenient normalization. This normalization makes it clear that the maximization is identical to the simpler maximization problem over the size of firms, K^j . Or more precisely,

$$\max_{K^j} \max_{\{g_{t+\tau}^{jk}, b_{t+\tau}^{jk}\}_{k \in \mathbb{K}^j}} E \left[\sum_{\tau=0}^{\infty} \delta^{\tau} \left(\bar{w}_t^j(K^j) - g_{t+\tau}^j(K^j) - b_{t+\tau}^j(K^j) \right) \right], \quad (16)$$

The objective function of the firm is not enough to determine individual incentives, however. We also need to know how each individual teacher is rewarded. We assume

¹⁹It can be shown that, as long as we are in the case with $\gamma > \underline{\gamma}$, “no switching” is a long-run equilibrium. To see this briefly, consider a symmetric long-run equilibrium. Every teacher is paid her expected output equal to $w_t^{jk} = \bar{m}_t^{jk} + f(\bar{g}_t)^j$ every period. In the long-run equilibrium, g is constant over time and across all teachers. Hence, the future expected utility of a teacher if she stays in the firm is given by: $U_t^{jk} = E_t \left[\sum_{\tau=0}^{\infty} \delta^{\tau} \left(\bar{m}_{t+\tau}^{jk} + f(g) - g - b \right) \right]$. Moreover, since ability is a random walk,

we have $E_t[\bar{m}_{t+\tau}^{jk}] = \bar{m}_t^{jk}$, and hence $U_t^{jk} = \left(\frac{\bar{m}_{t+\tau}^{jk} + f(g) - g - b}{1 - \delta} \right)$. Now suppose that a teacher deviates and switches to another firm. Compute his/her utility assuming that in all future periods, he/she is expected to, and will, exert good and bad effort equal to g and b . In this case, we continue to have $E_t[\bar{m}_{t+\tau}^{jk}] = \bar{m}_t^{jk}$, so her utility is still given by U_t^{jk} above, and there is no gain in switching. In addition, after switching teams, market perceptions of abilities in the new and the old teams of the teacher who switched will be negatively correlated. This will induce the teacher who switches to put in more effort (both good and bad). As long as $\gamma > \underline{\gamma}$, this will be rewarded by the market less than the cost of effort, and hence greater effort will reduce the utility of the switching teacher. XXX

that each teacher is paid according to:

$$\tilde{w}_t^{jk} = \overline{m}_t^{jk} + \overline{f(g_t)^j} \quad (17)$$

where \overline{m}_t^{jk} is given by (13). This wage rule rewards each teacher according to her contribution to the firm's revenues. Such a wage rule will result from a variety of different micro-foundations, for example, when each teacher makes a take-it-or-leave-it offer to the "firm". This wage rule also parallels the market wage rule, (10), thus making it clear that the advantage of firms does not come from manipulating the wage rule, but from obscuring information.

What are the benefits of having large firms? To answer this question, let us return to the updating equation, (13), which is similar to the updating equation in the market case, (4). The career-concerns coefficient for an individual teacher is different, however. In particular, in a firm of size K , the individual career-concerns coefficient is $\frac{\beta_F}{K}$. The reason for this decline is the "moral-hazard-in-teams" problem. For each incremental increase in her test score, a teacher only gets rewarded for a fraction $\frac{1}{K}$ of the value created for the team. Moreover, as the proof to Proposition 3 in the Appendix will show, $\beta_F = \beta$, and $\overline{\beta}_F = \overline{\beta}$. Since $\frac{\beta}{K}$ is decreasing in K , the power of incentives can be reduced by increasing firm size, and in the case where $\gamma > \underline{\gamma}$, there exists a K^* such that $\frac{\beta}{K^*} = \beta_{SB}$ where β_{SB} is the career-concerns coefficient that would ensure the second-best with $\alpha = 1$, as defined by equation (9). Moreover, given that the maximization problem in (16) is identical to the social surplus maximization problem (6), the firm will select $K = K^*$, and the second-best outcome is attained. This gives us the following result:

Proposition 3 Suppose that $\sigma_\theta^2 \rightarrow \infty$. Then for a firm of size K , the good effort level chosen by a teacher, g , is given by $g^F(K)$, where g^F is monotonically decreasing in K with $g^F(1) = g^M$ and $g^F(K) \rightarrow 0$ as $K \rightarrow \infty$.

When $\gamma > \underline{\gamma}$, there exists a unique equilibrium where firms have size equal to $K^* = \beta/\beta_{SB} > 1$ and where teachers exert the second-best level of good effort, g^{SB} .

The proof is given in the Appendix. Here we outline the intuition behind the result. As in the market equilibrium, a teacher is still paid her expected output. However, the marginal effect of test score at time t on future expected ability is lower in firms than in markets. In other words, firms lower the career-concerns coefficient from β to $\frac{\beta}{K}$, thus weakening individual incentives.

The reduction of career concerns effects under firms can thus completely resolve the “over-incentivization” problem. With firms of appropriate size, the second-best allocation of Section 2.3 is achieved. When firms compete to maximize their value, all firms will endogenously expand to the optimal size K^* , since by definition K^* gives the maximum of (16). As a result, the size of the firms here arises as a means to dull incentives.

4.2 Asymmetric information and the commitment problem—

$$\sigma_{\theta}^2 < \infty$$

The idea that the outside market only observes the average and not individual test scores of teachers is meant to capture the idea that the organizational structure of firms can be used to mask or suppress information. The question still remains, however, as to what extent the firm as a whole (or the principal/the owner) has access to information regarding an individual teacher’s test score. Proposition 3 above implicitly assumed that nobody in the firm is able to observe individual teachers’ test scores either. Alternatively, the allocation in Proposition 3 can be achieved if firms can announce a wage contract of the form (17), i.e., one that does not make any use of non-publicly available information, for all of their employees, and make a strict commitment to (not renegotiating) this wage contract.

We now relax the assumption of symmetric information by assuming that $\sigma_{\theta}^2 < \infty$. This implies that insiders now observe a noisy signal of individual teacher performance as well as the public signal coming from average firm performance. In addition, we also assume that firms cannot commit to not modifying the rewards of their employees if this is in their interests. This is plausible given the various ways in which firms can enter into side deals with their employees. Notice that without the asymmetry of information, firms had no ability to manipulate incentives by modifying employee rewards, so there was no need for firms to commit to wage contracts, hence no commitment problem. The commitment problem is introduced by the asymmetry of information.

