MICHIGAN
ROSS SCHOOL OF BUSINESS

Working Paper

Deadlines, Work Flows, Task Sorting, and Work Quality

Natarajan Balasubramanian
Syracuse University

Jeongsik (Jay) Lee
Georgia Institute of Technology

Jagadeesh Sivadasan
Stephen M. Ross School of Business
University of Michigan

Ross School of Business Working Paper
Working Paper No. 1199
August 2013

Deadlines, Work flows, Task Sorting, and Work Quality*

Natarajan Balasubramanian†, Jeongsik (Jay) Lee‡, and
Jagadeesh Sivadasan§

First draft December 2010
This draft August 2013

Abstract

We examine deadlines-induced behavior using large-scale, high frequency data on about 5 million U.S. patents and published applications. We motivate the study with a model of rational agents facing discontinuous incentives around deadline thresholds, without using time-inconsistent preferences invoked in behavioral economics models of deadline-related behavior. Consistent with our model predictions, we find notable clustering of more complex patent applications around potential deadlines at month-, quarter- and year-ends, along with a small to moderate decline in work quality around those periods.

JEL classification codes: O30, O31, O34, L25, M54

Keywords: Incentives, routines, productivity, time-inconsistent preferences, innovation, patents

*We thank Nick Bloom, John Bound, Charlie Brown, Bill Kerr, Joel Slemrod, and seminar participants at the University of Michigan, Indiana University and Georgia Institute of Technology for their comments. We thank Briana Sell and Dale Jackson for excellent research support. We also thank Jim Hirabayashi at the United States Patent and Trademark Office (USPTO) as well as a number of inventors and patent attorneys for generously sparing time to talk with us, and providing their inputs and comments. All errors are our own.

†Syracuse University, nabalasu@syr.edu

‡Georgia Institute of Technology, jeongsik.lee@scheller.gatech.edu

§University of Michigan, jagadees@umich.edu

1 Introduction

Repeated routines are an integral part of how organizations execute tasks (Nelson and Winter 1982). Mandatory financial reporting routines for publicly listed firms are perhaps the most visible example. Other examples include routines for internal management reporting, employee performance evaluations, capital and operating budgeting, and payroll and supplier payments processing. Closely related to these routines are 'deadlines', or specific end-dates to complete certain tasks, with penalties for delaying work beyond the deadline.[1]

While routine-related deadlines are widespread, empirical studies of their effects in workplace settings are relatively rare. Results from studies of responses to inter-temporal incentives (e.g., Asch 1990, Courty and Marschke 1997, Oyer 1998) suggest that work flows will be affected by such deadlines. Surprisingly, there is little evidence on the effects beyond those on work flows. In particular, two important and related questions have been under-investigated. The first, with implications for models of underlying worker behavior, is how sorting of tasks is influenced by deadlines. To our knowledge, there are no empirical studies that examine this question. The second, crucial for understanding consequences of deadlines for workplace productivity, is whether any shifts in work flows are associated with changes in work quality. Other than the notable exceptions of Cadena et al. (2011) and Carpenter et al. (2012), there are few studies on this topic. In this study, we use large scale, high frequency data on patents and patent applications to shed light on these questions.

We motivate our empirical investigations using a simple, but novel, model where agents trade-off penalties for missing deadlines with the costs of additional errors should they choose to accelerate task completion. The model predicts several deadline-related phenomena as results of rational responses to discontinuous incentives around deadline thresholds. Though our model is agnostic about why deadlines arise, in spirit, it departs from the broad motivation discussed in the behavioral economics literature: dissuading procrastination induced by "irrational" time-inconsistent preferences (Akerlof 1991, Clark 1994, O'Donaghue and Rabin 2001, Kaur, Kremer and Mullainathan 2010).[2] In so doing, we hope to offer an explanation of some deadline-related empirical regularities without invoking time-inconsistent preferences.

Broadly, our model makes three predictions. First, it unambiguously predicts bunching of

[1] We use 'routine' in a narrower sense than Nelson and Winter (1982) who define it to include anything 'regular and predictable'...[such as]...routines for producing things, [and] procedures for hiring and firing (p.14).

[2] Another motivation, not based on time-inconsistent behavior, is that deadlines help with coordination and control (Simon 1947, Nelson and Winter 1982, March and Olsen 1989, Knott and McKelvey 1999). Becker (2004) reviews the literature on organizational routines. Besides coordination and control, other potential objectives of routines include economizing on cognitive capacity (Simon 1947, Hayek 1952), reducing uncertainty (Becker 2004), helping improve learning (Langlois 1992), and storing tacit knowledge (Nelson and Winter 1982).

work close to deadlines; this arises from the discontinuity of incentives, which makes it worthwhile for agents to accelerate some of the tasks to finish them before the deadline. Second, under some conditions, the model predicts that more complex tasks are more likely to be accelerated for completion. Finally, if more complex tasks are accelerated, and if the deadline penalties are low enough, then tasks completed close to deadlines are on average more error prone.

Our empirical analysis uses large-scale data on granted patents (about 3 million) and published applications (about 1.9 million) in the U.S, and consists of three parts. First, we examine whether work flows, as measured by patterns in patent filings, are influenced by periodic deadlines. We then analyze characteristics of patents and patent applications to investigate if relatively more (or less) important tasks are completed near deadlines. Finally, we study whether there are any associated changes in work quality, i.e., whether the quality of patent applications completed close to deadlines are systematically different from those filed at other times.

