NBER WORKING PAPER SERIES

MORAL HAZARD, INCENTIVE CONTRACTS AND RISK: EVIDENCE FROM PROCUREMENT

Gregory Lewis
Patrick Bajari

Working Paper 17647
http://www.nber.org/papers/w17647

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2011

Previously circulated as "Incentives and Adaptation: Evidence from Highway Procurement in Minnesota." We are grateful to the Minnesota Department of Transportation for data, and to Rabinder Bains, Tom Ravn and Gus Wagner for their help. We would also like to thank John Asker, Susan Athey, Raj Chetty, Matt Gentzkow, Oliver Hart, Ken Hendricks, Jon Levin, Justin Marion, Ariel Pakes, Chad Syverson and participants at Harvard, LSE, MIT, Toronto, UC Davis and Wisconsin and the AEA, CAPCP, IIOC, Stony Brook, UBC IO, WBEC and the NBER IO / Market Design / PE conferences. Lou Argentieri, Jason Kriss, Zhenyu Lai, Tina Marsh, Maryam Saeedi, Connan Snider and Danyang Su provided excellent research assistance. We gratefully acknowledge support from the NSF (grant no. SES-0924371). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Moral Hazard, Incentive Contracts and Risk: Evidence from Procurement
Gregory Lewis and Patrick Bajari
NBER Working Paper No. 17647
December 2011, Revised August 2013
JEL No. D86,H57,L92

ABSTRACT

Deadlines and penalties are widely used to incentivize effort. We model how these incentive contracts affect the work rate and time taken in a procurement setting, characterizing the efficient contract design. Using new micro-level data on Minnesota highway construction contracts that includes day-by-day information on work plans, hours actually worked and delays, we find evidence of moral hazard. As an application, we build an econometric model that endogenizes the work rate, and simulate how different incentive structures affect outcomes and the variance of contractor payments. Accounting for the traffic delays caused by construction, switching to a more efficient design would substantially increase welfare without substantially increasing the risk borne by contractors.

Gregory Lewis
Department of Economics
Harvard University
125 Littauer Center
1805 Cambridge Street
Cambridge, MA 02138
and NBER
glewis@fas.harvard.edu

Patrick Bajari
University of Washington
331 Savery Hall
UW Economics Box 353330
Seattle, Washington 98195-3330
and NBER
Bajari@uw.edu

An online appendix is available at:
http://www.nber.org/data-appendix/w17647

1 Introduction

Public procurement is big business. In 2002, the European Commission estimated that public procurement spending amounted to 16.3% of the European Union's GDP (European Commission 2004). This fraction is typical of many developing and developed countries: in South Africa the World Bank assessed the share to be 13% of GDP (World Bank 2003).

Procurement costs and quality substantially depend on the terms of the contract between the procurer and the firm supplying the service. For example, in highway construction an important element of product quality is the time to completion, as ongoing construction delays commuters. Completion time depends both on factors under the contractor's control (e.g. inputs and work rate) and idiosyncratic shocks (e.g. bad weather, input delays, and equipment failure). Contractors can accelerate construction to get back on schedule after a negative shock, but this is costly, and so must be incentivized. Such incentives are typically provided by project deadlines and penalties for late completion. This fits neatly into the standard principal-agent framework: completion time depends both on costly (non-contractible) effort and a random shock, and the contractor's payment depends on the completion time.

In this paper, we examine how project deadlines affect contractor work rates and completion times, using data from state highway construction projects in Minnesota. Our dataset is unusually rich, as it contains day-by-day reports by the project engineers on weather conditions, delays, and planned and actual work hours. This allows us to get a measure of the shock by looking at how many hours of work were required to complete the project, relative to the best linear ex-ante prediction. Similarly we can construct a measure of the effort from the difference between their ex-post work rate and the best linear ex-ante prediction.

We find evidence of adaptation in response to the time incentives. While the distribution of shocks is continuous, the distribution of outcomes exhibits "bunching" at the project deadline, with many projects being completed exactly on time.[1] We also show that contractors increase their effort in response to negative shocks. This acceleration helps to avoid time overruns: contracts with bigger time penalties are less likely to be late.

We demonstrate that the current contract design is decidedly sub-optimal, once you take into account the time cost of commuter delays. To reach this conclusion, we structurally estimate the contractor's costs of acceleration following any shock. We use necessary optimization

[1] A recent literature in public economics has found similar "bunching" at kink points in the tax structure (Saez 2010, Chetty, Friedman, Olsen and Pistaferri 2011).

conditions to infer these costs: the contractor could have completed a day earlier or later, but chose not to, which implies restrictions on their marginal costs. Using the estimated costs, we perform counterfactual policy analysis.

The current policy is effectively a quota: complete on time or face penalties. Relative to a baseline of no time incentives at all, we find that the welfare gain from this quota policy is small, around $26,000 on a $1.2 million contract. We compare this to a contract that is linear in the completion time, charging the contractor 10% of the traffic delay cost for each day they take. This generates a much bigger welfare gain of $267,000 per contract. But because there is uncertainty, these higher powered incentives create risk. We are able to quantify this risk because we observe the shocks, and we find it to be relatively small: the standard deviation of contractor payments under the linear contract is only $12,000.

Our work complements the analysis of Lewis and Bajari (2011), who examined the use of scoring auctions to award contracts based on both time and price. That paper looked only at outcomes, whereas here we are able to examine the mechanisms behind the outcomes, and quantify risk. Both studies suggest substantial gains from improved contract design.

The theory literature on procurement has long emphasized the twin roles of asymmetric information and moral hazard (Laffont and Tirole 1993).[2] But the empirical literature has tended to focus on competition between firms for procurement contracts, which typically occurs through an auction.[3] By contrast, this paper stresses moral hazard, showing that it is important in practice: the welfare gains we find here are much larger in magnitude than the potential gains from shaving markups through improved auction design.

We also demonstrate that the moral hazard is at least in part "ex-post", using a testing framework similar to that of Chiappori and Salanié (2000). This sheds light on the timing assumptions in the theory, which is important for the optimal contract design problem. Finally, this paper forms part of the empirical literature on high-powered incentives and

[2]See also Laffont and Tirole (1986), McAfee and McMillan (1986) and Laffont and Tirole (1987). Recent theory papers have explored issues like contract renegotiation (Bajari and Tadelis 2001), make-or-buy (Levin and Tadelis 2010) and public private partnerships (Martimort and Pouyet 2008, Maskin and Tirole 2008).

[3]See for example Porter and Zona (1993) and Bajari and Ye (2003) on bid rigging, Hong and Shum (2002) on the winner's curse, Jofre-Bonet and Pesendorfer (2003) on estimation with forward-looking bidders, Krasnokutskaya (2011) on the econometrics of unobserved heterogeneity, Marion (2007) and Krasnokutskaya and Seim (2011) on bid preference programs, Gil and Marion (2009) on subcontracting, Li and Zheng (2009) on entry, De Silva, Dunne, Kankanamge and Kosmopoulou (2008) on the release of public information before the auction, Decarolis (2013a) and Decarolis (2013b) on comparisons between the first-price and average auction format. A notable exception is Bajari, Houghton and Tadelis (2013), which shows that contractual renegotiation imposes significant ex-post costs.

their effects on output, which has mainly focused on labor contracts within the firm (see e.g. Prendergast (1999), Lazear (2000) and Bandiera, Barankay and Rasul (2005)).

The paper proceeds as follows. Section 2 presents an overview of the highway procurement process. Sections 3, 4 and 5 respectively contain the theoretical, descriptive and policy analysis. Section 6 concludes. All tables are to be found at the end of the paper.

2 The Highway Construction Process

We emphasize key features of the process in Minnesota that inform our later modeling choices. Figure 1 gives a simple timeline, starting from when the contract is awarded. At that time, the winning contractor must post a "contract bond" guaranteeing the completion of the contract according to the design specification. As a result, defaults are rare.

Once the contract is awarded, the contractor plans the various distinct activities, such as excavation or grading, that make up the construction project. To do this, they work out how long each activity will take for a standard crew size, and then use sophisticated software to work out the optimal sequence to complete the activities in by using the "critical path method" (Clough, Sears and Sears 2005). The key feature of this technique is that some activities are designated as critical, and must be completed on time to avoid delay, while others are off the critical path and have some time slack. The critical activities are called the "project controlling operations" (PCOs).

The contractor presents his plan to the project engineer in the pre-construction meeting. It is considered good practice to choose a plan that allows some contingency time on the side (around 5% of the time allowed). But a busy contractor may select a plan that allows little or no margin for error, or alternatively plan to finish early and move onto another project. This may be affected by the time incentives that are offered. In Minnesota, the incentives are usually simple. The design engineer initially specifies a number of "working days" that the contractor is allowed to take to complete the contract. A "working day" is a day on which the contractor could reasonably be expected to work. Usually this means weekdays (excluding public holidays) with amenable weather conditions. When the contractor works, a working day is charged. When the contractor could have worked, but didn't, a working day is charged and a note is made of the hours of "avoidable delay". When working is difficult for reasons outside the contractor's control — for example due to poor weather conditions or

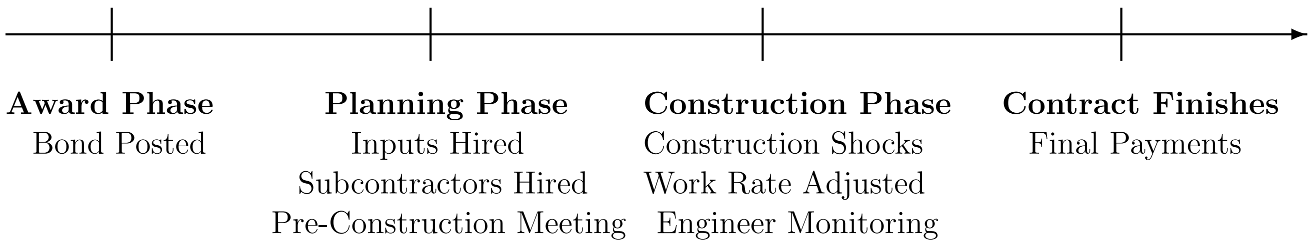

Figure 1: Construction Process

errors in the original project design — the project engineer may elect not to charge a working day. In this case a note is made of the hours of "unavoidable delay". The contractor may still choose to work on such days, and the hours of productive work are recorded; but the day does not count towards the project deadline.

Each additional day beyond the number of target working days is charged as a day late. Each day late incurs a constant penalty which depends only on the size of the contract. The penalties for being late are specified in the standard contract specifications, which we reproduce in Table 1. They are *standardized across all contracts* and *concave in project size*. The penalties were last increased in 2005. Notice that it is the big contracts that have the smallest penalties as a fraction of contract size, and are also most likely to finish late.

Once the planning is complete, the construction begins. During the process, the project engineer conducts random checks on the quality of the materials and monitors whether construction conforms to the design specifications. Productivity shocks, materials delays or unexpected site conditions may affect the rate at which any activity is completed, and the contractor must continually check progress against the planned time path. If necessary, the work rate may need to be amended, especially when there is delay on a critical path activity. At the end of the process, the contractor is paid the amount bid less any damages assessed for late completion. As we will see later, the penalties for late completion are rarely enforced. The reason for this low enforcement rate is that the project engineer may issue a change order during the project in which they agree to waive any time-related damages. They issue such a change order when they believe construction has proceeded to a point at which the road is available for safe use by commuters. For simplicity, we ignore the issue of enforcement in the theory below, but we return to this point in the data analysis.