With the asymmetry of information of this sort, and the resulting commitment problem, firms become less attractive because, given the *ex post* manipulation of incentives inside the team, there will be limits to how much they can reduce the power of the incentives. More specifically:

Proposition 4 Suppose $\sigma_\theta^2 < \infty$. There exists $\overline{\sigma_\theta^2}$, such that when $\sigma_\theta^2 > \overline{\sigma_\theta^2}$ and $\gamma > \underline{\gamma}$, there is a unique equilibrium in which firms have size equal to $K^{**}(\sigma_\theta^2) > 1$, where $K^{**}(\sigma_\theta^2)$ induces the second-best level of effort g^{SB} , and is decreasing in σ_θ^2 . When $\sigma_\theta^2 \leq \overline{\sigma_\theta^2}$, the second-best outcome cannot be achieved. When $\sigma_\theta^2 = 0$, the firm equilibrium leads to the market outcome, i.e., the good effort level g^M .

(See Appendix for the proof.) As internal signals become more precise, a profit-maximizing firm will always use that extra information to encourage teachers to exert further bad effort, thus improving outside perception of the average ability of its employees, and via this channel, its future revenues. This implies that choosing a firm size of K^* (as given by Proposition 3) is no longer a credible commitment to low-powered incentives and to the second-best level of good effort. Instead, the exact extent of the strength of incentives will be determined by the amount of information the firm has about each employee's performance. Since the firm's (the insiders') information about individual performance is also imperfect, i.e., typically, $\sigma_\theta^2 > 0$, average performance of the firm is still informative about each employee's ability. Therefore, firm size, by affecting how informative average performance is about individual ability, still influences how powerful each employee's incentives are. Generally, the larger the size of the firm, the less information there is about an individual's performance inside the firm, and the less powerful are equilibrium incentives. Therefore, a firm might still be able to credibly commit to low-powered incentives by further increasing its size to $K^{**}(\sigma_\theta^2)$, thus reducing teachers' incentives even after taking into account the *ex post* manipulation of these incentives. Nevertheless, the precision of internal signals puts a lower bound on how much the firm can dull incentives through "team production". In particular, if $\sigma_\theta^2 \leq \overline{\sigma_\theta^2}$ for some critical threshold $\overline{\sigma_\theta^2}$, then there is sufficiently good internal information about teacher performance that even a very large firm would not be able to dull incentives sufficiently. Therefore, in this case the asymmetry of information, and the resulting commitment problem, breaks the "firm equilibrium" of Proposition 3, which achieved the second-best.

The intuition for why the asymmetry of information and the associated commitment problem make firms less useful can be alternatively described as follows: when production is organized within firms, individual teachers have relatively weak incentives because of

the moral-hazard-in-teams problem. The firm as an entity, or its owner, however, has strong incentives, since it is the residual claimant of the profits. The problem is whether these high-powered incentives of the firm will trickle down to employees. With symmetric information, the firm has no way of increasing the incentives of its employees, thus there is no trickle-down of its high-powered incentives. The asymmetry of information introduces the possibility that the firm can manipulate its employees' incentives, and this, combined with the inability to commit to observable contracts, makes its high-powered incentives trickle down to the teachers. As a result, the benefits of firms in terms of dulling incentives are reduced or disappear.

5 Incentives in governments

The analysis above highlighted the potential role of firms in improving efficiency through their ability to suppress information. However, it also pointed out the limitations of firms to credibly commit to such a course of action in the presence of the informational asymmetries between the firm and the outside world. Let us now imagine a world with $\gamma > \underline{\gamma}$, so that markets provide too high-powered incentives, and with σ_θ^2 small so that firms cannot commit to dulling individual incentives because of severe asymmetries of information between insiders and outsiders. In particular, let us assume $\sigma_\theta^2 = 0$, which implies that firms will be unable to solve the over-incentivization problem, and g^M (the market equilibrium) is also the firm equilibrium. Can government operation reduce incentives below those implied by firms and markets?

The answer may be yes, since governments are often thought to provide relatively lower-powered incentives than markets and firms. There are a variety of reasons for low-powered incentives in governments, ranging from the absence of market discipline (Niskanen [1971], Hanushek [1996]) to an optimal design to avoid collusion and corruption (e.g., Crozier [1967], Tirole [1986], Banerjee [1997]). In our model, we can think of these concerns imposing a wage structure on government organizations of the form $w_t^i = \alpha^G m_t^i + \kappa$. If α^G were close to α^{SB} , i.e., incentives in government-run firms were close to the power of incentives necessary to achieve the second-best, government organization would be useful.²⁰

²⁰The analysis in this section simply points out when government operation may be preferred over markets and firms. We do not offer a political economy model determining when government operation

In addition to factors outside our model leading to low-powered incentives in government, there are also endogenous reasons in our model enabling governments to manipulate the power of incentives differently than profit-maximizing organizations. We now discuss these issues using a highly stylized model of government organization whereby the government (a politician) decides the size of schools and individual rewards. This politician is potentially self-interested. Moreover, similar to individual teachers in our analysis so far, she would like to convince the market (the voters) that she has high ability.

Denote the politician’s true ability and market’s perception of it at time t by a_t^p and m_t^p respectively. Further suppose that the politician has an objective function

$$U_t^{pol} = E_t \left[\sum_{\tau=0}^{\infty} \delta^\tau (m_{t+\tau}^p - C_{t+\tau}) \right], \quad (18)$$

where $C_{t+\tau}$ is the cost per student of the schooling system, or $C_{t+\tau} = w_{t+\tau}^{ij}$. This utility function implies that the politician always likes to convince parents (or the voters) that he has high ability, and faces a cost in terms of the expenditures on the education budget in this process. Suppose also that the ability of the politician evolves according to

$$a_{t+1}^p = a_t^p + \varepsilon_t^p,$$

where ε_t^p is i.i.d. with $\varepsilon \sim N(0, \sigma_p^2)$ just as in equation (1).

We will discuss incentives provided by government ownership/operation under two different scenarios: (1) “no politician effects”, so that the ability and actions of the politician do not matter for student performance (other than through the incentives provided to teachers); (2) “politician effects,” so that the ability of the politician influences the human capital of students, for example, because the politician takes actions that affect teacher incentives through other means, or influences the curriculum, etc.