Patent filings are a particularly interesting context because firms have strong incentives to avoid delays while simultaneously maintaining accuracy of the application. The date of filing the application is extremely critical in the patent application process; in most cases, this date becomes the "priority date" – the date that determines legal priority over rival claims on the same idea. Thus, in general, it is valuable to file earlier, as disputes related to delayed filing could be very costly.[3] However, submitting a premature, inaccurate or incomplete application to avoid delays is not costless. Such applications are returned by the USPTO with an 'Application Incomplete Notice', which adds to the review time and creates among others, a risk of loss of the priority date. Thus, firms and their lawyers have to carefully balance these tradeoffs when evaluating the timing of their filings.

If ideas arrive randomly (as assumed in most models of stochastic R&D) and are processed without accelerations or delays, the distribution of patent filings will be uniform over time. However, consistent with the presence of deadline penalties, we find strong clustering of patent filings in the data, with significant month-end, quarter-end, and year-end spikes in filing dates (Figure 2). For instance, the last 5 working days of a month account for 26.6% of all successful patent filings versus about 22.4% for the first 5 working days. Additional clustering is also observed at typical financial quarter-ends – March, June, September and December. These four months account for 36.3% of

[3] Under the laws in place during the period of our data, technically there were exceptions to the rule that the first to file gets priority for the idea. In particular, our data are all from the "first-to-invent" regime; i.e., in a dispute over priority, the inventor who can establish they were the first to invent (the invention specified under the patent) has priority. However, in practice even under the first-to-invent regime, the filing date was viewed by firms and lawyers as very important, as it had serious legal consequences for the inventors (Merrill, Levin and Myers 2004). We provide more details in Section 4. On September 16, 2011, the "Leahy-Smith America Invents Act" was signed into law, which made the filing date of paramount importance. This Act specifies a "first-to-file" system; i.e., in the event of a dispute, the inventor who files first is entitled to the patent.

all successful patent filings, compared to 33% in case of uniform filing. Similarly, firms file 12.9% of their patents in the month that their fiscal year ends, almost 50% higher than the expected uniform rate of 8.3% per month.

We then investigate reasons behind these work flow spikes. A number of different approaches strongly suggest that these clustering patterns are driven by routine-generated deadlines at large firms. First, corporate assignees exhibit these patterns but individual assignees, who are unlikely to have patent filing routines, do not. Second, patent filings of higher patent-volume firms exhibit significantly stronger clustering patterns. Third, changes in firms' fiscal years are systematically associated with changes in clustering patterns. Finally, detailed, semi-structured interviews with patent attorneys working for top patenting firms reveal a strong role for deadlines related to annual (and quarterly) corporate planning and reporting cycles and monthly billing cycles at law firms in generating the clustering patterns. These results are consistent with our model as well as with findings in the literature that document agents' modification of work efforts in response to incentives (e.g., Asch 1990, Courty and Marschke 1997, Oyer 1998).[4]

Turning to the analysis of task sorting, we examine three measures that are likely to be related to the quality (or importance) of the underlying idea: the log number of claims in the application, the log number of forward citations, and the probability of renewal of the patent. With all three measures, we find robust evidence that month-end applications are of higher complexity/importance (Figure 4). The economic magnitude is modest: for the last three working days, the number of claims is higher by 0.95%, the log number of cites by 0.6%, and the probability of renewal by 0.2%. These results are consistent with our model's prediction that under some conditions, when faced with discontinuous incentives around deadlines, agents sort tasks such that more complex work is more prone to be accelerated and completed just before deadlines.

The final piece of our analysis focuses on work quality. We use three measures of work quality: the probability of obtaining an 'Application Incomplete' notice from the USPTO, the probability of final approval, and the duration of review for successful applications. The first of these metrics is a shorter-term measure and a direct indicator of documentation quality. These notices are usually due to a failure to complete administrative formalities (e.g., omitting some of the required documents) and hence are a sign of deficiencies during the last stage of the patent filing process. Importantly, such notices require rework and re-filing by the lawyers and thus impose sizeable short-term costs on the filing firm. Consistent with our model prediction, we find that month-end applications are more likely to receive an 'Application Incomplete' notice from the USPTO: a 3.6% higher probability for

[4]In an interesting study, Oyer (1998) exploits variations in fiscal year-ends to document that manufacturing firm sales are higher at the end of the fiscal year and lower at the beginning, consistent with salespersons and executives responding to convex incentives which induce them to cluster effort in the last quarter of the firm's fiscal year.

the last 3 working days, relative to an overall average of 29%, which translates to a 12.4% higher incompletion rate.

Longer-term impacts on the ultimate fate of the application are more modest in terms of economic significance: the predicted approval probability of an application filed in the last three working days of a month is about 0.46% lower than that of applications filed by the *same* firm on other days. This translates into a 2.4% increase in rejection probability, given that the average approval rate is about 81% in our sample. Month-end applications also take slightly but systematically longer: the review length is about 1.5% greater for applications filed in the last 3 working days of the month.[5] The earlier findings on task sorting also rule out as an alternative explanation for lower month-end work quality: that patent attorneys complete work on the "higher value" ideas promptly, but postpone and then potentially rush work on lower value ideas close to deadlines.