3 Model

With this process in mind, we outline a model of ex-post moral hazard in highway construction. Contractors have private costs that depend on the amount of capital and the size of the work crew they employ, as well as on the number of hours per day they choose to work. Road construction inflicts a negative externality on commuters, so the efficient outcome requires accelerated construction relative to the private optimum. Faster construction in turn requires either increased scale (more capital and labor) or an increased work rate. We assume that the scale is determined at the start of the project, so that as productivity shocks occur, the firm can only adapt to those shocks by changing their work rate. The work rate is not contracted on, and so this adaptation is a form of ex-post moral hazard. Time incentives will affect the privately optimal work-rate, and we would like to know what the socially efficient contract design is.

We introduce a two-period model. In the first period, the contractor chooses a level of capital K, representing all the factors of production that will be fixed over the length of the project (hired equipment, project manager etc). They also fix the labor L. Following this, a shock θ is realized. This shock is anything that was unanticipated ex-ante by the contractor about the amount of work needed to complete the project. We will refer to it as a productivity shock below. Given K, L and θ, the project takes $H(K, L, \theta)$ man-hours to complete.

In the second period, the firm chooses a uniform work rate s (in hours per day). This in turn determines the number of days $d = H(K, L, \theta)/sL$ the project will take, since the number of man-hours of work completed each day is just sL. We impose some economically motivated restrictions on the total hours $H(K, L, \theta)$. Capital substitutes for labor, decreasing the number of man hours required ($H_K < 0$). Labor has declining marginal product, so that adding additional labor increases the number of man hours required ($H_L > 0$), though decreasing the number of days taken ($H_L < H/L$). Last, a good productivity shock corresponds to a low θ ($H_\theta > 0$).

The work-rate decision will be influenced by the time incentives laid out in the contract. We consider time incentives that take the following form: a target completion date d^T and a penalty c_D for each day late. These form of incentives are widely used in highway procurement; other forms of time incentives are called "innovative". One innovative design is the "lane rental" contract. In this design, the contractor pays a rental rate for each day of construction that closes a lane. For construction jobs that require continuous lane closure, this is a special case with a deadline of $d^T = 0$ and lane rental rate c_D.

The contractor is risk-neutral, and pays daily rental rates of r per unit of capital, and an hourly wage $w(s)$ to each worker. For algebraic simplicity, the wage function $w(s)$ is assumed to take the linear form $w(s) = \underline{w} + bs$, a base wage plus an increment that depends on the work rate, reflecting overtime, bonuses for night-time work etc. Overall the contractor's ex-post private costs for a given set of time incentives are:

$$\begin{aligned} C(s, K, L, \theta) &= \underbrace{H(K, L, \theta)w(s)}_{\text{labor costs}} + \underbrace{rdK}_{\text{capital costs}} + \underbrace{\max\{d - d^T, 0\}c_D}_{\text{time penalties}} \\ &= H(K, L, \theta)w(s) + rK\frac{H(K, L, \theta)}{sL} + \max\left\{\frac{H(K, L, \theta)}{sL} - d^T, 0\right\} c_D \end{aligned} \tag{1}$$

Traffic delay costs are assumed to be linear in the days taken, with the daily cost equal to a constant c_T. Linearity is a good approximation if traffic and delays are constant over time.

Discussion: We have assumed that the productivity shock is realized before the work rate decision is made, so that the contractor can choose when the contract will be completed. An alternative would be to make the contractor decide on his work rate before the productivity shock is realized, so that the completion time is stochastic. Both of these are imperfect approximations to a more complex dynamic process. The latter timing assumption is closer to standard principal-agent models, where the agent chooses an effort level that induces a distribution of (contractible) outcomes. As we will later show, the ex-post moral hazard model is better able to rationalize the observed data (in particular the large fraction of contracts that finish exactly on time).

We have also assumed the contractor is risk-neutral. This assumption is innocuous for most of the analysis, since we work with the ex-post profit function to see how incentives affect adaptation. But for the welfare calculations at the end of this section, the ex-ante joint welfare will vary with the contractor's risk preferences. We discuss the implications of alternative models at that time.

Finally, we assume that the quality of the construction is unaffected by the time incentives. One may worry that the contractor may shirk on quality to save time. This is a version of the famous multi-tasking problem of Holmstrom and Milgrom (1991). In highway procurement, this is less of a concern, as the government employs a project engineer to monitor the

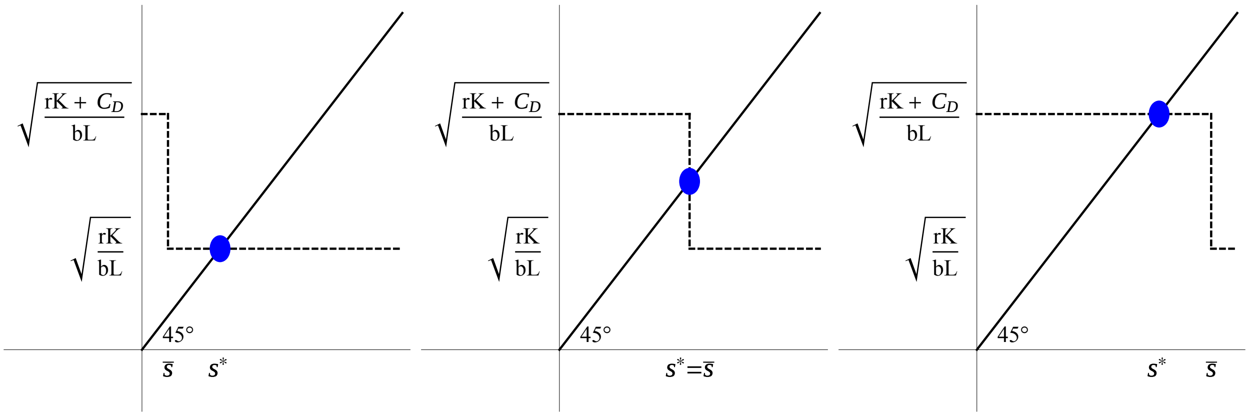

Figure 2: **Optimal Work Rates.** The figure depicts how work rates change with the number of hours of work required to complete the project. In the left panel, a favorable productivity shock means that a slow work rate would suffice for on-time completion, but the contractor works faster to economize on capital costs and will finish early. The middle panel shows a contract where no productivity shock has occurred, and the contractor works at a rate that leads to on-time completion. In the right panel, a negative productivity shock implies a fast work rate is necessary for on-time completion, but the contractor optimally chooses to work slower and will finish late.

construction and ensure that the finished project meets the contract specifications. Low quality construction is penalized by additional penalties laid out in the contract terms, and these penalties can be enforced against the contract bond.[4]

Adaptation: We start the analysis by looking at how the work rate s is chosen, given the realization of the productivity shock. Define $\bar{s} = \frac{H}{d^T L}$, the (ex-post) work rate required for on-time completion. Then taking a first-order condition in s, and dealing with various boundary issues resulting from the presence of the max operator in the objective function, we get the following expression for the optimal s^*

$$s^* = \begin{cases} \sqrt{\frac{rK}{bL}} & \text{if } \bar{s} < \sqrt{\frac{rK}{bL}} \\ \bar{s} & \text{if } \bar{s} \in \left[\sqrt{\frac{rK}{bL}}, \sqrt{\frac{rK+c_D}{bL}}\right] \\ \sqrt{\frac{rK+c_D}{bL}} & \text{if } \bar{s} > \sqrt{\frac{rK+c_D}{bL}} \end{cases} \tag{2}$$

There are three cases, corresponding to contracts in which the required work rate for on-

[4]Lewis and Bajari (2011) found that there was no difference in the number of quality violations detected between California highway construction contracts auctioned using scoring auctions (which emphasize cost and time) and standard auctions (which emphasize only cost).

time completion $\bar{s}$ is low (good productivity draws), those where it is intermediate (average productivity draws), and those where is it high (poor productivity draws). These cases are depicted in Figure 2 as the left, middle and right panels, respectively. When the contract is unexpectedly easy to complete, the contractor could work slowly and still complete on time, avoiding the wage premium for accelerated work. The countervailing incentive is that this ties up capital over a longer period, which is costly. Balancing these incentives, the contractor chooses $s^* = \sqrt{\frac{rK}{bL}}$, which is increasing in the rental rate and the capital-labor ratio, and decreasing in the slope of the wage premium. On the other hand, given a middling draw, the contractor chooses the work rate to complete on-time ($s^* = \bar{s}$), since accelerating to be early is too costly, and slowing down to be late incurs time penalties. Finally, when facing a poor productivity shock, the contractor chooses a work rate of $s^* = \sqrt{\frac{rK+c_D}{bL}}$ and finishes late. Higher time penalties imply faster work rates in this case.

So the contractor adapts his work rate in line with the productivity shocks, although the range of adaptation is bounded. Specifically, for fixed capital and labor, there are a range of shocks θ for which $\bar{s} \in \left[\sqrt{\frac{rK}{bL}}, \sqrt{\frac{rK+c_D}{bL}}\right]$, and in that range, the contractor will accelerate or decelerate work as needed to keep production on time. But no positive productivity shock could induce a slower work rate than $\sqrt{\frac{rK}{bL}}$, nor could a sufficiently negative one induce him to work faster than $\sqrt{\frac{rK+c_D}{bL}}$. To reduce construction time even after bad shocks, one needs high penalties c_D. These increase the maximum work rate of the contractor.

One important caveat to this analysis is that it is *short-run*, in that we are holding the capital and labor inputs fixed. If the procurer were to consistently offer more aggressive time incentives, contractors would learn to use more capital or labor than is standard, in order to get the jobs done faster (this would be a form of ex-ante moral hazard). Indeed the evidence presented in Lewis and Bajari (2011) suggests that when high-powered time incentives are offered, the contractors may be willing to adopt entirely non-standard work schedules, signing costly rush orders for inputs, using big work crews, and working 24 hours a day. These long-run changes can only reduce their costs below the short-run levels. Since we have no data on capital and labor, our analysis will focus on the short-run.

Welfare Analysis: For welfare analysis, it is useful to separate the time incentives from the contractor's other costs (capital rental and wages), and look at how much money the contractor can save by slowing down and completing the contract one day later. Writing

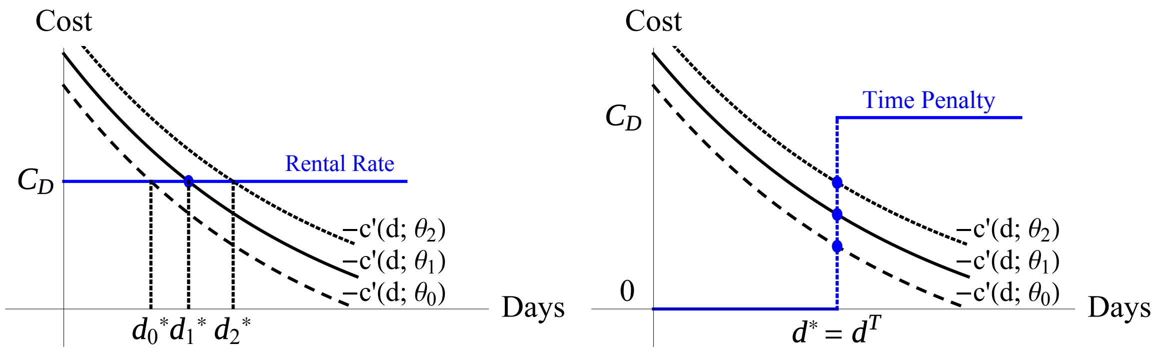

Figure 3: **Completion Time in Lane Rentals and Standard Contracts.** Both panels depict the marginal benefit to delay curve $-c'(d;\theta)$, drawn for three different productivity shocks. In the left panel, a lane rental contract imposes a constant cost of delay c_D, so the contractor optimally completes at d_0^*, d_1^* and d_2^* respectively, in each case equating marginal benefit and cost of delay. In the right panel, the incentive structure is standard, with damages charged after the target completion time d^T. In all cases the contractor will optimally complete exactly on time.