5.1 Government operation with “no politician effects”

We start with the case where the politician’s ability or action does not matter for student performance—except through the politician’s manipulation of teachers’ incentives. This immediately implies that no action that the politician can take will send a positive signal

will actually arise in equilibrium.

about her ability. This removes the commitment problem faced by firms. Therefore, the politician can choose, and commit to, the allocation of Proposition 3. In particular, she can set up $J = n/K^*$ schools of size K^* each, and promise a payment of $\overline{w}_t^j = \overline{m}_t^j + \overline{f(g_t)^j}$ to each teacher, replicating the allocation of Proposition 3, which, as shown earlier, coincides with the second-best. Alternatively, the politician can simply set the wage of each teacher equal to $w_t^i = \alpha^{SB} m_t^i + \kappa$.

This is true even when the politician has superior information about teacher test scores relative to the market, i.e., observes a signal s_t^{jk} like in (14) with $\sigma_\theta^2 < \infty$. Although she can reward individual teachers on the basis of this additional signal, s_t^{jk} , and encourage them to exert more good and bad effort, the resulting increase in test scores will not lead to a better voter belief about her ability, since a higher average test score will not constitute a positive signal about ability, a_t^p . As a result, the politician has no reason to manipulate teacher incentives. This enables government organizations to commit to low-powered incentives.

5.2 Government operation with “politician effects”

The above analysis may be criticized because there is no room for the actions of the politician to influence outcomes other than through her effect on teacher effort. This means that the politician has no incentive to “inflate” student performance, and could easily provide, and commit to, the second-best incentives for teachers. In general, decisions taken by education ministers or prime ministers can have important influences on aggregate outcomes, for example, through teacher selection, by affecting incentives in other dimensions or by influencing the curriculum. We allow for this possibility in a simple way by assuming that the ability of the politician also matters for the human capital attained by the children. In particular, assume that the human capital of a student taught by teacher k in school j is

$$h_t^{jk} = (1 - \lambda) a_t^{jk} + \lambda a_t^p + \overline{f(g_t)^j}, \quad (19)$$

where a_t^p is the ability of the politician in charge of the schooling system. This formulation implies that the politician’s ability influences the human capital of all the children in the school system, because of some other dimension of incentives that the politician provides to teachers, or because of his decisions. Consequently, the politician has an

incentive to inflate test scores in order to improve others' perception of his own ability.

The point that we want to make is that even in this case, the government may have a comparative advantage in providing low-powered incentives. When an individual school inflates its own test scores, this has a negative effect on other schools because of the relative performance evaluation used by the market to remove the effect of the common shock, η_t . This intensifies the negative externality, and encourages private schools to give high-powered incentives to their teachers. In contrast, with government operation, the politician is in charge of the whole school system, so when citizens (voters) update their beliefs about the ability of the politician, the common shock is not filtered out, and acts as an additional source of noise, thus weakening the incentives of the politician.

More formally, parents (or voters) observe all test scores, and update their beliefs regarding the ability of the politician according to the equation

$$m_{t+1}^p = m_t^p + \beta^p (z_t - m_t^p)$$

where

$$z_t = \frac{1}{n'} \sum_{i=1}^{n'} \bar{s}_t^i - (1 - \lambda) \bar{A} - \overline{f(g_t)} - \gamma \overline{f(b_t)} = \lambda a_t^p + \eta_t$$

and

$$\beta^p = \frac{\lambda \left(1 + \sqrt{1 + 4 \left(\frac{\sigma_\eta^2}{\lambda^4 \sigma_p^2} \right)} \right)}{\lambda^2 \left(1 + \sqrt{1 + 4 \left(\frac{\sigma_\eta^2}{\lambda^4 \sigma_p^2} \right)} \right) + 2 \left(\frac{\sigma_\eta^2}{\lambda^4 \sigma_p^2} \right)}. \quad (20)$$

where n' is the number of firms in the economy, \bar{A} is the average ability of teachers in the population, and \bar{s}_t^i refers to the average test score of firm i at time t . These updating equations have an intuition similar to (4) and (5). The updating is now about the ability of the politician. For updating, only the average test score in the population is relevant, and along the rational expectations equilibrium path, this average test score is equal to $\lambda a_t^p + (1 - \lambda) \bar{A} + \overline{f(g_t)} + \gamma \overline{f(b_t)} + \eta_t$. The career-concerns coefficient of the politician, β^p , is different from that of firms (or individual teachers), β , because learning now is about the ability of the politician, which may have a different distribution, and more importantly, because noise comes from the aggregate shock, η_t , not from the student performance shocks, the θ_t 's. Here the absence of relative performance evaluation (yardstick competition) with government operation is important. The reason why σ_η^2 did

not feature in the updating equations (4) and (5) is that relative performance evaluation eliminated this aggregate shock. With government operation, relative performance evaluation is not possible, since everything is run by the government,²¹ and this makes (the perception of) government performance dependent on the realization of the aggregate shock. As a result, the politician receives credit for only part of the improvements in test scores, weakening his incentives, and therefore, indirectly those of the whole government organization. The greater σ_η^2 , i.e., the more important the aggregate shock, the smaller β^p , and the weaker the incentives in governments. In the limit, as $\sigma_\eta^2 \rightarrow \infty$, the politician has completely flat incentives.

We have determined the incentives that a politician running government schools faces from the voters, but the real outcome that we are interested in is the level of effort chosen by the teachers under government ownership, g^G . This effort level will be determined by the incentives trickling down to the individual teacher level. Given the politician's own incentives in (20), we can determine the wage schedule that the politician will offer to each of his teachers. In particular, assume that the politician offers each teacher a linear wage function of the form,

$$w_t^{ij} = \alpha_{t+\tau}^p m_{t+\tau}^{ij} + \kappa,$$

where κ is some constant. First, consider the case where the level of incentives provided to teachers $\alpha_{t+\tau}^p$ is observable. Then, even though the politician can manipulate teacher incentives, he will receive no benefit from this, since voters will effectively observe the level of good and bad effort exerted by teachers. In this case, the results would be identical to that with no politician effects, and the politician would simply choose $\alpha_{t+\tau}^p = \alpha^{SB}$ and achieve the second-best.

However, parallel to our treatment of firms where teacher incentives inside the firm are not observed by outsiders, it may be more reasonable to presume that $\alpha_{t+\tau}^p$ s are not observable citizens. Interestingly, even in this case, government operation can provide weaker incentives than markets and firms. We now discuss this case.

²¹This argument needs to be qualified when local politicians run local school districts, for example, as in the U.S. In this case, there will be some amount of competition even with government operation. Nevertheless, given the importance of district-specific shocks, the extent of yardstick competition might be much less than the case of private ownership, with competition between private schools, thus qualitatively leading to the same type of comparison as that emphasized in this section.