Our work contributes to related literatures in a number of ways. First, we present a simple model of discontinuous incentives around deadlines that is able to explain key empirical regularities without invoking time-inconsistent preferences of agents. Second, by providing systematic evidence from an extremely time-sensitive setting, we add to the literature on the role of deadlines in altering work flows (e.g. Oyer 1998). Third, we contribute to the very limited literature on the relationship between deadline-driven behavior and output quality (Cadena et al. (2011), Carpenter et al. (2012)). To the best of our knowledge, our work is the first to use such large scale, high frequency data on work flows encompassing many firms to examine the effect of work clustering on work quality.

Fourth, while our study lacks advantages of randomized field and laboratory experiments (such as in Ariely and Wertenbroch (2002) and Cadena et al. (2011)), we contribute by looking at a very similar work process at a daily level across thousands of firms over several decades. Furthermore, the richness and scope of the data not only provide a very fine-grained look at an important work process but also allow us to use firm, period and other fixed effects along with other controls to condition for a variety of potential omitted variables, and rule out alternative explanations. Fifth, ours is one of the first studies to analyze the relationship between task characteristics and deadlines, and show that more complex tasks are more likely to be completed close to deadlines. This is particularly relevant for building a model of underlying worker behavior because models of procrastinating behavior such as O'Donaghue and Rabin (2001) have implications

[5]The weaker effects on the longer-term measures are not surprising. The quality or novelty of the underlying idea, which is set well in advance of the filing date, strongly affects the longer-term outcomes. Hence, they are likely to be noisier indicators of application draft quality, compared to the shorter-run 'Application Incomplete' measure which better captures effects of drafting errors or oversights. Therefore, while work completion close to deadlines is accompanied by significantly more errors, these are not so severe that it significantly delays or very substantially impacts the ultimate success of the application.

about which types of tasks may be more prone to procrastination. Finally, by studying the relatively under-examined process aspects of patenting, our paper also incidentally contributes to the literature on patents spawned by the seminal work of Hall, Jaffe and Trajtenberg (2001).

The paper is organized as follows. In Section 2, we present a model that generates predictions about the impact of deadlines on work flows, task sorting and work quality. Section 3 describes various data sources, while Section 4 provides a brief description of the patent application rules and processes. Section 5 analyzes deadline-driven behavior in the flow of patent applications. Section 6 examines task sorting around the month-end deadlines. In Section 7, we present evidence on work quality around deadlines along with robustness checks relating to lower work quality at close to month-ends. Section 8 briefly discusses a simulation that links our model to the empirical findings. Finally, Section 9 discusses and concludes.

2 A Model of Deadlines

In this section, we present a model of deadline-driven behavior that motivates our empirical examination and offers a potential explanation for our findings. In the model, agents complete independent tasks, possibly of varying complexities, that arrive randomly over time. The optimal time taken to complete a given task depends on the costs and benefits to the agent from completing that task. We model deadlines as periodically occuring 'checkpoints', when a monitor imposes a cost on the agent for having incomplete tasks. To keep the scope manageable, we do not model the monitor's decisions, and treat deadlines as exogenous to the agents.[6] Note that we are agnostic about the need for deadlines. The model does not endogenize the provision of the deadline penalty; thus deadlines may be related to coordination motives (Nelson and Winter, 1982) or designed to solve self-control problems (O'Donaghue and Rabin, 2001).

2.1 Tasks and Task Arrival Process

Tasks vary in complexity, defined by a continuous scalar, $x \in [1, X]$, with a CDF that has strictly positive support over the entire domain. The benefits, net of all costs other than errors (described below), to completing a task are assumed to be given by $\frac{a(x)}{t}$, with $a(x)$ positive, bounded and continuous, and $a'(x) > 0$. Thus, more complex tasks are more valuable, and the benefits of completing a task are inversely proportional to the time taken to complete the task. Task execution involves errors, which are costly. We assume that agents exercise the same level of diligence on all tasks and at all times so that, conditional on time spent, the error rate depends only on complexity.

[6] For instance, this is likely to be the case if these tasks are one of many processes in a firm, which then imposes monthly, quarterly and annual deadlines for co-ordination across these processes.

In particular, we assume an error function (normalized in cost units) as follows:

$$E(x,t) = \frac{b(x)}{2}\left(\frac{1}{t}\right)^2 \tag{1}$$

where x is complexity of the task, t is the duration and $b(x)$ is a positive, bounded and continuous function. Thus, speeding up execution increases error rate and associated costs quadratically. Since errors are more likely on complex tasks, we assume $b'(x) > 0$. Tasks arrive at every instant, t, with the number of arrivals per instant distributed identically, independently and uniformly over the set {0,1,2,...,M}. The complexity distribution is uncorrelated with the time of arrival, so that both number and complexity of applications are i.i.d over time.