$c(d;\theta) = H(K,L,\theta)\left(\underline{w} + b\frac{H(K,L,\theta)}{dL}\right) + rKd$ for the cost of completing in d days for a given K, L, and taking a first order condition, we get:

$$-c'(d;\theta) = \frac{bH(K,L,\theta)^2}{d^2L} - rK \tag{3}$$

We refer to this as the marginal benefit of delay. It is strictly decreasing in d. The interpretation is that while an extra day of construction is useful to a contractor facing a tight work schedule (low d), it is less useful when the pace of construction is already rather slow.

In Figure 3 we depict how the time incentives affect the contractor's choice of completion time. The left panel shows the marginal benefit of delay curves for three different productivity shocks under a lane rental contract. Here the contractor faces a daily penalty of c_D right from day one, and thus the incentive structure is flat. Profit maximization implies that he equate the marginal costs and benefits of delay, and so for each shock θ_i he completes at d_i^*.

If the rental rate is set equal to the daily traffic delay costs c_T, the contractor internalizes the negative externality inflicted by the construction, and the social planner's problem is identical to that faced by the contractor. Accordingly, the contractor will hire capital and labor to minimize expected social costs (the sum of private and traffic delay costs), and efficiently choose the work rate given the realization of the productivity shock. Ex-post, the input choices may be sub-optimal, as they are not perfectly adapted to the productivity shock, but this is unavoidable given the timing.

The right panel of Figure 3 shows the same benefit curves under the standard incentive structure. Before the target date d^T, the contractor has no marginal cost of an extra day, since this is not penalized at all. But after the target, each additional day taken attracts c_D in time penalties, and therefore the marginal costs of delay jump discontinuously up from zero to c_D at d^T. In the figure this implies that for all three different productivity shocks the contractor will complete exactly on time. There is no incentive to complete early, as delay remains valuable; but also no reason to be late, as delay is not sufficiently worthwhile to offset the time penalties. This implies that completion times should be "sticky" at the target date: we should see many contracts finishing exactly on time.

In contrast to the simple lane rental design, the standard contract design will almost certainly lead to inefficient outcomes. The contractor should efficiently adapt to different productivity shocks by choosing different completion times, but the wedge in incentives makes this privately sub-optimal. In addition, there is little incentive to hire additional capital or labor at the planning stage to increase the probability of quick completion, since finishing early is not rewarded. This makes it difficult for the procurer to design a good incentive structure. On the one hand, setting $c_D = c_T$ at least ensures efficient adaptation for bad productivity draws (it sets the right penalty). But it may be preferable to distort short-run incentives with $c_D > c_T$, setting unreasonably high penalties. This increases the ex-ante incentive to hire additional capital, and thereby gives the contractor ex-post incentives to finish quickly. This is a second-best solution, creating a short-run distortion to offset a long-run distortion.

This analysis is similar to Weitzman (1974) on regulating a firm with unknown costs of compliance. The twist in this two-period model is that the contractor is also ex-ante uncertain, and only learns his costs after the incentive structure has been chosen. The lane rental is essentially a Pigouvian tax, and remains efficient in an ex-ante sense. The standard design is like a quota, and has the usual problem that the regulator has to set it without knowing the underlying costs of the contractor. This leads to inefficiency.

Risk Aversion: An important alternative way to look at this problem is to use a standard principal-agent model (Holmstrom and Milgrom 1987): the contractor is risk-averse, his "effort" is his work rate, the output is the number of days taken, and the productivity shock is the source of output uncertainty. When the space of contracts offered by the principal is restricted to incentive contracts that are a function of the completion time alone — the case here — the work rate is not contracted upon, which allows moral hazard. As we know from that literature, it is no longer optimal to transfer all the risk to the contractor by using the efficient lane rental: giving such high-powered incentives in the presence of productivity shocks increases the variance of the contractor's payments, lowering their expected utility. Weaker incentives are to be preferred under risk aversion.

It is hard to assess how important risk aversion is in describing contractor's preferences, although papers on skew bidding suggest that they are at least partially risk averse (Athey and Levin 2001, Bajari et al. 2013). Fortunately, up until the welfare calculations at the end of the paper, none of the empirical analysis relies on the assumption of risk neutrality.

4 Descriptive Analysis

The above theory indicates how contractors should adapt to productivity shocks, and how such adaptation is mediated by the contract design. In the remainder of the paper, we analyze data from contracts let by the Minnesota Department of Transportation (Mn/DOT). Our dataset is unusually detailed, as it includes daily reports by the project engineer on how construction is progressing. This enables us to test for ex-post moral hazard, seeing if contractors adapt their work rate in response to productivity shocks, and exhibit the "stickiness" in completion times predicted by the model. Having shown that the theory is largely confirmed, we estimate the contractor's short-run cost curves and use these to run some counterfactual simulations of alternative policies, such as lane rentals.

4.1 Data and Variables

The data comprises a selected set of highway construction contracts let by Mn/DOT during the period 1996-2005. It was provided to us by Mn/DOT themselves, as a set of files in a proprietary program called FieldOps that Mn/DOT project engineers used to record the daily progress of construction on their projects. We restricted attention to working day contracts for bridge repair, construction or resurfacing.[5] This yielded a sample of 466 contracts.

[5] About 30% of the contracts were "calendar day" contracts, in which the project deadline is a fixed date. Unfortunately the data quality in these contracts is bad, as project engineers typically don't bother to record diary data since it is unnecessary to keep track of working days. Of the remaining contracts, another 40% are for more superficial work that is unlikely to significantly impact commuters. See the supplementary appendix for more details on how the data is constructed.

The dataset includes daily information on the number of hours worked each day, the planned work schedule, the number of hours of avoidable or unavoidable delays recorded, what the weather conditions were like, and what the current project controlling operation was. We also see how working days were charged, and therefore can deduce whether the project finished early or late. Although our dataset is a panel, our analysis mainly uses the cross-sectional variation in contract outcomes, shocks and incentives.

We define the following time-related outcome variables: *hours worked* is the total number of hours worked, the analog of H in the theory; *unavoidable delays* is the total number of unavoidable delay hours; *unavoidable delay days* are the total workdays that were not charged due to unavoidable delays; *days worked* is the total number of days on which the contractor worked a positive number of hours; *days charged* is the total number of working days charged by the project engineer, the analog of d in the theory; *engineer days* is the number of days allowed, the analog of d^T in the theory; *work rate* is hours worked divided by days worked, the analog of s in the theory; and *engineer work rate* is the total planned hours divided by the engineer days.

There are also a number of contract characteristics that we observe. These include the contract value (equal to the contractor's winning bid), the time penalty (determined as in Table 1) and the contractor identity. We use the contractor identity to construct some additional firm-project-specific controls: the current backlog of the contractor[6], the firm capacity (calculated as their maximum backlog over the sample period), overlap with other projects[7], and whether the firm is located in or out of state. Throughout the analysis we have to make comparisons across heterogeneous contracts, and so we will often "normalize" a variable by dividing through by the engineer's days. We offer a structural motivation for these normalizations in the policy analysis section below.

We augmented our dataset by collecting data from the National Climatic Data Center (NCDC), on the daily amount of rainfall and snowfall at every monitoring station in their database for the period 1990-2010. Matching each project to data from the closest monitoring station, we construct four weather related measures for each contract. Two of these are ex-ante: *historical daily rainfall* is an average over the planned construction period of

[6]This is calculated as the sum of the outstanding contract value of all contracts this firm is working on, where the outstanding contract value for a contract is determined as the initial contract value, multiplied by the hours of work remaining, divided by the total hours of work for that contract.

[7]This is calculated as the fraction of days of construction on which the firm will also be working on at least one other project, if construction is carried out as planned by the design engineer.

the average daily rainfall for the full 20-year period; *historical chance of snow* is the average chance of snow across workdays in the planned construction period. The remaining two are measured ex-post: the actual average rainfall over the construction period, and an indicator for if it snowed during construction.

For the welfare analysis we need an estimate of the daily traffic delay cost. We calculated a contract-specific measure by multiplying the average daily traffic around the construction location by an estimate of the time value of commuters ($12/hour), and a conservative estimate of the delay that construction will cause them. Because estimating the delay required detailed manual work on Google Maps, we constructed these estimates only for a subsample of 87 contracts (the delay subsample). More details on both sample selection and the estimation of the traffic delay costs are available in the supplementary appendix.

We present summary statistics on the contracts in Table 2. A typical contract has value of about $1.2 million, and is of relatively short duration, around 37 days. During the contract, contractors work for 356 hours, at an average work rate of 9.3 hours a day (almost identical to the engineer work rate). A substantial number of both avoidable and unavoidable delays are recorded. Contracts are generally completed on time, although in the event that they are completed late, damages are assessed in only 24% of cases. As noted earlier, damages are often waived by the project engineer in the middle of the construction process, via a change order. This means that the contracts that are late are more likely to have been those on which the penalties were waived; conversely the (latent) enforcement rate on the contracts that were completed exactly on time was presumably much higher. We address this selection problem in the structural analysis. Damages, when assessed, range from $500 to as high as $29,000. By comparison, we project delay costs to commuters ranging from $0-124,000.

4.2 Graphical Analysis

The starting point for our analysis is an examination of the raw data. Look at the top left panel of Figure 4, which is a histogram of the days late across contracts. Recall that our theory predicts that the contract completion time will be "sticky" around the deadline, so that many contracts will be completed exactly on time. In the data, a full 11% of the contracts finish exactly on time, while 55% finish early and only 34% finish late.