The question which arises is, given her own incentives determined by β^p , what level of α^p will the politician choose for the teachers? We can answer this question by maximizing the objective function of the politician given by (18). Since the government acts as a monopolist, it will only give each teacher their minimum reservation utility. Let \underline{u} be the spot reservation utility of a teacher putting in zero effort. Then the government must pay each teacher a wage equal to $(\underline{u} + g_t + b_t)$ each period where g_t and b_t are the effort levels that α^p induces. In other words, $w_t^{ij} = \alpha_t^p m_t^{ij} + \kappa = (\underline{u} + g_t^{ij} + b_t^{ij})$ each period. Plugging the wage function into (18) and maximizing with respect to $\alpha_{t+\tau}^p$ gives us the following result:

Proposition 5 Suppose that $\sigma_\theta^2 = 0$ and $\gamma > \underline{\gamma}$. Then both markets and firms lead to the same inefficiently high level of effort $g^M > g^{SB}$, with g^M given by Proposition 2. On the other hand, the level of effort under government operation is g^G , given by:

$$f'(g^G) = \frac{1 - \delta(1 - \beta^p)}{\delta\beta^p}.$$

We have $g^G < g^M$ if and only if $\beta^p < \beta$. β^p is decreasing in σ_η^2 and increasing in λ .

The proposition establishes that government operation often provides weaker incentives than firms and markets, even when politicians have an interest in inflating test scores, and the manipulation of teacher incentives by the politician is not observed by voters. The reason is the presence of common shocks (hence the importance of σ_η^2), which were filtered out in markets, but not under government operation. Common shocks are not filtered out, thereby increasing the amount of noise in the performance, and thus weakening the incentives of politicians. These weaker incentives then trickle down to the teachers.

More specifically, when $\beta^p < \beta$, government organization provides less high-powered incentives than markets and firms, because the politician has less to gain by inflating test scores. This is likely to be the case when aggregate shocks are large, i.e., when σ_η^2 is large, and when the contribution of the politician to aggregate test scores, λ , and the room for the politician to prove that he has high ability, σ_p^2 , are limited. This also suggests that government operation may be beneficial in reducing incentives in activities where there is more scope for unproductive signaling effort and politicians have limited room to manipulate aggregate performance to improve their standing. In contrast, when

σ_η^2 is small, and/or when λ and σ_p^2 are large, politicians can manipulate incentives more than profit-maximizing firms, and government ownership/operation is likely to lead to a deterioration in the allocation of resources.

6 Conclusion

This paper has provided a simple model to analyze incentives in markets, firms and governments when consumers only have access to noisy signals of output quality. Our analysis suggests that competitive markets work well when effort and investment are important relative to the ability of agents to manipulate quality (i.e., small γ in terms of Proposition 1). When output signals/quality can be manipulated, markets typically yield too high-powered incentives, and consequently agents distort the composition of their efforts.

Firms can help in such situations by suppressing some information, thus dulling incentives. Nevertheless, the ability of firms to reduce the power of incentives may be limited because firms themselves are profit-maximizing organizations and would like to signal to the market that the average performance of the firm in the future will be good. Therefore, firm owners (or firms themselves as entities) will have an incentive to encourage their employees to distort their efforts. In other words, the high-powered incentives of firms coming from profit maximization are likely to trickle down to their employees.

This is where governments may potentially lead to better outcomes due to their ability to provide relatively low-powered incentives to workers. This may arise naturally if governments provide low-powered incentives for other reasons (which is often argued to be the case, and suggested as a reason for the inefficiency of government operation in many activities). More interestingly, in the presence of common shocks, there is a natural reason for governments to have a comparative advantage in providing low-powered incentives—the presence of common shocks leads to yardstick competition among firms, whereby an improvement in one firm’s signal (or more generally, observed output) has a negative impact on the evaluation of the remaining firms. Profit-maximizing firms do not take this externality into account, contributing further to high-powered incentives. Government ownership internalizes this negative externality, reducing the power of in-

centives. Intuitively, a self-interested politician knows that she will obtain only limited credit for improvements in aggregate performance if there are large aggregate shocks. This consideration is absent when production is organized in firms, since what matters is, at least in part, comparison to other firms.

Our analysis also suggests a theory for why certain occupations should supply their services in markets, while others should be organized in firms, and yet others should be under government operation. There are limits to this theory, however, since in actual practice, many other factors are undoubtedly important, and the boundaries of markets, firms, and governments are not simply, or perhaps even mainly, determined as a way of regulating the power of incentives. For example, governments may run certain functions for rent-seeking reasons. Nevertheless, the arguments developed in this paper might suggest a reason for why government ownership in some activities make be less costly than in others, and thus help us understand which activities we are more likely to see government involvement. Overall, the importance of the forces emphasized here is therefore an empirical question. We believe that an empirical investigation of relative efficiency of markets, firms and governments in different activities, taking into account issues of relative output quality and composition of effort, would be a fruitful area for future research.

A major element missing from our analysis is political economy: the fact that politicians have non-profit maximizing incentives and respond to their political constituencies may have important bearing on the incentives provided in government organizations, and the types of actions to which government organizations can and cannot commit. Work incorporating the interaction between the power of incentives and political economy considerations may be another fruitful area for future research.

7 Appendix

Proof of Lemma 1: Although in the text we focus on the case where $n \rightarrow \infty$, here we solve for the general case with n finite first. Parents with prior m_t^i about teacher i observe the vector Z_t . Let v_t be the variance of m_t^i . Since m_t^i and Z_t are distributed normally, we can use the normal updating formula to update the ability prior and compute m_{t+1}^i conditional on Z_t . This formula is given by:

$$m_{t+1}^i = m_t^i + \Sigma_{12}\Sigma_{22}^{-1}(Z_t - M_t),$$

where $M_t = [m_t^1 \dots m_t^n]^T$, $\Sigma_{12} = [0 \ 0 \dots v_t \dots 0]$, with the convention that v_t corresponds to the i th component of the vector, and

$$\Sigma_{22} = \begin{bmatrix} (v_t + \sigma_\theta^2 + \sigma_\eta^2) & & \sigma_\eta^2 & & \sigma_\eta^2 \\ & \sigma_\eta^2 & & \dots & \sigma_\eta^2 \\ & & \sigma_\eta^2 & & \\ & & & \sigma_\eta^2 & \\ & & & & (v_t + \sigma_\theta^2 + \sigma_\eta^2) \end{bmatrix}.$$