2.2 Optimal Task Duration and Work Flow without Deadlines

For a task with complexity x, optimal task duration (t^*) is obtained by maximizing net benefit Π:

$$\Pi = \frac{a(x)}{t} - \frac{b(x)}{2}\left(\frac{1}{t}\right)^2 \tag{2}$$

This yields optimum duration as:

$$t^* = \frac{b(x)}{a(x)} \tag{3}$$

We assume that $a(x)$ and $b(x)$ are such that $\frac{b(x)}{a(x)}$ and $\frac{a(x)^2}{b(x)}$ are strictly increasing in x, and $\frac{b(1)}{a(1)} > 1$. Together, these assumptions imply that the optimal task duration is higher for more complex tasks, that the maximized net benefits are higher for more complex tasks, and that there is a one-to-one correspondence between x and t^*.

Definition 1. *Define $\tau(t^*)$ as the PDF of the distribution of optimal task durations, corresponding to the complexity distribution over the set [1,X]. Let $\overline{\tau}$ be the mean optimal task duration and $\underline{T}$ and $\overline{T}$ be the minimum and maximum durations, i.e., $t^*(1) = \underline{T}; t^*(X) = \overline{T}$. It can be verified that $\tau(.)$ is continuous and has positive support over $[\underline{T}, \overline{T}]$.*

Given our assumptions, the completion of tasks is uniform. In particular, the expected mass of tasks completed every instant is identical, and equal to $\frac{M}{2}\int_{\underline{T}}^{\overline{T}} \tau(t^*)dt^* = \frac{M}{2}$. The mean complexity of tasks, as measured by the optimal task duration, is $\overline{\tau}$.

2.3 Deadlines

Deadlines occur once for every length of time $D > 0$, defined as one period. At this point, the monitor imposes as an extra penalty (cost) on agents for having incomplete tasks at the deadline. In particular, we assume that this cost is $\gamma c(x)$, where $\gamma > 0$ and $c(x)$ (> 0) is bounded, continuous, and non-decreasing in complexity (i.e., the deadline cost is same or higher for more complex tasks).[7] For tasks that span multiple deadlines, we assume that the cost is imposed only on the last deadline.

We assume that a task of duration t^* arriving at instant d is completed at instant $d + t^*$. Let z be the time between the planned end time and the closest prior period-end deadline. Then $z = Mod(t^* + d, D)$, if $t^* > (D - d)$ where d is the instant that the task arrives. Since the arrival of tasks is uniform, for any given t^*, z will be distributed uniformly over the set $[0, t^*]$ if $t^* \leq D$, and over the set $[0, D]$ if $t^* > D$. Note that tasks with $t^* > D$ span across more than one period, and hence, the penalty is relevant only at the last deadline. Therefore, the acceleration of such tasks is limited to a maximum of D.

Assuming agents are forward-looking and have rational expectations, they choose whether or not to accelerate a task when it arrives. In particular, a task of complexity x is rescheduled to be finished by prior period-end (i.e., accelerated) if the following inequality holds:

$$\frac{a(x)}{t^* - z} - \frac{b(x)}{2}\left(\frac{1}{t^* - z}\right)^2 > \frac{a(x)}{t^*} - \frac{b(x)}{2t^{*2}} - \gamma c(x) \tag{4}$$

Rearranging terms and using result (3) we get:

$$f(t^*)\left[\frac{z^2}{2t^{*2}(t^* - z)^2}\right] < \gamma \tag{5}$$

where $f(t^*) = \frac{b(x)}{c(x)}$. Let $\underline{F}$ and $\overline{F}$ be the lower and upper bounds of $f(t^*)$ respectively.

2.4 Work Flows

Lemma 1. *For any given t^* and γ, we can find a z close enough to zero that inequality (5) holds.*

This follows from the fact that the numerator z^2 can be made arbitrarily close to zero while keeping the denominator bounded below at a positive value.

Lemma 2. *For any given t^* and γ, there is an upper bound, $z^*(t^*, \gamma)$ above which it is not optimal to accelerate so as to finish the task by the deadline. The upper bound has the following properties: (i) $z^* = t^*\left[\frac{t^*\mu}{1+t^*\mu}\right]$ where $\mu = \sqrt{\frac{2\gamma}{f(t^*)}}$ (ii) z^* is increasing in γ for any given t^* (iii) $\frac{z^*}{t^*}$ is increasing in γ for any given t^* (iv) $\operatorname{sgn}\left[\frac{\partial^2(z^*/t^*)}{\partial t^*\partial\gamma}\right] = \operatorname{sgn}\left[\frac{\partial(z^*/t^*)}{\partial t^*}\right]$ and $\operatorname{sgn}\left[\frac{\partial^2 z^*}{\partial t^*\partial\gamma}\right] = \operatorname{sgn}\left[\frac{\partial z^*}{\partial t^*}\right]$ for low enough γ and $\operatorname{sgn}\left[\frac{\partial^2(z^*/t^*)}{\partial t^*\partial\gamma}\right] = -\operatorname{sgn}\left[\frac{\partial(z^*/t^*)}{\partial t^*}\right]$ for high enough γ*

An upper bound exists because the left hand side of inequality (5) is increasing in z for any given

[7] For instance, this may be appropriate if a firm bills its clients monthly, and bills them more for more complex tasks. So, missing the month-end billing deadline would mean waiting for another month to receive payments.

t^* and γ. The upper bound is obtained by setting the left hand side to γ and solving for z^*. See appendix for proof of the other properties.