To explore this connection to the theory further, we produce a number of other graphs that share a common logic and structure. The idea is to compare average outcomes from

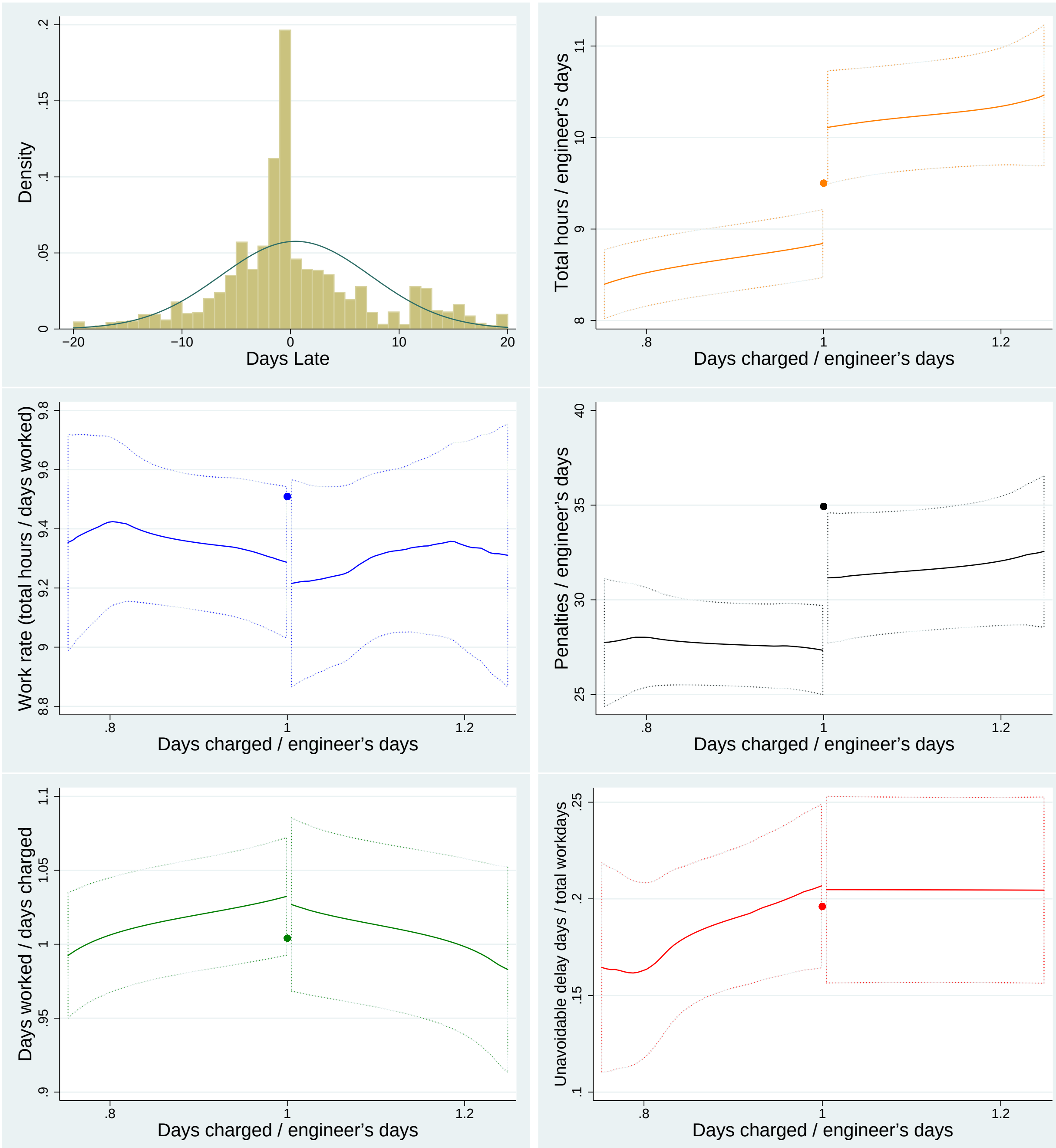

Figure 4: **Graphical Analysis.** The top left figure shows a histogram of the days charged minus engineer's days (i.e. days late), with a normal density function superimposed. Contracts completed exactly on time are included in the bar to the left of zero. In the remaining 5 figures, the x-axis is days charged divided by engineer's days (denoted $\widetilde{d}$), so that an on-time contract has $\widetilde{d} = 1$. Each of these figures plots the average value of the y-axis variable for on-time contracts as a dot, and the results of separate local linear regressions of the y-variable on the x-variable for early ($\widetilde{d} < 1$) and late ($\widetilde{d} > 1$) contracts. 95% confidence intervals for each regression are shown as dotted lines.

contracts that finished "just" early, to those that finished exactly on time, to those that finished "just" late. According to the theory, these contracts should differ from each other primarily in the size of the shock they experienced or the penalties for being late; and we should accordingly observe differences in contractor behavior across these groups (they should work faster with bad shocks or high penalties). However if the theory is incorrect and contractors are unresponsive to time incentives, one might expect little difference along these dimensions across contracts that differ only slightly in their completion time. To be clear, this is *not* a regression discontinuity design, as the forcing variable is endogenous. This is an initial look at the predictions of the theory, the empirical counterparts of Figures 2 and 3.

We implement this idea in the following way. The x-variable is the days charged divided by the engineer's days (denoted $\widetilde{d}$). For varying outcome variables (plotted on the y-axis in separate graphs), we run local linear regressions of the outcome variable on $\widetilde{d}$, separately for early and late contracts (using Stata's default choices of bandwidth and kernel). We also plot the average outcome for the contracts that finish exactly on time (as a dot).

The results, shown in the remaining panels of Figure 4, are striking. In the top right panel, the outcome variable is the normalized total work hours. Notice that the normalized hours jumps discontinuously from the left at $\widetilde{d} = 1$ and jumps again as we move to the right. This is *exactly* what the theory predicts: for positive shocks, contractors finish early; for moderate negative shocks, contractors accelerate construction and finish on time, but for bigger negative shocks they end up finishing late regardless. 95% confidence intervals are shown as dotted lines, and so we can see that these jumps are statistically significant.

The theory also predicts that the equilibrium work rate should be lowest on average in contracts that finish early, higher in those that are completed on-time and higher still in contracts that finish late (see equation (2)). We find partial support for this in the data: on-time projects have higher work rates than either early or late contracts (middle left panel, almost statistically significant at 5%). This suggests that one important way in which contractors respond to negative shocks is to accelerate their work rate, as in the model.

But it is puzzling that contracts in which the work was done on time have higher work rates than contracts that finish late. The reason is that these groups of contracts have systematically different time incentives. The middle right panel shows that the on-time contracts have higher normalized penalties than the contracts either finishing early or late (again statistically significantly), so the incentives to accelerate were stronger in these contracts.

We use this same approach to check if there is any evidence of adaptation along margins other than the work rate. In the model, contractors can only work faster or slower. But in practice, they can also adapt on the extensive margin, working on days that they are not required to (e.g. on weekends or days on which the project engineer does not charge them due to unavoidable delays). The outcome variable we use to test for this is the ratio of days worked to days charged. We find that if anything contracts completed on time have less work on uncharged days (bottom left panel), though this is not statistically significant.

Another way to adapt is by convincing the project engineer to chalk some days up to unavoidable delays (thereby effectively extending their project deadline). To check for this, we construct an outcome variable that is the number of workdays on which unavoidable delays were awarded, normalized by engineer's days. We find no evidence that early, on-time and late contracts differ along this dimension (bottom right panel).[8]

4.3 Testing for Moral Hazard

The above analysis is suggestive, but informal. It does not control for observable differences across contracts, nor does it rule out alternative explanations for the patterns in the data. We now develop a more formal approach to testing for ex-post moral hazard.

Let h_t be the total hours of work done on project t, normalized by the engineer's days (i.e. $h_t = \frac{H_t}{d_t^T}$). Let s_t be the work rate on the project. Let Ω_t be the contractor's ex-ante information set (i.e. everything they know about the project before construction begins, including their choices of labor and capital). Decompose the realized hours and work rate into an ex-ante expectation and an innovation:

$$\begin{aligned} h_t &= E[h|\Omega_t] + \theta_t \\ s_t &= E[s|\Omega_t] + u_t \end{aligned} \tag{4}$$

According to the theory, θ_t and u_t should be positively correlated: when the contract requires more work ex-post than was expected ex-ante, the contractor accelerates construction.

[8]In the supplementary appendix we continue this line of analysis, testing whether the normalized unavoidable delay days are correlated with contract characteristics, including firm and project engineer identity. The only significant correlations are negative: contracts with higher engineer work rates and contractors with more overlapping projects are awarded fewer delay days. The firm fixed effects are not jointly significant, although the project engineer fixed effects are, indicating some degree of heterogeneity across engineers in how they award delays.

Now, suppose further that the econometrician observes a collection of covariates x_t that is sufficient for the contractor's information, and that the conditional expectations are linear in the covariates:[9]

$$\begin{aligned} h_t &= x_t\beta + \theta_t \\ s_t &= x_t\gamma + u_t \end{aligned} \tag{5}$$

Then a test for ex-post moral hazard is to regress h_t and s_t on the covariates x_t and test for positive correlation in the residuals. This is similar to the Chiappori and Salanié (2000) test for asymmetric information in insurance markets, where they test for correlation between accident outcomes (an ex-post outcome) and the decision to purchase insurance. A different implementation of the same basic idea is to regress h_t on x_t in a "first-stage" and then use the estimated shock $\hat{\theta}_t$ as an additional regressor in a regression of s_t on x_t. Because this approach fits more cleanly into our later structural model, we run a series of these regressions and report the results in Table 3.

Consider columns (1) and (4). Column (1) is a first-stage, showing that the only statistically significant predictor of the normalized contract hours is the normalized contract value (bigger contracts require more work). Column (4) is the corresponding second-stage, and we find a statistically significant and positive correlation between the ex-post work rate and the residual from the first-stage. This suggests ex-post moral hazard.

The difficulty we face with this interpretation is that it could be that the contractor knew something that we have not controlled for (e.g. that they were going to deploy less capital than usual on this project), and therefore also planned to work harder ex-ante. More precisely, the correlation test is a test against the joint null that the econometrician's covariates are sufficient for the contractor's information set *and* that there is no ex-post moral hazard.

One way to be more confident in our conclusion is to add more controls and test if the correlation persists. This is what we do in columns (2) and (5) — where we add firm fixed effects — and in columns (3) and (6) — where we additionally include project engineer fixed effects. Though we find that both kinds of fixed effects are statistically significant (via Wald tests), the positive and significant correlation of the residuals with the work rates remains.[10]

[9]The linear specification assumption is unnecessary; with sufficient data these tests could instead be implemented non-parametrically. We prefer the linear specification here because we have many covariates.

[10]We are glossing over a technical issue here in the interests of simplicity: the hours residual is estimated in a first-stage and therefore the standard errors on the coefficients are too small, since they don't account for first-stage error. In the supplementary appendix we instead implement the regressions as a pair of seemingly

Another approach is to look for variables whose realization is plausibly unknown to the contractor ex-ante (i.e. is not in their information set) and that are also correlated with the total amount of work needed for the project, and then see if their realization is also positively correlated with the work rate. We consider two such sets of variables in Table 4, in specifications (1) and (2) respectively. The first is the *residual* hours of unavoidable delay on the project (i.e. a residual from the first-stage regression of unavoidable delay hours on contract characteristics x_t). While some amount of unavoidable delays are presumably anticipated by the contractor, the actual realization should be a surprise.

Then in the next three columns we regress the hours worked, work rate and ratio of days worked to days charged on the unavoidable delay residual and the same set of covariates. We find that in contracts with more delays, contractors end up working more total hours, have a higher ratio of days worked to charged, and work at a slower rate overall. Our interpretation is that unavoidable delays create extra work for the contractor by requiring them to shuffle their construction plans around. They do this by "smoothing" construction, doing some work on the days in which the project engineer is giving them a free day (due to delays) and are consequently able to work slower overall. This possibility is not in the model, where all adaptation is on the intensive margin (work rate) rather than extensive margin (which days to work). Nonetheless, if the unavoidable delays were unanticipated by the contractor, this is evidence in favor of adaptation.