Σ_{22} is an $(n \times n)$ matrix with $(v_t + \sigma_\theta^2 + \sigma_\eta^2)$ as the diagonal term, and σ_η^2 as all the non-diagonal terms. Σ_{22}^{-1} can be written as:

$$\Sigma_{22}^{-1} = \frac{1}{b} \begin{bmatrix} a & 1 & 1 \\ 1 & \dots a \dots & 1 \\ 1 & 1 & a \end{bmatrix}$$

where

$$b = (n-1)\sigma_\eta^2 - \frac{(v_t + \sigma_\theta^2 + \sigma_\eta^2)^2}{\sigma_\eta^2} - (n-2)(v_t + \sigma_\theta^2 + \sigma_\eta^2), \text{ and}$$

$$a = - \left[\frac{(v_t + \sigma_\theta^2 + \sigma_\eta^2)}{\sigma_\eta^2} + (n-2) \right].$$

Plugging in the value of Σ_{12} and Σ_{22}^{-1} , we obtain:

$$m_{t+1}^i = m_t^i + \beta(z_t^i - m_t^i) - \bar{\beta}(\bar{z}_t^{-i} - \bar{m}_t^{-i}), \tag{A1}$$

where:

$$\bar{z}_t^{-i} = \frac{1}{(n-1)} \sum_{j \neq i} z_t^j$$

$$\bar{m}_t^{-i} = \frac{1}{(n-1)} \sum_{j \neq i} m_t^j$$

$$\beta = \frac{v_t a}{b} \text{ and } \bar{\beta} = -\frac{v_t(n-1)}{b}.$$

Note that $1 > \beta \geq \bar{\beta} > 0$.

Next, we need to solve for v_t . Like m_t^i , v_t is also updated each period after the realization of Z_t . This updating formula is given by:

$$v_{t+1} = v_t - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} + \sigma_\varepsilon^2,$$

where Σ_{12} and Σ_{22} are defined above and $\Sigma_{21} = [\Sigma_{12}]^T$. Therefore, we have

$$v_{t+1} = v_t - \frac{a}{b} v_t^2 + \sigma_\varepsilon^2$$

Since we are interested in the long-run stationary equilibrium, we impose the condition $v_{t+1} = v_t = v$. This stationarity condition implies:

$$v^2 = \frac{b}{a} \sigma_\varepsilon^2$$

which can be expanded into:

$$v^3 + v^2[(n-1)\sigma_\eta^2 + (\sigma_\theta^2 - \sigma_\varepsilon^2)] - v[n\sigma_\eta^2\sigma_\varepsilon^2 + 2\sigma_\theta^2\sigma_\varepsilon^2] - [\sigma_\varepsilon^2(\sigma_\theta^2)^2 + n\sigma_\eta^2\sigma_\theta^2\sigma_\varepsilon^2] = 0 \quad (\text{A2})$$

We now look at the case where $n \rightarrow \infty$. Then (A2) simplifies to:

$$v^2 - v\sigma_\varepsilon^2 - \sigma_\theta^2\sigma_\varepsilon^2 = 0,$$

which can be explicitly solved as:

$$v = \frac{\sigma_\varepsilon^2 + \sqrt{\sigma_\varepsilon^4 + 4\sigma_\theta^2\sigma_\varepsilon^2}}{2}. \quad (\text{A3})$$

Plugging (A3) back into (A1), and taking the limit and $n \rightarrow \infty$, we obtain the updating equation (A1), $(\bar{z}_t^i - \bar{m}_t^i) \rightarrow \eta$, and $\bar{\beta} \rightarrow \beta$, thus $\bar{\beta}(\bar{z}_t^i - \bar{m}_t^i) \rightarrow \beta\eta$, and we have

$$m_{t+1}^i = m_t^i + \beta (z_t^i - m_t^i - \eta)$$

where

$$\beta = \frac{v}{v + \sigma_\theta^2} = \frac{1 + \sqrt{1 + 4\left(\frac{\sigma_\theta^2}{\sigma_\varepsilon^2}\right)}}{1 + 2\left(\frac{\sigma_\theta^2}{\sigma_\varepsilon^2}\right) + \sqrt{1 + 4\left(\frac{\sigma_\theta^2}{\sigma_\varepsilon^2}\right)}}. \quad (\text{A4})$$

It can be easily checked from (A4) that $0 < \beta < 1$, $\frac{\partial \beta}{\partial \sigma_\varepsilon^2} > 0$, and $\frac{\partial \beta}{\partial \sigma_\eta^2} < 0$. ■

Proof of Proposition 1: The second-best is given by the solution to (7). Notice that (7) is a point-wise maximization problem over time. The constraint can then be written separately for each τ as:

$$f'(g_{t+\tau}) \sum_{\tau'=0}^{\infty} \delta^{\tau'} \frac{\partial w_{t+\tau+\tau'}}{\partial s_{t+\tau}} = 1, \text{ and } \gamma f'(b_{t+\tau}) \sum_{\tau'=0}^{\infty} \delta^{\tau'} \frac{\partial w_{t+\tau+\tau'}}{\partial s_{t+\tau}} = 1 \quad (\text{A5})$$

The above conditions can be combined to give $f'(g_{t+\tau}) = \gamma f'(b_{t+\tau})$ for all τ which implies that:

$$b_{t+\tau} = f'^{-1} \left(\frac{f'(g_{t+\tau})}{\gamma} \right). \quad (\text{A6})$$

The inverse of $f'(x)$ exists due to the concavity of $f(x)$. Equation (A6) defines the *feasible* pairs of $(g_{t+\tau}, b_{t+\tau})$ even when the wage function is not differentiable. Restriction to differentiable wage functions therefore does not change our second-best solution. Given (A6), we can simplify our maximization problem (7) into the unconstrained problem:

$$\max_{g_{t+\tau}} \sum_{\tau=0}^{\infty} \delta^\tau \left(\bar{A} + f(g_{t+\tau}) - g_{t+\tau} - f'^{-1} \left(\frac{f'(g_{t+\tau})}{\gamma} \right) \right). \quad (\text{A7})$$

The above is a well-defined maximization problem with a unique global maximum. Moreover, because of the additive nature of (A7), at the optimum, $g_t^{SB} = g_{t'}^{SB} = g^{SB}$ for all t, t' . Differentiating (A7) with respect to g_t , we obtain:

$$f'(g^{SB}) = 1 + \frac{(1/\gamma)f''(g^{SB})}{f''(f'^{-1}(f'(g^{SB})/\gamma))}$$

Because of the concavity of $f(x)$, the expression on the RHS is greater than 1, implying $g^{SB} < g^{FB}$.