Proposition 1. *In a regime with a deadline penalty, there is clustering of task completions at the period-end. Tasks of all complexities get accelerated, though not all to the same extent. The additional mass of tasks completed at the deadline is* $\frac{M}{2}\int_{\underline{T}}^{D}\left[\frac{z^*}{t^*}\right]\tau(t^*)dt^* + \frac{M}{2}\int_{D}^{\overline{T}}\left[\frac{z^*}{D}\right]\tau(t^*)dt^*$.

The proof directly follows from the lemmas, and z being distributed uniformly for any given t^*.

Corollary 1. *The clustering of task completions is increasing in* γ.

2.5 Task Sorting

Unlike in the above analyses of work flows, whether more complex or less complex tasks get accelerated to meet deadlines depends on $f(t^*)$. In particular, the slope of $f(t^*)$ determines if the fraction of tasks to be accelerated (as determined by $\frac{z^*}{t^*}$ for tasks with $t^* \leq D$, and $\frac{z^*}{D}$ for tasks with $t^* > D$) is increasing or decreasing in complexity.

Proposition 2. *In a regime with a deadline penalty, the relation between clustering and task complexity is as follows. If* $f'(t^*) < \frac{2f(t^*)}{t^*}$ *for all* t^**, there is greater clustering of higher complexity tasks at the deadline. If* $\frac{2f(t^*)}{t^*} < f'(t^*) < \frac{4f(t^*)}{t^*} + 2\sqrt{2\gamma \underline{F}}$ *for all* t^**, tasks at the ends of the complexity distribution are accelerated to a greater extent. If* $f'(t^*) > \frac{4f(t^*)}{t^*} + 2\sqrt{2\gamma \bar{F}}$ *for all* t^**, there is greater acceleration among tasks of lower complexity.*

Corollary 2. *If* $f(t^*)$ *is decreasing, then more complex tasks will be accelerated more than less complex tasks.*

See proof in the appendix. Broadly speaking, under the first condition, $z^*(t^*, \gamma)$ and $\frac{z^*}{t^*}$ are increasing in t^*. Therefore, the fraction of tasks accelerated to be completed at the deadline ($\frac{z^*}{t^*}$ for tasks with $t^* \leq D$, and $\frac{z^*}{D}$ for tasks with $t^* > D$) is increasing in t^*. This, in turn, implies greater clustering of more complex tasks. Under the second condition, z^* is increasing in t^* but $\frac{z^*}{t^*}$ is decreasing. Therefore, tasks at both ends of the complexity distribution are accelerated more.

Now we consider what happens to average task complexity of accelerated tasks as γ increases.

Proposition 3. *Consider a deadline-penalty regime where the penalties are s.t. higher (lower) complexity tasks are accelerated. In such a regime, as* γ *increases, initially at low values of* γ*, the average complexity of accelerated tasks increases (decreases). Eventually, beyond sufficiently high values of* γ*, the average complexity of accelerated tasks tends to* $\overline{\tau}$.

The first part follows from Lemma 2(iv). For tasks with $t^* \leq D$, the second part follows from the latter part of the same lemma. For tasks with greater complexity, suppose z^* is increasing in t^*.

Then, as γ increases, $z^* = D$ for tasks of sufficiently high complexity. This implies that all tasks of those complexities, irrespective of the instant of arrival, are accelerated. As γ increases further, the same condition is achieved for tasks of lower complexity, and eventually at a high enough γ, all tasks with $t^* > D$ are accelerated, irrespective of the instant of duration. Thus, beyond this γ, the mean complexity of tasks with $t^* > D$ stays constant. However, at this γ and beyond, not all tasks with $t^* \leq D$ are accelerated, and their mean continues to tend to $\overline{\tau}$. A similar argument holds if z^* is decreasing in t^*.

2.6 Work Quality

Proposition 4. *In a regime with a deadline penalty, conditional on complexity, there is a higher expected error rate for tasks completed at the deadline. As γ increases, for tasks with $t^* \leq D$, the error rate increases without bound. For other tasks, the error rate increases but is bounded.*

Note that, for a given t^* and z, the additional error due to acceleration is $\frac{b(x)}{2(t^*-z)^2} - \frac{b(x)}{2t^{*2}}$, which is greater than 0. Thus, conditional on complexity, the error rates are higher as long as $\gamma > 0$. Since z^* is increasing in γ, the error rate is increasing in γ. The boundedness of error on complex tasks follows from the fact that the maximum acceleration for tasks with $t^* > D$ is D.

Proposition 5. *In a deadline-penalty regime with a low enough γ s.t. for any given complexity, there is some fraction of tasks that are unaccelerated, and if $f'(t^*) < \frac{2f(t^*)}{t^*}$, unconditional on complexity, there is a higher expected error rate for tasks completed at the deadline.*

To prove this, note that at optimal duration, error rate is $E(x) = \frac{b(x)}{2}\left(\frac{1}{a(x)}\right)^2$. This, by assumption, is increasing in complexity x, which implies that error rate of unaccelerated tasks is increasing in complexity. In the presence of deadlines, and with these assumptions, a greater fraction of complex tasks are accelerated to be completed at the period-end deadlines. Since these accelerated tasks have a higher error rate, the average error rate for tasks completed at the period-end is higher than the mean baseline error rate.