The last three columns test whether weather shocks cause adaptation. We define rain difference and snow difference as the difference between realized and historical weather conditions, and use them as additional regressors. The evidence here is more mixed. We find a significant positive correlation between unexpected snow and total project hours, and a corresponding decrease in work rate, presumably because it is impossible to work when it is snowing. On the other hand, unexpected rain is basically uncorrelated with total hours and work rate, but is positively correlated with the day ratio, suggesting that contractors again smooth construction in response to rain. Taken together, the combined evidence from the direct tests with multiple controls (which are powerful but susceptible to unobserved heterogeneity) and the indirect tests based on delay and weather shocks (which are less powerful but more robust) are convincing evidence of ex-post moral hazard.[11] Notice that nothing in our analysis

unrelated regressions and test for correlation in the residuals, which avoids this problem. We still get a statistically significant positive correlation. We thank a referee for pointing this issue out.

[11]We also test whether for contractors working on multiple projects, shocks on one project affect work rates on other projects. The results are presented in the supplementary appendix. We find no evidence of

rules out adaptation on margins other than work rate (e.g. labor adjustments), though if these other margins are substitutes, it just makes it harder to detect work rate adaptation.

4.4 Time Incentives and Completion Times

We next present evidence that the contractors are responding to time incentives in choosing to adapt. We showed earlier in Figure 4 that there is an atom of on-time completions, and that these contracts have significantly higher time incentives than other contracts. We now pose a basic question: when incentives are bigger, are contractors late less often?

Table 5 shows the results from linear regressions of an indicator for the contract being late on the normalized time penalties and other covariates. In all specifications — including those with firm and project engineer fixed effects — we find that contracts with higher penalty rates, relative to the engineer's days, are significantly less likely to be late. The effects are quite big: they imply that for an average contract, doubling the penalties would reduce the probability that the contract is late by around 20%. These results together with the earlier graphical analysis are suggestive of a causal effect of disincentives on adaptation.

But we cannot rule out other explanations because we don't have a clean quasi-experiment (i.e. a convincing source of exogenous variation in the normalized time penalties). Perhaps contractors complete on time because there are non-pecuniary costs of late completion, such as acquiring a poor reputation.[12] In what follows, we attribute all of the adaptation that we see in the data to the incentive structure (this is implicit in the first order conditions we develop below). This may lead us to overestimate the responsiveness of contractors to incentives. We try to get a sense of how reasonable our results are later in the paper.

5 Policy Analysis

Having found evidence that contractors adjust their work rate in order to meet project deadlines, we would now like to assess different policy proposals for alleviating the negative

such interconnections, although this may be simply due to a lack of power to detect them (the coefficients have the right signs, but they are small and statistically insignificant). We also find no evidence that the atom in on-time completions is entirely due to contractors with multiple projects shifting their work around; when restricting the sample to contractors working on a single project, we still find an atom.

[12] We suspect this is not the case: reputation doesn't play much of a role in public procurement, because federal regulations give little discretion for selecting contractors on any basis other than cost or quality.

externality caused by construction. Our strategy for doing this is as follows. We estimate the contractor's short-run private costs of acceleration by looking at how their behavior changes as damages vary, using necessary conditions motivated directly by our theoretical model. With these in hand, we consider simple counterfactual policy changes, including perfect enforcement of penalties, accelerated targets and higher penalties. Of particular interest is a realistic case in which the lane rental is a constant fraction of the traffic delay cost, which is constrained efficient when the procurer faces budget constraints.

Our analysis is entirely short-run. We see how contractors adapted to different shocks under different incentive structures, given whatever capital and labor choices they had already made. This doesn't tell us how their input choices might change under a counterfactual policy, and so we are forced to hold everything constant. Fortunately, the bias can be signed: since long-run costs should be no more than short-run costs in expectation (since they are solving the same optimization problem with fewer constraints), we will tend to overestimate contractor acceleration costs, and therefore underestimate counterfactual welfare gains.

5.1 Estimation

Recall from the theory that the contractor acts to equate the marginal benefit of delay $-c'(d)$ with the marginal costs of delay, which are determined by the time incentives. We observe the number of days taken d. We also know the exact form of the time incentives. What we don't know is the exact form of the marginal benefit of delay function. So we make a linear specification choice, using the engineer days to normalize across contracts:

$$-c_t'(d) = \alpha d + d_t^T \left(\widetilde{x}_t \gamma + \delta \theta_t + \varepsilon_t\right) \tag{6}$$

where $\alpha < 0$, $\widetilde{x}_t$ is the same set of observables as x_t *excluding* the normalized penalties, θ_t is a latent productivity shock that we will estimate (see below) and ε_t is a latent cost shock. We assume that both θ_t and ε_t are unobserved ex-ante by the contractor, and observed ex-post (i.e. the contractor knows their marginal benefit function when they choose d).

In our specification, the marginal benefits of delay are linear in the days taken d, and the intercept is proportional to the contract target date d^T. What we have in mind is that contracts with a longer duration are typically more labor intensive, and so the benefit of delay in terms of reduced wages scales with the contract length. These scale, linearity and normality assumptions, although restrictive, have the advantage of being simple. We will show that we can fit the data quite well despite these restrictions. Nonetheless we will have to be careful about out-of-sample extrapolation that relies solely on these parametric choices. As noted earlier, we don't have plausible quasi-experimental variation, and must rule out endogeneity of the time incentives:

Assumption 1 (No Endogeneity) *$\varepsilon_t \sim N(0, \sigma^2)$, independent of $c_{D,t}$, $\widetilde{x}_t$ and θ_t.*

Measuring risk: To evaluate how much risk the contractor faces ex-ante, we want to estimate the productivity shock and include it as a covariate. This is admittedly only one source of risk (another might be wage volatility), but it is one we can measure. Recall our notation: h_t is the normalized hours; Ω_t is the contractor's information set; x_t are a set of observables (including the normalized time penalties); and θ_t is a latent productivity shock.

Assumption 2 (No unobserved heterogeneity) $h_t = E[h|\Omega_t] + \theta_t = x_t\beta + \theta_t$

This assumption allows us to estimate the productivity shock that the contractor faced in each contract. It is a pretty strong assumption, since it rules out asymmetric information between the contractor and the econometrician. To the extent that the assumption is violated, we will tend to overestimate the amount of risk faced by contractors, since part of what we call a shock will in fact have been anticipated ex-ante. But will show later that out estimates do not change much when we change the set of controls, which is re-assuring. In running counterfactuals and assessing risk, we will also need to be able to simulate what might have happened had other shocks occurred. To do this, we make an assumption on the distribution of productivity shocks:

Assumption 3 (Independence) *θ_t and x_t are independent.*

By construction, our estimate of the shock $\hat{\theta}_t$ will be orthogonal to x_t (it is a residual from a first-stage regression), but we strengthen to independence for the counterfactuals.[13]

Enforcement: Because contractors are typically informed via change order of whether the contract will be enforced or not, we assume that contractors know this when making their

[13]This assumption is in principle unnecessary, since we can estimate the conditional distribution of shocks. But it simplifies the analysis and we haven't found much evidence of correlation.

completion time decisions (so in some cases the time incentives are flat, and others have the familiar discontinuity). We model the enforcement decision as the realization of a Bernoulli random variable with parameter p. We assume p is independent of all other variables (i.e. enforcement is essentially random).[14] Notice that although we observe enforcement outcomes on late contracts, this doesn't immediately identify p, since late contracts are likely to be late precisely because the contractor knew they were not going to be charged any penalties. For this reason p is estimated jointly with the other parameters.

Estimation approach: The most straightforward procedure would be to use the first order conditions from the theory model developed in the earlier sections. But in fact the contractor cannot choose a continuous completion time and so the first order conditions needn't hold exactly. This is not just a technical point: one reason why we may see a mass of on-time completions *even in the absence of time incentives* is because completing on-time is better than completing a day later or earlier. Allowing for this possibility is important to deal with cases in which the penalties are not enforced. So instead we work with necessary conditions. Define:

$$\Delta^+(d) = 1(\text{enforced})1(d+1 > d^T)(d+1-\max\{d, d^T\})c_D$$
$$\Delta^-(d) = 1(\text{enforced})1(d > d^T)(d-\max\{d-1, d^T\})c_D$$

These are respectively the additional penalties from waiting another day and the penalties saved by completing a day earlier. The contractor could have completed at either $d-1$ or $d+1$, but chose to complete at d. So the following conditions are both necessary:

$$c(d) < c(d+1) + \Delta^+(d) \text{ and } c(d) < c(d-1) - \Delta^-(d)$$

Given parameters and enforcement realizations, these conditions imply bounds on the latent shocks ε which rationalize the observed completion times, and since these are normally distributed and independent of the covariates, we can deduce the likelihood of the data.

We use a two-step estimation approach. We estimate the shocks $\hat{\theta}_t$ in a first stage by regressing h_t on x_t and obtaining the residuals (as we did in the descriptive section above). Then, treating these estimates as data, we proceed by maximum likelihood in a second stage.

[14]The correlation between contract characteristics and enforcement is weak, at least for the (selected) sample of late contracts; see the supplementary appendix.

One difficulty is that while the enforcement realizations for late contracts are observed, they are latent for on-time contracts. To deal with this, we employ the EM algorithm, alternating between two steps: an expectation step, where we compute the vector of posterior probabilities that each on-time contract was enforced, given the current parameter estimates; and a maximization step, where we maximize the expected log likelihood given the current estimate of the posterior. We terminate the algorithm when the change in the enforcement probability $\widehat{p}$ across iterations is < 0.0001. We bootstrap both stages to get confidence intervals that account for the first-stage estimation error.[15]

Identification: It is worth thinking through how the model is identified. After making the assumption that marginal benefits of delay scale with contract length, we have a considerable amount of variation in the normalized time penalty c_D/d^T (and moreover under assumption 1 this is "good" variation). We get some additional variation from the fact that some contracts are observably enforced and others are not. This variation pins down the slope of the marginal benefit of delay curve. Consider for example two similar contracts that have experienced a similar negative productivity shock, but have different normalized time penalties enforced. Then the extent to which the contractor speeds up in the case with more costly time penalties tells us how expensive acceleration is, or conversely how valuable delay is on the margin. Once the slope is pinned down, the intercept of the delay curve $x\gamma + \delta\theta$ is identified by the distribution of time outcomes for a given set of covariates and shocks.

The slope is identified solely off outcomes for on-time and late contracts: for example, we never see how much acceleration there would be if a bonus was awarded for being early. This implies that the marginal benefit of delay curve is locally identified around the targeted days, but to get marginal benefits for $\widetilde{d} < 1$, we are relying on the linear functional form. With this in mind, most of our counterfactual simulations require only this local identification.

Results and Model Fit: We estimate three specifications, each one including more fixed effects.[16] Our results are in Table 6, and are similar across all specifications. We find that marginal benefits of delay are declining over time, and that they are increasing in the size of

[15] We derive the likelihood function and the posterior in the supplementary appendix.

[16] The estimation of the last model with firm fixed effects is tricky because there are 80 parameters to estimate and the likelihood function is hard to maximize. We used a two-part procedure for each maximization step of the EM algorithm, finding a parameter estimate that maximizes an approximation to the likelihood function (a much easier problem) in the first part, and then in the second part running a local optimization procedure from that solution with the correct likelihood function.