Given g^{SB} , it is easy to solve for the optimal wage structure. In fact there is a continuum of optimal wage structures, as long as they satisfy the condition (from (A5)), $f'(g^{SB}) \sum_{l=1}^{\infty} (\delta^l \frac{\partial w_{t+k+l}}{\partial s_{t+k}}) = 1$. This condition can be satisfied by the wage schedule $w_t^i = \alpha^{SB} m_t^i + \kappa$. To see this, note that from (4), we can write

$$m_{t+l} = (1 - \beta)^l m_t + \beta(1 - \beta)^{l-1} s_t + \beta(1 - \beta)^{l-2} s_{t+1} + \dots + \beta s_{t+l-1} + \text{constant}.$$

We can then write $\frac{\partial w_{t+k+1}}{\partial s_{t+k}} = \alpha\beta(1-\beta)^{l-1}$, which implies that at the second-best

$$\alpha^{SB} = \frac{1-\delta(1-\beta)}{f'(g^{SB})\beta\delta}.$$

Since g^{SB} is decreasing in γ , there exists a $\underline{\gamma}$ such that for $\gamma > \underline{\gamma}$, $\alpha^{SB} < 1$. ■

Proof of Proposition 3: With firms, parents observe J signals, represented by:

$$\bar{Z}_t = [\bar{z}_t^1 \ \bar{z}_t^2 \ \dots \ \bar{z}_t^J]^T,$$

where

$$\bar{z}_t^j = \bar{a}_t^j + \bar{\theta}_t^j + \eta_t.$$

There are J firms in the economy with each firm j having a size K_j . Then using the normal updating formula, ability for teacher k in firm j is updated using:

$$m_{t+1}^{jk} = m_t^{jk} + \bar{\Sigma}_{12} \bar{\Sigma}_{22}^{-1} (\bar{Z}_t - \bar{M}_t),$$

where $\bar{M}_t = [\bar{m}_t^1 \ \dots \ \bar{m}_t^J]^T$, $\bar{\Sigma}_{12} = [0 \ 0 \ \dots \ \frac{v_{jt}^F}{K_j} \ \dots \ 0]$, with $\frac{v_{jt}^F}{K_j}$ corresponding to the j th component of the vector, and

$$\bar{\Sigma}_{22} = \begin{bmatrix} \left(\frac{v_{1t}^F + \sigma_\theta^2}{K_1} + \sigma_\eta^2\right) & & \sigma_\eta^2 & & \sigma_\eta^2 \\ \sigma_\eta^2 & & \dots & & \sigma_\eta^2 \\ & & \sigma_\eta^2 & & \left(\frac{v_{Jt}^F + \sigma_\theta^2}{K_J} + \sigma_\eta^2\right) \end{bmatrix}.$$

$\bar{\Sigma}_{22}$ is a $(J \times J)$ matrix with $\left(\frac{v_{jt}^F + \sigma_\theta^2}{K_j} + \sigma_\eta^2\right)$ as the diagonal term, and σ_η^2 as all the non-diagonal terms. Now take the limit $n \rightarrow \infty$, which, since $K_j < \infty$ for all j , implies $J \rightarrow \infty$. Then $\bar{\Sigma}_{22}^{-1}$ can be written as:

$$\bar{\Sigma}_{22}^{-1} = \begin{bmatrix} \frac{K_1}{v_{1t}^F + \sigma_\theta^2 + K_1 \sigma_\eta^2} & 0 & 0 \\ 0 & \dots \frac{K_j}{v_{jt}^F + \sigma_\theta^2 + K_j \sigma_\eta^2} \dots & 0 \\ 0 & 0 & \frac{K_J}{v_{Jt}^F + \sigma_\theta^2 + K_J \sigma_\eta^2} \end{bmatrix}.$$

Plugging in the value of $\bar{\Sigma}_{12}$ and $\bar{\Sigma}_{22}^{-1}$, we get:

$$m_{t+1}^{jk} = m_t^{jk} + \beta_F (\bar{z}_t^j - \bar{m}_t^j) - \beta_F (\bar{z}_t^{-j} - \bar{m}_t^{-j})$$

where:

$$\begin{aligned}\bar{z}_t^{-j} &= \frac{1}{(J-1)} \sum_{l \neq j} z_t^l \\ \bar{m}_t^{-j} &= \frac{1}{(J-1)} \sum_{l \neq j} m_t^l. \\ \beta_F &= \frac{v^F}{v^F + \sigma_\theta^2}\end{aligned}$$

The variance is then updated by:

$$v_{j,t+1}^F = v_{jt}^F - \bar{\Sigma}_{12} \bar{\Sigma}_{22}^{-1} \bar{\Sigma}_{21} + \bar{\sigma}_\epsilon^2$$

where $\bar{\sigma}_\epsilon^2 = \frac{\sigma_\epsilon^2}{K_j}$. Assuming stationarity, i.e. $v_{j,t+1}^F = v_{jt}^F = v_j^F$, we get:

$$v_j^F \rightarrow v = \frac{\sigma_\epsilon^2 + \sqrt{\sigma_\epsilon^4 + 4\sigma_\theta^2\sigma_\epsilon^2}}{2} \equiv v^F.$$

Notice that the stationary firm variance is the *same* as the individual teacher stationary variance under the market equilibrium. Similarly, as $J \rightarrow \infty$, $\beta_F \rightarrow \frac{v^F}{v^F + \sigma_\theta^2} \equiv \beta$, $\bar{\beta}_F \rightarrow \beta$, and $(\bar{z}_t^{-j} - \bar{m}_t^{-j}) \rightarrow \eta_t$. In other words, as $J \rightarrow \infty$, the career concerns coefficient for the entire firm is exactly the same as the career concerns coefficient for an individual teacher under market equilibrium. However, the career-concerns coefficient for an *individual* k in firm j is given by $\frac{\beta}{K_j}$, and is decreasing in K_j .