Given the scope of this paper, we do not analyze the optimality of penalty structures and deadline frequencies. A brief analysis of the impact of decreasing D is presented in the appendix. We now turn to the data.

3 Data

We draw from a number of data sources and methods. The primary source of our patent data is the USPTO. We purchased from the USPTO a DVD that contains data on all utility patents granted between January 1976 and August 2009 (a total of 3,209,376 patents). From this DVD, we obtained the patent number, U.S. classes and subclasses, the number of claims and application year of each patent. Further, these data were used to compute the number of citations to a patent. For most of our analysis, we use only data on patents assigned to "organizations," identified by assignee codes '2' and '3'. Further, within these, we excluded patents assigned to universities and multiple assignees. Finally, we eliminated patents that were not applied on a USPTO working day (0.6% of patents).

We supplemented the USPTO patent data with the NBER Patent Data (2009), which is an updated version of Hall, Jaffe and Trajtenberg (2001). This dataset provides a dynamic matching between the firm name in Compustat and the assignee name in the USPTO patent data.[8] This dataset contains comprehensive information on the U.S. utility patents granted between 1976 and 2006. In order to assign ownership after 2006, we used the assignee names and identifiers from the USPTO data, and assumed that ownership of assignees did not change after 2006. In case a new assignee was observed in the USPTO data after 2006, it was treated as a new firm.

To examine the potential role of patent attorneys, we obtained data on patent attorneys associated with each patent from the Patent Network Dataverse (Lai, D'Amour and Fleming, 2009). These data are relatively complete only after 1995. In the data, Lai et al. also disambiguate the name of inventors and assign unique identifiers to each inventor that appear in patents granted between 1975 and 2008. In a robustness check, we use this information to control for inventor-level heterogeneity by exploiting inventor mobility between firms.

We also used application data that include both successful patent applications and those that were finally rejected. To analyze the probability of approval of applications, we purchased detailed data on patent applications from Fairview Research LLC. The data include information on all U.S. Pre-Grant Applications published by the USPTO between January 1, 2001 and December 31, 2010. In addition to the application filing date and application number, the database includes application type (utility or design), number of claims, publication type, and importantly (to control for firm fixed effects), standardized assignee names.[9] To merge the application data with other patent databases, we did some further standardization of the assignee names.[10]

To examine potential differences in the complexity of examination process across patents, we collected from Google Patents the transaction history of patent applications.[11] The USPTO records

[8] Firms often file patents under different names ("assignees"). Further, firm ownership may change over time due to mergers and acquisitions, whereas the USPTO assignee name is frozen at the time of patent grant. The NBER Patent Data enables us to reflect the ownership changes and thereby minimize measurement errors due to misidentification.

[9] Assignee names are often not available in USPTO data because they are not tracked closely by the USPTO; assignment is made in many cases some time after the application is initially filed, and while this information is made public, the published application on the USPTO site often does not report the assignee. Fairview Research tracks the assignee information available in other USPTO documents and attributes it to the patent application.

[10] All code used for data cleaning and analysis are available on request from the authors.

[11] The site's URL is http://www.google.com/googlebooks/uspto-patents-pair.html. Google, under the agreement with the USPTO, has made bulk downloads of these 'PAIR' data publicly available. As of August 2012, the Google

every administrative action associated with each patent application starting with the initial exam team assignment to the final grant or abandonment. This enables a tracing of the entire review process for each application up to the most current transaction. To allow for a sufficient time between filing and the final decision, we limited to the applications filed by U.S. firms during the period of 2001-04. As of January 2012, which was the time of our data collection, about 52% of the applications were available for downloads. Even with this subset, downloading the entire documents posed technical challenges as many of the files were very large in size (some over 8 GB in zipped format). Trading off these challenges against diminishing benefits of increasing sample size, we downloaded a 25% random sample of the available applications, stratified by file size.

We supplemented the above data with information from a series of interviews with practitioners (inventors and attorneys) who all have had extensive first-hand experience in filing patent applications. The purpose of these interviews, conducted between December 2010 and March 2011, was to obtain detailed information on the patenting process inside firms, as well as to gain insights on the clustering of patent applications (documented in Section 5.1 below). The list of interviewees included: (a) legal staff (mostly Intellectual Property (IP) attorneys) at seven large patent-intensive firms; (b) legal staff from six independent IP law firms; and (c) four inventors each of whom had multiple inventions patented at different global electronics and semiconductor firms. The interviews, conducted either via phone or in-person, lasted from a minimum of 20 minutes to a maximum of two hours. All of these interviews were semi-structured; we first asked the interviewee several standardized questions based on our observations and then opened it up to unstructured responses. The information collected through these interviews formed the basis of our discussion on the process of patent application in Section 4.2. The responses of the interviewees about observed patent flows are discussed in Section 5.5. Because we promised respondents confidentiality, names of respondents and individual firms are kept anonymous.

4 Patent Application Rules and Processes

Before proceeding to the empirics, we provide a brief institutional background by describing the general process of patent application and discussing the importance of filing on time. We particularly focus on the factors that affect the timing of patent application and discuss how routines and deadlines, both internal and external, may play a role in this process. We first discuss various rules that regulate the timing of patent application.[12] Based on our interviews with practitioners, we then provide detailed steps of a typical patenting process from the inception of ideas to patent filing.