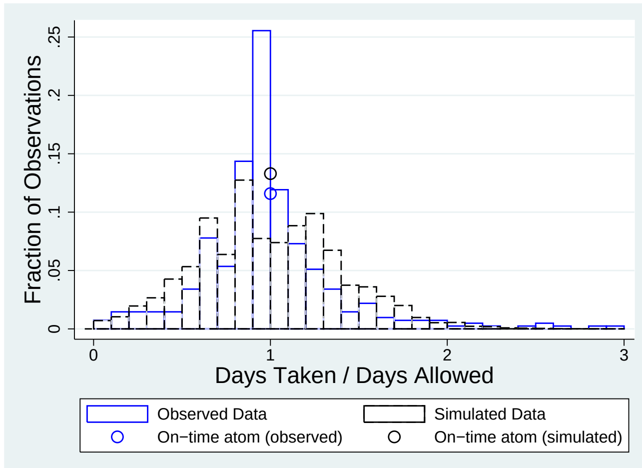

Figure 5: **Model Fit** Histograms of the normalized completion times $\widetilde{d}$ in the actual data (blue, solid) and those simulated from the structural model (black, dashed), including the atoms at $\widetilde{d} = 1$ (shown as circles).

the productivity shock. The latter is consistent with the theory in equation (3): when there is more work to be done, completing on the same schedule would require a higher work rate, and thus higher wage costs. It follows that delay is more valuable.

Our preferred estimate is the simplest one without any fixed effects, given in the first column (we want to avoid overfitting). We use this for the discussion and simulations in the rest of the paper. For an average contract, which lasts 37 days, our estimate of the marginal benefit of delay is $5250 - 4922d$, implying that the average marginal benefit to delay at the project deadline is \$128. This benefit to delay should be interpreted as the difference between the labor costs that would be saved and the increased capital costs from another day of construction. This compares with an average cost of delay of \$220, so that for many contracts the time penalties will actually bite.

We estimate that the enforcement rate p is 45%, implying that for the set of contracts completed on the margin (less than a day early), the enforcement rate was high: around 81%. So a large part of the atom of on-time completions is explained by the time penalties.

We also report how a one standard deviation shock affects the marginal benefit of delay. The estimated impact is large: a negative shock of that magnitude increases the benefit of delay by \$1220. This is slightly smaller than the standard deviation of the error term, so that the ex-post shocks explain almost as much of the cost differences as unobserved factors do.

The model fits quite well, although there are some problems caused by assuming a normal distribution for the error term. We examine fit in a number of ways. In Figure 5, we show histograms of the normalized completion times $\widetilde{d}$ in the actual data (blue, solid) and those simulated from the structural model (black, dashed), including the atoms at $\widetilde{d} = 1$ (shown as circles). This is intended as an informal check on the shape of the distribution. The model does a good job of capturing the incentives to be on time (the mass of on-time completions is very similar), but under-predicts the probability that a contract will finish just early or just late, as there are some non-integer completion times in the data (resulting from partial charges by the project engineers) that the model cannot simulate.

A more rigorous test is to compare some sample and predicted moments. We do this in Table 7. The strengths and weaknesses of the model are quite clear. We do a good job on predicting how long the contract will take, even conditional on contract size. But the model over-predicts the fraction of contracts that will be late, and under-predicts how late they will be conditional on being late. This is probably a consequence of the thin tails of the normal error distribution.

Finally, we compare the fit of the structural model to a simple linear model. We regress the total number of days d_t on the covariates $d_t^T, c_{D,t}, x_t, \hat{\theta}_t$ and a constant, getting an estimate $\hat{d}_t^{OLS}$ and corresponding $R^2_{OLS} = 0.85$. Then we construct an estimate $\hat{d}_t^{MODEL}$ as the average completion time for contract t across simulations from the structural model, yielding $R^2_{MODEL} = 0.88$ (in both cases, the R^2 is equal to one minus the ratio of the residual to the total sum-of-squares). This is encouraging, as it shows that the structural model actually fits better in the sense of R^2 than a simple reduced form, which need not be the case.

5.2 Counterfactuals

Now that we have contractor costs, we consider four counterfactual policy changes and estimate them on the subsample of contracts for which we have traffic delay cost estimates. In the first counterfactual, we consider what would happen if the penalties were enforced with certainty. In the remaining counterfactuals, we assume perfect enforcement and change other parameters. The second tightens the deadline to 90% of the current target, without changing the time penalties. The third policy is more contract specific: changing the penalties to 10% of the traffic delay cost of the contract. This policy is almost neutral with respect to average penalties: the existing average daily penalty on a contract in the sub-sample is \$1,140, while under this counterfactual policy it would be \$1,425. The final policy is a “lane rental contract", where the contractor pays a penalty each day from the beginning of the contract. The penalty is set equal to 10% of the traffic delay cost. This policy is a member of a class of constrained efficient policies that maximize welfare subject to the constraint that the total costs to the contractors not exceed a certain amount. This class is of interest because Mn/DOT faces a budget constraint of its own, and any costs it imposes on contractors through stronger time incentives will be passed-through, possibly at a rate higher than one. We do not consider the fully efficient policy of a lane rental equal to traffic delay cost, as this takes us too far out of sample.

To simulate counterfactual outcomes in a way that accounts for productivity shocks, we create copies of each contract, and assign to each copy a unique element of the vector of estimated shock realizations $\hat{\theta}$ (i.e. we create a dataset in which all contracts were hit with the estimated shock from contract 1, then from contract 2...). Then for all the contracts in the subsample and their copies — a dataset of size 87×466 — we sample from the estimated distribution of ϵ and average to get predicted outcomes for each contract-shock realization-policy triple. The outcomes are the mean completion time, commuter gain from shorter construction, additional private costs incurred by the contractor in accelerating the contract (excluding penalties), and the penalties paid by the contractor. We also calculate a welfare gain, defined as the difference between the gain to commuters and the increased costs incurred by contractors. In all cases, these are measured relative to a baseline with no penalties for late completion. We report the means of these statistics across contracts and shock realizations for each policy, and the average standard deviation in contractor costs, where the standard deviation is across shock realizations and the average is across contracts.

The results are in Table 8. We find that simply enforcing the penalties has a pretty big impact on completion times, and more than doubles the net gain. Tinkering further with the penalties and setting them equal to 10% of the delay cost actually slows down completion time, but raises welfare a little further. Tighter deadlines accelerate completion and raise welfare. But the big changes come from the 10% lane rental policy, which generates more than ten times the net gain. This is not surprising, since it aligns incentives not only when the contract is running late, but also when things are going well.

One concern is that the subsample of contracts for which we have estimated traffic delay costs is not representative of our full sample. To address this, we use propensity score re-weighting. First, we use a probit to estimate the probability that a contract is in the subsample given project characteristics, historical weather etc (no fixed effects). Second, we weight the relevant counterfactual statistics for each observation in the subsample by the inverse of the estimated probability of inclusion. If the unobservables have the same distribution across the full sample and subsample, this will correct for selection. Looking at the bottom panel of Table 8 we see that the results are similar in percentage terms, though smaller in magnitude. The welfare gains of a lane rental relative to the current policy are just over \$265,000 per contract, which is big relative to the current contract size of \$1.2M.

We also report the standard deviation of the contractor costs (as defined above). While the mean penalty is just a transfer from the contractor to the procurer and doesn't enter welfare calculations even under risk aversion, the standard deviation is a measure of the riskiness of the contract, and thus matters. We find that this risk increases as we move to full enforcement, and then increases again as a lane rental is introduced. In the case of a lane rental, the standard deviation of contractor costs is just over \$12,000. To put this into context, if the contractor has a 10% markup on an average \$1,2M contract, then a one standard deviation shock would wipe out just over 10% of their profit on the contract. This is perhaps not an unreasonable level of risk for the contractor to assume.

Discussion: How reasonable are our estimates? To get a sense of this, we compare our results with those in Lewis and Bajari (2011). That paper examines what happened when the California Department of Transportation offered stronger time incentives through the use of scoring auctions. When the time incentives offered were equivalent to a lane rental of roughly \$14,000 per day on much bigger contracts (nearly \$22M on average), contractors completed in approximately 60% of the allotted time. Our lane rental counterfactual here predicts that lane rentals of about a tenth the size (\$1,425) applied to contracts of about a twentieth the size (\$1.2M a contract) would accelerate completion times from 37.8 days to 30.6 days, which is about 19%. So if anything our counterfactuals here suggest less responsiveness in the Minnesota case. This may be because we are only picking up "short-run" responses through changes in work rate; the true "long-run" response to a change in policy may involve changes in the level of capital or labor used.

There are some important limitations of the above analysis. In some cases we are relying on the functional form of the parametric model. This is particularly true of the lane rental policy, where the marginal benefits of delay prior to the contract deadline are inferred based on the slope α of that function, which in turn was identified from late contracts. If the slope is not constant (e.g. marginal benefits are decreasing) we may underestimate the benefits to delay, and hence overestimate the effectiveness of the policy. To address these concerns we have mainly focused on small counterfactual policy changes, so that if the function is locally linear around $\widetilde{d} = 1$, our estimates will be approximately correct. We also assume the traffic delay costs are constant over the life of the project, whereas Mn/DOT's own enforcement policy suggests that they may decline once the construction reaches some milestone.

We ignore the fact that under the new contract regime, contractors may be selected for on the basis of their ability to complete quickly, and therefore the winning bidders may actually have lower marginal costs than those we estimated earlier (although they may have higher fixed costs). Moreover, they may make different decisions with respect to the hiring of labor and rental of capital than they currently do, enabling them to complete quicker without incurring the high costs we project when we hold labor and capital decisions fixed. This suggests we may underestimate the welfare gains from more high-powered time incentives.

But with these incentives comes an increased need for monitoring the contractors. This may be costly, since Mn/DOT would have to employ additional personnel on site. At least for the small changes considered here, this seems unlikely to be necessary. Finally, we do not account for general equilibrium effects due to the bidding up of input prices, or for the potential deadweight loss incurred in raising the funds for this accelerated procurement.

6 Conclusion

We have shown that moral hazard plays an important role in highway procurement. Contractors will adapt their work rate in the face of negative shocks, but only if incentivized to do so. This implies that careful contract design is a first-order concern. The current design is far from optimal: we estimate welfare gains of just over $265,000 per contract (over 20% of the contract value) from switching to a linear incentive scheme. While imposing risk on the contractors is a natural concern with high-powered incentive schemes, in this case the risk is small: the standard deviation in contractor payments will increase to $12,000, which is a small fraction of the contract value.

This paper has focused on highway contracts, because they are cases where the design is close to fully specified and adaptation is generally one-sided. This leaves room for interesting empirical research on cases where the design itself may need adaptation in the face of a shock, such as in military procurement or less routine construction.

References

Athey, Susan and Jonathan Levin, "Information and Competition in U.S. Forest Service Timber Auctions.," *Journal of Political Economy*, 2001, *109* (2), 375–417.

Bajari, Patrick and Lixin Ye, "Deciding Between Competition and Collusion," *Review of Economics and Statistics*, 2003, *85* (4), 971–989.

– **and Steven Tadelis**, "Incentives Versus Transaction Costs: A Theory of Procurement Contracts," *The RAND Journal of Economics*, 2001, *32* (3).