From this, it is straightforward to see that $g^F(1) = g^M$ and $g^F(K_j) \rightarrow 0$ as $K_j \rightarrow \infty$. Moreover, $g^F(K_j)$ is monotonically decreasing in K_j . The firm will now endogenously set $K_j = K^*$ such that $g^F(K^*) = g^{SB}$. To see this, note that the firm partners try to maximize:

$$\max_{K_j} E_t \left[\sum_{\tau=0}^{\infty} \delta^\tau (\bar{m}_{t+\tau}^j + f(g_{t+\tau}(K_j)) - g_{t+\tau}(K_j) - b_{t+\tau}(K_j)) \right]. \quad (\text{A8})$$

As we saw in Proposition 1, (A8) is maximized at $K_j = K^*$, such that $g(K^*) = g^{SB}$, providing the second-best solution. ■

Proof of Proposition 4: Let $n \rightarrow \infty$ so that with each firm of finite size, $J \rightarrow \infty$. As before, this assumption implies that the common shocks can be perfectly filtered

out, so to simplify notation, we ignore the common shocks. Since there is asymmetric information now, we must distinguish between internal and public information. The internal information on an individual k in firm j can be summarized by:

$$z_t^{jk} = a_t^{jk} + \theta_t^{jk} + \tilde{\theta}_t^{jk},$$

while the public information is given by: $\bar{z}_t^j = \bar{a}_t + \bar{\theta}_t^j$. The firm has access to both internal and public information. Recall that each teacher also gets her full surplus, i.e., $w_t^{jk} = m_t^{jk} + f(g_t^{jk})$.

Given the internal and public signal, the updating formula used by the firm becomes:

$$m_{t+1}^{jk} = m_t^{jk} + \begin{bmatrix} v_t^F & \frac{v_t^F}{K} \end{bmatrix} \begin{bmatrix} (v_t^F + \sigma_\theta^2 + \sigma_\theta^2) & \frac{(v_t^F + \sigma_\theta^2)}{K} \\ \frac{(v_t^F + \sigma_\theta^2)}{K} & \frac{(v_t^F + \sigma_\theta^2)}{K} \end{bmatrix}^{-1} \begin{bmatrix} (z_t^{jk} - m_t^{jk}) \\ (\bar{z}_t^j - \bar{m}_t^j) \end{bmatrix},$$

which implies

$$m_{t+1}^{jk} = m_t^{jk} + \beta^{asy}(z_t^{jk} - m_t^{jk}) + \bar{\beta}^{asy}(\bar{z}_t^{j(-k)} - \bar{m}_t^{j(-k)}),$$

where

$$\beta^{asy} \equiv \left(\frac{v_t^F (K-1)(v_t^F + \sigma_\theta^2) + v_t^F \sigma_\theta^2}{(v_t^F + \sigma_\theta^2)((K-1)(v_t^F + \sigma_\theta^2) + K\sigma_\theta^2)} \right)$$

defines the career concerns coefficient with asymmetric information, superscript $-k$ refers to the average excluding the k th teacher, and

$$\bar{\beta}^{asy} \equiv \frac{v_t^F \sigma_\theta^2 K^2}{(v_t^F + \sigma_\theta^2)((K-1)(v_t^F + \sigma_\theta^2) + K\sigma_\theta^2)(K-1)}.$$

Once again the stationary variance v^F will be given by $v_{t+1}^F = v_t^F = v^F$, which after applying the normal updating formula is given by the implicit equation,

$$\frac{(v^F)^2(K-1)(v^F + \sigma_\theta^2) + (v^F)^2\sigma_\theta^2}{(v^F + \sigma_\theta^2)((K-1)(v^F + \sigma_\theta^2) + K\sigma_\theta^2)} = \frac{\sigma_\varepsilon^2}{K}.$$

To emphasize dependence on firm size, let us write the career concerns coefficient above, β^{asy} , as $\beta^{asy}(K)$. Let K^{**} be the value of K that makes $\beta^{asy}(K) = \beta_{SB}$. Then we have that $\frac{\partial K^{**}}{\partial \sigma_\theta^2} < 0$. In other words, as the firm learns more about an individual teacher, it becomes harder to sustain the second-best level of effort, and firm size needs to increase.

Since $\frac{\partial \beta^{asy}(K)}{\partial K} < 0$, and $\beta^{asy}(K = 1) = \beta_M$, all we need to show[? SHOW-LOOK COMBO IS WEIRD] for $\beta^{asy}(K) = \beta_{SB}$ in equilibrium is to look at the limit $\lim_{K \rightarrow \infty} \beta^{asy}(K)$ and show this is less than β_{SB} . As $K \rightarrow \infty$, we have:

$$\beta^{asy}(K) \rightarrow \frac{v^F}{v^F + \sigma_\theta^2 + \sigma_\theta^2},$$

and

$$\left(\frac{v^F}{v^F + \sigma_\theta^2 + \sigma_\theta^2} \right) < \beta_{SB} \Leftrightarrow \sigma_\theta^2 > \overline{\sigma_\theta^2} \equiv \left(\frac{v^F}{\beta_{SB}} - v^F - \sigma_\theta^2 \right).$$

Therefore, if $\sigma_\theta^2 > \overline{\sigma_\theta^2}$, the economy can achieve the second-best allocation. However, for $\sigma_\theta^2 \leq \overline{\sigma_\theta^2}$ (i.e., severe asymmetric information), the commitment problem implies that the second-best can never be achieved. ■

Proof of Proposition 5: The politician's problem can be defined as follows:

$$\max_{\alpha_{t+\tau}^p} U_t^{pol} = E_t \left[\sum_{\tau=0}^{\infty} \delta^\tau (m_{t+\tau}^p - w_{t+\tau}^{ij}) \right],$$

where

$$w_{t+\tau}^{ij} = \alpha_{t+\tau}^p m_{t+\tau}^{ij} + \kappa_{t+\tau} = (\underline{u} + g_{t+\tau}^{ij} + b_{t+\tau}^{ij})$$

$$m_{t+1}^p = m_t^p + \beta^p (z_t - m_t^p).$$

Keeping market expectations of $g_{t+\tau}^{ij}$ and $b_{t+\tau}^{ij}$ fixed, and maximizing over $\alpha_{t+\tau}^p$, we get the first order condition:

$$\left(f'(g_{t+\tau}^{ij}) \frac{\partial g_{t+\tau}^{ij}}{\partial \alpha^p} + \gamma f'(b_{t+\tau}^{ij}) \frac{\partial b_{t+\tau}^{ij}}{\partial \alpha^p} \right) \left(\frac{\delta \beta^p}{1 - \delta(1 - \beta^p)} \right) = \left(\frac{\partial g_{t+\tau}^{ij}}{\partial \alpha^p} + \frac{\partial b_{t+\tau}^{ij}}{\partial \alpha^p} \right).$$

We know that teachers' first-order conditions imply that $f'(g) = \gamma f'(b)$. This simplifies the above expression to:

$$f'(g^G) \left(\frac{\delta \beta^p}{1 - \delta(1 - \beta^p)} \right) = 1,$$

thus completing the proof. ■

Bibliography

Acemoglu, Daron. [1998] "Credit Market Imperfections and Separation of Ownership from Control," *Journal of Economic Theory*, February, 78(2), 355-381.