PAIR bulk download project is still in progress.

[12]The specifics of the patenting procedure vary across jurisdictions; except wherein applicable, we limit our discussion to the case of the U.S., as this is the context of our empirical analysis.

4.1 Rules Affecting the Date of Patent Application

A patent is essentially a government-granted monopoly right to an invention. For a *complete* application, the filing date of a non-provisional patent application is the date on which the application is submitted for examination. Completeness is determined by the USPTO. Specifically, there are three main elements that must be present in an application to be complete: the specification, at least one claim, and the drawing. In addition, the application must also contain the inventor oath or declaration and the appropriate filing fees (McWha and Ryan, 2008).[13] For mailed filings, the postmarked date is the date of application. For electronic filings (which is the dominant form of filing today), it is the date of electronic submission. In principle, the filing of a patent application establishes the right of priority and the filing date of the first patent application becomes the priority date of the invention.[14] Based on this priority, the claimant can file a subsequent application in another jurisdiction for the same invention.

The filing date is also crucial in the event of a dispute over who was the first to invent. If two inventors generated the same invention independently, (at the time of our study) the U.S. applied a first-to-invent (FTI) rule to determine who is granted the patent.[15] Even under this system, however, the first person to file an application retained the *prima facie* right to the grant of a patent. If a subsequent applicant wanted to claim priority for the same invention, they had to institute interference proceedings to determine who the first inventor was. However, this procedure was "costly and often very protracted; frequently it moves from a USPTO administrative proceeding to full court litigation" (Merrill et al. 2004, pp. 124-125). Also, in practice, FTF has been the basis for an overwhelming majority of applications in the U.S., with less than 0.1 percent of the cases ending up in interference proceedings (Merrill et al. 2004). Hence, in general, the inventor had strong incentives to file an application as soon as possible to establish effective priority.[16]

4.2 Patent Application Processes

A patent can be legally assigned to an organization (typically the inventor's firm), or retained by the individual inventor(s). In the former case, it is the organization who owns the claims, while

[13]The Manual of Patent Examining Procedures (MPEP), 1.53b reads "...the filing date of an application for patent filed under this section...is the date on which a specification as prescribed by 35 U.S.C.112 containing a description pursuant to 1.71 and at least one claim pursuant to 1.75, and any drawing required by 1.81(a) are filed in the Patent and Trademark Office. *No new matter may be introduced into an application after its filing date*" (emphasis added).

[14]Under some conditions, the priority date is not the date of filing. Some rules allow for non-standard applications such as provisional applications, continuations and divisions to claim a priority date different from the actual filing date. While it is hard to identify such non-standard *applications* (because of limited information in the applications data), we checked and found key results robust to excluding such non-standard *approved patents* from our analysis.

[15]As of October 2011, all countries but the U.S. adopt the first-to-file (FTF) system that assigns the right to whoever first applies for a protection of the given invention by filing a patent application.

[16]Specifics of how a filer can establish an earlier priority date are set out in USPTO MPEP, 715.07(a). The U.S. switched to a FTF system on March 6, 2013, when the America Invents Act came into effect.

in the latter it is the inventor who owns all claims to the invention.[17] Hence, depending on the inventor's affiliation, the application process can be quite different.

For individual assignees, there is no typical application process; the inventor decides whether and when to apply for the protection of an invention.[18] Also, the inventor faces few organizational constraints in filing an application. The inventor may work with a patent lawyer for the actual filing but key decisions lie with the inventor.

In contrast, the patent application process for R&D personnel at corporate firms is more complex. The process typically begins with the inventor's disclosure of an invention to an internal review committee, and ends with either the abandonment of the idea or the filing of a patent application. Figure 1 illustrates the process flow for a typical firm that has established procedures for patent prosecution.[19] This "disclosure" document is not yet a full description of the idea; it could a be one-page memo or a short technical note that outlines the subject matter of the invention. The review committee typically comprises senior managers in the R&D department and the legal team that is responsible for patent prosecutions. The committee regularly meets to discuss and determine if ideas submitted from various research teams are worth pursuing for a patent.[20] If an idea is approved by the committee, it is then drafted into a formal patent application document. The inventor may write the first draft and pass it on to the legal staff, or the legal staff may write up the draft based on the information the inventor provides. In either case, typically multiple rounds of meetings take place between the inventor and the lawyer before the application document is finalized. Many firms entirely outsource this portion of the patent prosecution to external law firms, while some use a mix of internal and external lawyers.[21] Once the patent application is finalized, the lawyer (either internal or external) files it with the patent office in the corresponding jurisdiction (e.g., USPTO, EPO), which then gives a unique identifier to each application and assigns it to a patent examiner for a review.[22] During the examination process, the patent examiner may issue office actions that demand follow-up actions from the applicant

[17] We focus on firms here. For ease of exposition, we refer to inventors for whom there is no assignee in the USPTO data (i.e. with assignee code = 1) as "individual assignees" or "inventor-owners," and inventors whose patent is assigned to an organization on the patent document as "R&D personnel."

[18] About 14% of patents granted during 1976-2009 were unassigned, i.e., the ownership was retained by the inventors.

[19] Note that this flow is only broadly representative of the process followed by the firms we spoke with; specific practices may vary across firms.