– **, Stephanie Houghton, and Steven Tadelis**, "Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs," 2013. forthcoming in American Economic Review.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *Quarterly Journal of Economics*, August 2005, *120* (3), 917–962.

Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Chiappori, Pierre-Andre and Bernard Salanié, "Testing for Asymmetric Information in Insurance Markets," *The Journal of Political Economy*, feb 2000, *108* (1), 56–78.

Clough, Richard H., Glenn A. Sears, and S. Keoki Sears, *Construction Contracting*, John Wiley & Sons, 2005.

De Silva, Dakshina G., Timothy Dunne, Anuruddha Kankanamge, and Georgia Kosmopoulou, "The impact of public information on bidding in highway procurement auctions," *European Economic Review*, 2008, *52* (1), 150 – 181.

Decarolis, Francesco, "Awarding Price, Contract Performance and Bids Screening: Evidence from Procurement Auctions," 2013. forthcoming in American Economic Journal: Applied Economics.

– , "Comparing Public Procurement Auctions," 2013. working paper, Boston University.

European Commission, "A report of the functioning of public procurement markets in the EU: benefits from the application of EU directives and challenges for the future," Technical Report, European Commission 2004.

Gil, Ricard and Justin Marion, "The Role of Repeated Interactions, Self-Enforcing Agreements and Relational [Sub]Contracting: Evidence from California Highway Procurement Auctions," 2009. Working paper, University of Santa Cruz.

Holmstrom, Bengt and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 1987, *55* (2), 303–328.

– **and** – , "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics and Organization*, 1991, *7*, 24–52.

Hong, Han and Matthew Shum, "Increasing Competition and the Winner's Curse: Evidence from Procurement," *The Review of Economic Studies*, 2002, *69* (4), 871–898.

Jofre-Bonet, Mireia and Martin Pesendorfer, "Estimation of a Dynamic Auction Game," *Econometrica*, 2003, *71* (5), 1443–1489.

Krasnokutskaya, Elena, "Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity," *Review of Economic Studies*, 2011, *71* (1), 293–327.

– **and Katja Seim**, "Bid Preference Programs and Participation in Highway Procurement Auctions," *American Economic Review*, 2011, *101.*

Laffont, Jean-Jacques and Jean Tirole, "Using cost observation to regulate firms," *Journal of Political Economy*, 1986, *94*, 614—641.

– **and** – , "Auctioning Incentive Contracts," *The Journal of Political Economy*, 1987, *95* (5), 921–937.

– **and** – , *A Theory of Incentives in Procurement and Regulation*, The MIT Press, 1993.

Lazear, Edward P., "Performance Pay and Productivity," *American Economic Review*, December 2000, *90* (5), 1346–1361.

Levin, Jonathan and Steven Tadelis, "Contracting for Government Services: Theory and Evidence from U.S. Cities," *Journal of Industrial Economics*, 2010, *58* (3), 507–541.

Lewis, Gregory and Patrick Bajari, "Procurement Contracting with Time Incentives: Theory and Evidence," *Quarterly Journal of Economics*, August 2011, *126* (3).

Li, Tong and Xiaoyong Zheng, "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions.," *Review of Economic Studies*, 2009, *76* (4), 1397 – 1429.

Marion, Justin, "Are Bid Preferences Benign? The Effect of Small Business Subsidies in Highway Procurement Auctions," *Journal of Public Economics*, 2007, *91*, 1591—1624.

Martimort, David and Jerome Pouyet, "To build or not to build: Normative and positive theories of public–private partnerships," *International Journal of Industrial Organization*, 2008, *26* (2), 393—411.

Maskin, Eric and Jean Tirole, "Public-Private Partnerships and Government Spending Limits," *International Journal of Industrial Organization*, 2008, *26* (412—420).

McAfee, R. Preston and John McMillan, "Bidding for contracts: a principal-agent analysis," *RAND Journal of Economics*, 1986, *17* (3), 326–338.

Porter, Robert and Douglas Zona, "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 1993, *101* (3), 518—538.

Prendergast, Canice, "The Provision of Incentives in Firms," *Journal of Economic Literature*, 1999.

Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2*, 180–212.

Weitzman, Martin L., "Prices vs. Quantities," *Review of Economic Studies*, 1974, *41* (4), 477–491.

World Bank, "Country Procurement Assessment Report for South Africa," Technical Report, World Bank 2003.

Table 1: Damage Specifications and Time Outcomes

	Damages per day ($)		Time Outcomes	
Contract Value ($)	1996-2004	2005	# Obs	# Late
Below 25K	75	150	2	0
25K-50K	125	300	10	0
50K-100K	250	300	26	0
100K-500K	500	600	151	33
500K-1M	750	1000	99	37
1M-2M	1250	1500	99	42
2M-5M	1750	2000	61	30
5M-10M	2500	3000	18	11

Damage specifications are taken from the Mn/DOT standard specifications for construction contracts. These were last amended in 2005. Time outcomes are based on the diary data.

Table 2: Summary Statistics: Mn/DOT Highway Construction Contracts, 1996-2005

	Mean	Std. Dev.	Min	Max
Contract value (thousands)	1199.0	1476.1	15.84	9546.2
Engineer days	37.35	26.88	10	198
Days worked	38.66	37.53	1	296
Days charged	38.68	35.36	1	271
Engineer's hour estimate	353.4	265.0	80	1756.3
Total hours worked	355.8	345.9	8.500	3015.5
Avoidable delays (hours)	64.99	104.4	0	921.3
Unavoidable delays (hours)	83.77	122.3	0	1099
Unavoidable delay days / engineer's days	0.186	0.251	0	1.600
Work rate (hrs/day)	9.327	1.503	4.232	14.36
Engineer work rate	9.396	1.427	6.595	15.19
Firm backlog / firm capacity	0.283	0.274	0	1
Overlap with other projects	0.389	0.409	0	1
Contract late?	0.328	0.470	0	1
Penalty applied if late?	0.235	0.426	0	1
Damages (if applied)	6335	7115.7	500	29000
Delay subsample				
Daily traffic	11362.2	16483.5	380	100000
Projected delay (mins)	10	7.229	0	20
Daily traffic delay cost	14248.3	20584.0	0	124000
Number of contracts	466			

Summary statistics for selected highway construction contracts let by the Minnesota Department of Transportation between 1996 and 2005. Capacity is measured as the firm's maximum backlog over the sample period. Overlap is the fraction of planned days that overlap with the planned construction for another project the firm was awarded. Work rate is calculated as total hours worked divided by total days worked. Unavoidable delays are those that were outside of the contractor's control; avoidable delays are those that were preventable in the judgement of the project engineer. Unavoidable delay days are working days that were not charged due to unavoidable delays. The delay subsample consists of 87 contracts for which we collected more detailed data on detour options, and projected a daily cost to commuters from traffic delays.

Table 3: Ex-post moral hazard: work rate adjustment

	Hours worked / engineer's days			Work rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Time penalty	-0.017	-0.005	-0.022	0.002	0.002	-0.000
	(0.018)	(0.017)	(0.018)	(0.006)	(0.007)	(0.008)
Engineer work rate	-0.173	-0.190	-0.301*	0.255***	0.226***	0.209***
	(0.120)	(0.133)	(0.163)	(0.049)	(0.059)	(0.073)
Contract value ($K/day)	0.086***	0.089***	0.097***	0.010**	0.010**	0.007
	(0.011)	(0.012)	(0.014)	(0.004)	(0.005)	(0.006)
Historical daily rainfall	0.002	0.011	0.057**	0.005	0.003	-0.009
	(0.020)	(0.023)	(0.029)	(0.011)	(0.013)	(0.015)
Historical chance of snow	0.036	0.031	-0.002	-0.008	-0.001	0.007
	(0.022)	(0.027)	(0.023)	(0.019)	(0.021)	(0.022)
Firm capacity > $3M	0.017			0.250		
	(0.410)			(0.172)		
In-state contractor	-0.526			0.291		
	(0.508)			(0.268)		
Firm backlog / firm capacity	0.126	-0.292	-0.622	-0.033	0.133	0.408
	(0.626)	(0.818)	(0.829)	(0.255)	(0.343)	(0.388)
Overlap with other projects	-0.297	0.290	0.278	0.307	0.375	0.075
	(0.424)	(0.516)	(0.535)	(0.189)	(0.246)	(0.289)
Hours residual				0.080***	0.098***	0.133***
				(0.021)	(0.026)	(0.033)
District/Work/Year FE	yes	yes	yes	yes	yes	yes
Firm FE	no	yes	yes	no	yes	yes
Project Engineer FE	no	no	yes	no	no	yes
Wald Test: Firm FE (p-value)	-	0.00	0.02	-	0.29	0.11
Wald Test: Engineer FE (p-value)	-	-	0.00	-	-	0.01
R^2	0.40	0.53	0.68	0.32	0.37	0.57
N	466	400	358	466	400	358

In the first 3 columns, the dependent variable is the total hours taken, normalized by the engineer's days. In the next three columns, the dependent variable is the work rate, defined as total hours worked divided by total days worked. In (2) and (5) the estimation sample consists only of contracts where each contractor participated in at least 3 contracts in the sample; additionally in columns (3) and (6) each project engineer has managed at least two contracts. All results are from OLS regressions, and standard errors are robust. Significance levels are denoted by asterisks (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Capacity is measured as the firm's maximum backlog over the sample period. Overlap is the fraction of planned days that overlap with planned construction on other projects the firm is contracted for. Hour residuals in columns (4)-(6) are residuals from the regressions in (1)-(3) respectively.

Table 4: Ex-post moral hazard: additional tests

	(1) Hours	(1) Work rate	(1) Day ratio	(2) Hours	(2) Work rate	(2) Day ratio
Time penalty	-0.017	0.002	0.002	-0.011	-0.001	0.002*
	(0.015)	(0.006)	(0.001)	(0.015)	(0.007)	(0.001)
Engineer work rate	-0.173	0.255***	-0.024**	-0.146	0.240***	-0.024**
	(0.111)	(0.047)	(0.009)	(0.111)	(0.049)	(0.009)
Contract value ($K/day)	0.086***	0.010**	0.004***	0.083***	0.012***	0.004***
	(0.010)	(0.004)	(0.001)	(0.010)	(0.004)	(0.001)
Historical daily rainfall	0.002	0.005	0.001	0.014	-0.002	0.001
	(0.024)	(0.010)	(0.002)	(0.024)	(0.011)	(0.002)
Historical chance of snow	0.036	-0.008	-0.002	1.319***	-0.704***	-0.013
	(0.043)	(0.018)	(0.004)	(0.488)	(0.215)	(0.041)
Firm capacity > $3M	0.017	0.250	-0.033	0.029	0.240	-0.030
	(0.393)	(0.166)	(0.033)	(0.395)	(0.174)	(0.033)
In-state contractor	-0.526	0.291	-0.049	-0.486	0.266	-0.047
	(0.611)	(0.258)	(0.051)	(0.613)	(0.270)	(0.051)
Firm backlog / firm capacity	0.126	-0.033	0.097**	0.062	-0.001	0.100**
	(0.582)	(0.245)	(0.048)	(0.584)	(0.257)	(0.049)
Overlap with other projects	-0.297	0.307*	-0.063*	-0.157	0.236	-0.068*
	(0.431)	(0.182)	(0.036)	(0.435)	(0.191)	(0.037)
Unavoidable delay residual	0.179***	-0.193***	0.020***			
	(0.065)	(0.027)	(0.005)			
Rain difference				-0.004	0.001	0.001**
				(0.007)	(0.003)	(0.001)
Snow difference				1.315***	-0.713***	-0.010
				(0.499)	(0.219)	(0.042)
District/Work/Year FE	yes	yes	yes	yes	yes	yes
R^2	0.41	0.37	0.25	0.41	0.31	0.24
N	466	466	466	466	466	466

Results of two sets of regressions. In each set (1) and (2), the dependent variables are the normalized hours worked (total hours / engineer's days), the work rate (total hours / total days) and the ratio of days worked to days charged. All regressions are by OLS. Standard errors are robust and significance levels are denoted by asterisks (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Capacity is measured as the firm's maximum backlog over the sample period. Overlap is the fraction of planned days that overlap with planned construction on other projects the firm is contracted for. The unobserved delay residual is the residual from an OLS regression of the unobserved delays on all the other controls. Rain (snow) difference is the difference between the realized annual rainfall (indicator for snow occuring) and the historical average rainfall (probability of snow).