Acemoglu, Daron and Thierry Verdier. [2000] "The Choice Between Market Failures and Corruption," *American Economic Review*, 90(1), 194-211.

Banerjee, Abhijit. [1997] "A Theory of Misgovernance," *Quarterly Journal of Economics*, 112(4), 1289-1332.

Barberis, Nicholas; Maxim Boycko; Andrei Shleifer; and Natalia Tsukanova. [1996] "How does Privatization Work? Evidence from the Russian Shops," *Journal of Political Economy*, 104(4), 764-790.

Bates, Robert. [1981] *Markets and States in Tropical Africa*, Berkeley, CA: University of California Press.

Bowles, Samuel and Herbert Gintis. [1976] *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*, New York: Basic Books.

Chevalier, Judith and Glenn Ellison. [1999] "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, May, 114(2), 389-431.

Coase, Robert. [1937] "The Nature of the Firm," *Economica*, 4(16), 386-405.

Crozier, Maurice. [1967] *Bureaucratic Phenomenon*, University of Chicago Press.

Dewatripont, Mathias; Ian Jewitt; and Jean Tirole. [1999] "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, January, 66(1), 183-198.

Diamond, Peter and Salvador Valdes-Prieto. [1994] "Social Security Reforms" in Bosworth, Barry P., Rudiger Dornbusch and Raul Laban, *The Chilean Economy: Policy Lessons and Challenges*, Washington, D.C: The Brookings Institution, 257-328.

Figlio, David N. and Joshua Winicki. [2002] "Food for Thought: The Effects of School Accountability Plans on School Nutrition," National Bureau of Economics Research Working Paper: 9319.

Gibbons, Bob. [1998] "Incentives in Organizations," National Bureau of Economic Research Working Paper: 6695.

Gibbons, Bob and Murphy, Kevin. [1992] "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, June, 100(3), 468-505.

Glennerster, Howard. [2002] "United Kingdom Education 1997-2001," *Oxford Review of Economic Policy*, 18(2), 120-136.

Glewwe, Paul; Nauman Ilias; and Michael Kremer. [2003] "Teacher Incentives," National Bureau of Economic Research Working Paper: 9671.

Green, Jerry and Nancy Stokey. [1983] "A Comparison of Tournaments and Contests," *Journal of Political Economy*, 91(3), 349-64.

Grossman, S. and O. Hart. [1986] "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94(4), 691-719.

Harbison, Ralph W. and Eric A. Hanushek. [1992] *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*, New York: Oxford University Press.

Hanushek, Eric A. [1996] "Outcomes, Cost, and Incentives in Schools," in Hanushek, Eric A. and Dale W. Jorgenson, eds. *Improving America's Schools: The Role of Incen-*

tives, Washington, D.C.: National Academies Press.

Hart, O. and J. Moore. [1990] "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6), 1119-1158.

Hart, O.; A. Shleifer; and R. Vishny. [1997] "The Proper Scope of Government: Theory and an Application to Prisons," *Quarterly Journal of Economics*, November, 112(4), 1127-1161.

Holmstrom, B. [1979] "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74-91.

Holmstrom, B. [1999] "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, January, 66(1), 169-182.

Holmstrom, B., and Milgrom P. [1991] "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics and Organization*, 7 (Special Issue): 24-52.

Holmstrom, B. and P. Milgrom. [1991] "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics and Organization*, 7 (Special Issue), 24-52.

Holmstrom, B., and P. Milgrom. [1994] "The Firm as an Incentive Structure," *American Economic Review*, 84(4), 972-991.

Hsieh, Chang-Tsi and Miguel Urquiola. [2002] "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program," unpublished.

Jacob, B. [2002] "Accountability, Incentives, and Behavior: The Impact of High Stakes Testing in the Chicago Public Schools," National Bureau of Economic Research Working Paper: 8968.

- Jacob, B. and S. Levitt. [2002] "Teacher Cheating," University of Chicago, mimeo.
- Kihlstrom, R. and M. Riordan. [1984] "Advertising As a Signal," *Journal of Political Economy*, 92(3), 427-450.
- Kremer, Michael. [1997] "Why Are Worker Cooperatives So Rare?," National Bureau of Economic Research Working Paper: 6118.
- Kremer, Michael and Andrei Sarychev. [1998] "Why Do Governments Operate Schools," unpublished.
- Ladd, Helen and Edward Fiske. [2000] *When Schools Compete: A Cautionary Tale*, Washington, D.C.: The Brookings Institution.
- Levin, J. and S. Tadelis. [2002] "A Theory of Partnerships," Stanford University Working Paper.
- Lott, Jr., John. [1999] "Public Schooling, Indoctrination and Totalitarianism," *Journal of Political Economy*, 107(6), 127-157.
- Meyer, M. and J. Vickers. [1997] "Performance Comparisons and Dynamic Incentives," *Journal of Political Economy*, June, 105(3), 547-581.
- Milgrom P. and J. Roberts. [1986] "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 94(4), 796-821.
- Niskanen, William A. [1971] *Bureaucracy and Representative Government*, Chicago: Aldine/Atherton.
- Prendergast, Canice. [2003] "The Limits of Bureaucratic Efficiency," forthcoming *Journal of Political Economy*.

Pritchett, Lant. [2002] “When Will They Ever Learn?: Why All Governments Produce Schooling,” unpublished.

Rajan, R. and L. Zingales. [1998] “Power in a Theory of the Firm,” *Quarterly Journal of Economics*, May, 113(2), 387-432.

Shleifer, A. [1985] “A Theory of Yardstick Competition,” *Rand Journal of Economics*, 16(3), 319-327.

Shleifer, Andrei and Robert W. Vishny. [1994] “Politicians and Firms,” *Quarterly Journal of Economics*, November, 109(4), 995-1025.

Stein, J. [1989] “Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior,” *Quarterly Journal of Economics*, November, 104(4), 655-669.

Tirole, Jean. [1986]. “Hierarchies and Bureaucracies: On the Role of Collusion in Organizations”. *Journal of Law Economics and Organization*, 2, 181-214.

Williamson, O. [1985] *The Economic Institutions of Capitalism*, New York: Free Press.