Table 5: Determinants of late contract completion

	Contract late					
	(1)	(2)	(3)	(4)	(5)	(6)
Time penalty	-0.006***	-0.007***	-0.007**	-0.006***	-0.007***	-0.007**
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Engineer work rate	-0.028*	-0.042**	-0.056**	-0.028*	-0.042**	-0.056**
	(0.016)	(0.020)	(0.027)	(0.016)	(0.020)	(0.025)
Contract value ($K/day)	0.006***	0.007***	0.009***	0.006***	0.007***	0.009***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Historical daily rainfall	-0.006	-0.001	0.003	-0.006*	-0.001	0.003
	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)
Historical chance of snow	-0.004	-0.008	-0.014*	-0.004	-0.008	-0.014*
	(0.005)	(0.007)	(0.008)	(0.005)	(0.007)	(0.008)
Firm capacity > $3M	0.076			0.076		
	(0.064)			(0.059)		
In-state contractor	0.080			0.080		
	(0.081)			(0.078)		
Firm backlog / firm capacity	0.045	-0.104	-0.226*	0.045	-0.104	-0.226*
	(0.074)	(0.106)	(0.122)	(0.067)	(0.099)	(0.118)
Overlap with other projects	-0.022	0.056	0.218**	-0.022	0.056	0.218**
	(0.064)	(0.082)	(0.108)	(0.060)	(0.079)	(0.103)
Hours residual				0.055***	0.052***	0.048***
				(0.006)	(0.008)	(0.010)
District/Work/Year FE	yes	yes	yes	yes	yes	yes
Firm FE	no	yes	yes	no	yes	yes
Project Engineer FE	no	no	yes	no	no	yes
Wald Test: Firm FE (p-value)	-	0.00	0.03	-	0.00	0.02
Wald Test: Engineer FE (p-value)	-	-	0.00	-	-	0.00
R^2	0.21	0.31	0.49	0.32	0.38	0.53
N	466	400	358	466	400	358

In all regressions the dependent variable is an indicator for the contract finishing late. In (2) and (5) the estimation sample consists only of contracts where each contractor participated in at least 3 contracts in the sample; additionally in columns (3) and (6) each project engineer has managed at least two contracts. All results are from OLS regressions, and standard errors are robust. Significance levels are denoted by asterisks (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Capacity is measured as the firm's maximum backlog over the sample period. Overlap is the fraction of planned days that overlap with planned construction on other projects the firm is contracted for. Hour residuals in columns (4)-(6) are from the regressions in columns (1)-(3) of Table 3 respectively.

Table 6: Estimates of Marginal Benefit of Delay

	Days Taken / Days Allowed (raw coefficients)		
Hours residual	10.29	10.62	9.85
	(7.19,13.43)	(7.65,13.90)	(1.67,14.00)
Engineer work rate	-1.27	-3.71	-3.67
	(-4.57,1.62)	(-8.20,-0.35)	(-8.35,0.72)
Contract value ($K/day)	0.44	0.40	0.58
	(0.09,0.53)	(0.08,0.58)	(0.21,0.94)
Historical average rain	-0.31	-0.44	-0.49
	(-0.83,0.14)	(-1.14,0.09)	(-1.34,0.18)
Historical average snow	0.86	1.02	0.62
	(0.15,2.08)	(-0.05,2.14)	(-2.06,2.59)
Firm capacity > $50M	18.29	8.23	
	(7.69,32.70)	(-2.55,20.85)	
In-state contractor	8.69	-1.37	
	(0.52,21.43)	(-13.13,15.10)	
Firm backlog / firm capacity	16.75	13.23	-5.11
	(5.47,32.07)	(2.20,28.97)	(-48.18,8.60)
Overlap with other projects	-6.44	-4.10	3.59
	(-17.30,4.57)	(-15.11,8.48)	(-41.00,7.12)
Slope (α)	-131.80	-135.94	-136.94
	(-188.28,-94.05)	(-189.50,-99.91)	(-203.56,-96.32)
Std. deviation of error (σ)	41.86	40.02	35.84
	(30.66,53.32)	(28.36,50.08)	(24.12,44.09)
Enforcement probability (p)	0.45	0.45	0.45
	(0.38,0.51)	(0.38,0.51)	(0.37,0.50)
	Average marginal benefit function		
Intercept	5250.14	5360.97	5290.78
Slope	-4922.03	-5076.78	-5010.97
Impact of 1SD hours shock	1220.96	1124.59	916.46
Std. deviation of error	1563.12	1494.54	1311.52
District/Work/Year FE	no	yes	yes
Firm FE	no	no	yes
Sample size	466	466	400

Maximum likelihood estimates of the marginal benefit of delay function. The top panel has the raw coefficients; the bottom panel shows the implied function for an average contract. An average contract has value $1.2M and should be completed in 37 days. The covariates are defined as in the notes to the previous tables. Hour residuals are residuals from a regression of the normalized hours taken on all other RHS covariates, where the set of covariates varies across columns in the table. 95% confidence intervals are given in parentheses, where the endpoints correspond to the relevant quantiles from the marginal distribution of that coefficient across bootstrap simulations.

Table 7: Sample and Simulated Moments

	Data	Simulations		
	Mean	Mean	95% L.B.	95% U.B.
Days taken / Days allowed	0.99	0.99	0.96	1.02
Days taken / Days allowed if value >$1M	1.08	1.09	1.05	1.13
Days taken / Days allowed if value <$1M	0.93	0.93	0.89	0.97
Fraction of contracts late	0.33	0.42	0.36	0.46
Fraction late if value >$1M	0.47	0.61	0.53	0.67
Fraction late if value <$1M	0.24	0.29	0.24	0.35
Days taken / Days allowed if late	1.34	1.19	1.17	1.23
Days taken / Days allowed if late & value >$1M	1.31	1.19	1.17	1.24
Days taken / Days allowed if late & value <$1M	1.37	1.18	1.17	1.22
	Structural Model		OLS	
Simulated R^2 for Days taken	0.88		0.85	

Comparison of observed and simulated moments. The first column contains moments observed in the data; the second column contains moments simulated using the maximum likelihood estimates of the model parameters. Columns 3–4 give the lower and upper bounds of a 95% confidence interval for these moments, calculated as the relevant quantiles of the distribution of moments obtained when simulating with each of the bootstrap coefficient estimates in turn. The bottom part of the table shows a simulated R^2 for days taken. The OLS comparison uses OLS to predict the days taken, using all the covariates from the structural model (including the engineer days and a constant).

Table 8: Counterfactual Welfare Estimates under Alternative Policies

	Current Policy	Fully Enforced	10% Tighter Deadline	New Penalties (10% Delay Cost)	Lane Rental (10% Delay Cost)
			Subsample with delay data		
Days taken	40.13	37.82	37.00	38.41	32.37
	(38.49,41.11)	(36.37,38.70)	(35.27,38.00)	(36.96,39.25)	(29.45,34.51)
Commuter gain ($K)	35.26	78.15	92.58	115.88	376.47
	(20.74,46.80)	(51.89,100.04)	(62.12,119.41)	(80.63,132.72)	(282.92,452.46)
Acc. cost ($K)	1.31	2.92	3.51	3.75	16.71
	(0.78,1.72)	(1.94,3.64)	(2.40,4.37)	(2.79,4.17)	(13.15,18.75)
Penalties paid ($K)	3.00	6.67	9.09	3.97	35.26
	(1.99,3.55)	(4.34,7.78)	(6.41,10.33)	(2.60,4.75)	(27.14,42.16)
Net gain ($K)	33.94	75.22	89.08	112.13	359.76
	(19.95,45.08)	(49.90,96.43)	(59.71,114.98)	(77.66,128.64)	(269.77,433.54)
Std. deviation costs ($K)	3.53	7.84	8.88	6.97	14.57
	(2.65,4.11)	(6.27,8.96)	(7.30,10.01)	(5.34,8.14)	(12.57,16.37)
			Reweighted to match full sample		
Days taken	37.76	35.80	35.05	36.27	30.57
	(36.41,38.73)	(34.62,36.65)	(33.62,36.00)	(35.13,37.06)	(28.06,32.50)
Commuter gain ($K)	27.75	61.32	73.26	91.94	306.83
	(16.55,36.85)	(40.92,78.94)	(49.49,94.84)	(64.05,107.22)	(230.78,371.12)
Acc. cost ($K)	1.02	2.27	2.75	3.01	13.60
	(0.62,1.34)	(1.54,2.87)	(1.91,3.47)	(2.28,3.34)	(10.70,15.45)
Penalties paid ($K)	2.40	5.32	7.32	3.03	29.73
	(1.62,2.91)	(3.60,6.36)	(5.36,8.46)	(2.04,3.74)	(22.97,35.49)
Net gain ($K)	26.72	59.05	70.51	88.93	293.22
	(15.92,35.50)	(39.36,76.10)	(47.59,91.33)	(61.75,103.86)	(220.09,355.50)
Std. deviation costs ($K)	2.95	6.55	7.47	5.75	12.42
	(2.24,3.47)	(5.29,7.52)	(6.19,8.45)	(4.44,6.71)	(10.79,13.97)

Counterfactual welfare results under different policies. The top panel of results is estimated using the data subsample for which we have social cost measures; the second panel imputes counterfactual results for the full sample by reweighting the counterfactual moments for each contract by the inverse of the probability of appearing in the subsample. This probability is estimated by a probit, conditioning on the observables used elsewhere in the paper. The first column is simulated outcomes under the current policy, with random enforcement. The remaining simulations all assume perfect enforcement. The second column maintains the current deadlines and penalties. In the third column, the contract deadline is accelerated by 10%. In the fourth, the daily penalties for late completion are set equal to 10% of the social cost, while the final policy is a lane rental at 10% of the social cost. All statistics are calculated relative to a scenario where no time incentives are used. For example, "Commuter Gain" for any contract is the average number of days saved by using time incentives times the social cost. "Acceleration Cost" is the estimated additional cost to the winning contractor of accelerating construction relative to a no time incentives scenario. "Penalties paid" are damages the contractor pays for being late. "Net Gain" is the difference between commuter gain and contractor acceleration costs. "Std. deviation costs" is the standard deviation of the contrator costs plus penalties, where the standard deviation is across draws from the empirical distribution of estimated productivity shocks.