[20] This step of assessment often involves more than a simple selection or rejection. For those ideas that are selected for the next step, the committee could further sort them into several "grade" levels depending on the expected value of the idea. Lower-graded ideas may be pursued for patenting in a limited number of jurisdictions.

[21] Our interviews suggest that most firms have accelerated outsourcing of the application work. We discuss more details on billing arrangements for external lawyers in Section 5.5. Based on data by Lai et al. (2009), and assuming that applications with both the name of a law firm and names of lawyers on their front page were filed by external lawyers, approximately 78% of patents filed after 1995 were by external lawyers.

[22] Since we are primarily interested in the timing of application, we do not discuss the internal processes in the patent office.

such as clarification on the invention's specifications. These requests are mostly handled by the lawyers who drafted the application. For many firms, external lawyers are also responsible for the maintenance of granted patents such as alerting the client to renewal notices and, upon the client's renewal decision, paying renewal fees on behalf of the client.

5 Deadlines and Work Flow Clustering

We now turn to the empirical tests. In this section, we first document distinct clustering patterns in patent filings at close to quarter- and month-ends over the calendar year. We then provide robust evidence that these spikes in work flows are systematically driven by routine-generated deadlines, especially at large corporate firms.

5.1 Work Flow Patterns over the Calendar Year

As discussed above, innovating firms face a strong incentive to file a complete patent application at the earliest possible date. Then, if ideas arrive randomly over time, we should expect to see the distribution of patent applications to be uniform over the calendar year. Contrary to this expectation, however, we find significant deviations from a uniform filing pattern.

In Figure 2, we plot each calendar day's average share of published applications filed between 2001 and 2009. Also plotted is the corresponding average share of granted patents (based on filing date) for the same period. The vertical lines denote month-ends. The figure shows a substantial upward spike in filings at the end of every month. September and June show the highest month-end spikes, while the spiking pattern is somewhat muted for May, November and December.[23] From this figure it is clear that, on the aggregate, the filing of patent applications significantly deviates from a uniform distribution and has periodic clusters at month-ends.

While consistent with our model prediction in the presence of deadline penalties, the clustering patterns do not necessarily imply that they are driven by routine-generated deadlines or deadline-related penalties. In the following sections, we provide a number of tests that help establish a strong association between clustering patterns and routines-related deadlines. First, we examine differences between patterns for individual assignees and corporate firms; we expect significantly stronger clustering patterns for corporates since the latter are likely to have established routines for dealing with patenting. Next, we look at the effect of firm patent volumes; because routines and related deadlines are likely to be more established in firms with greater patent volumes, we

[23] All three months without a substantial spike have U.S. holidays at the end of the month - Memorial day in May, Thanksgiving Day in November and Christmas in December. In May, even though there is no sharp month-end spike, we notice a gradual increase in the share from the beginning of the month to the end of the month. In December, the month-end spike is shifted to the days before December 25, suggesting that individuals try to complete their work before Christmas. The four *downward* spikes correspond to January 1, July 4 (U.S. Independence Day), November 11 (Veterans Day in the U.S.) and December 25.

expect the clustering effect to be positively correlated with firm size, as measured by annual patent volumes. Third, we examine the effect of fiscal year ends; we expect firms with different fiscal year ends to have different patterns of clustering, and we expect changes in clustering patterns for firms that change fiscal year-ends. Finally, we present a synopsis of responses from patent practitioners, which support a key role for deadlines at large firms and work arrangements between large firms and external attorneys for the observed clustering of patent filing. We also examine robustness of key empirical results to a range of additional tests.

5.2 Comparison between Individual Inventors and Corporate Firms

To compare filing patterns between firms and individual inventors, we first compute the daily average share of granted patents separately for corporates and individual assignees. Figure 3 plots the difference between the daily shares for these two categories of assignees. *Every* month-end shows an upward spike indicating that the patent filing rate for corporates (relative to individual assignees) increases sharply towards the end of the month.[24] These spikes are particularly noteworthy at the end of each quarter. Equally noteworthy is the significant difference between individual assignees and corporates in December, which is the end of the fiscal year for most firms. Hence, Figure 3 strongly suggests that the patent filing work flows for corporates are significantly different from those of individual assignees, and that the clustering patterns observed in Section 5.1 are driven by corporate assignees.

We test the differences formally using the following OLS specification:

$$\delta_d = \beta_1 D_{1-7} + \beta_2 D_{8-15} + \beta_3 D_{16-23} + \beta_4 D_{24+} + \epsilon_d \qquad (6)$$

where D_{m-n} is a dummy variable that is 1 for days m to n of the month, and 0 otherwise; $\delta_d = \overline{s}_d^F - \overline{s}_d^I$ where $\overline{s}_d^I$ is the average daily share of annual patents (or applications) for individual assignees defined as $\frac{\sum_{t=1976}^{2009} n_{dt}^I}{\sum_{t=1976}^{2009} n_t^I}$ where n_{dt}^I is the number of successful patents (or published applications) applied by individuals in day d of year t and n_t^I is the number of successful patents (or published applications) applied by individuals in year t.[25] $\overline{s}_d^F$ is defined similarly for corporates.

Table 1 presents the results. In column 1, the coefficients on Days 1-7 and Days 8-15 are strongly negative while the coefficient on Days 24+ is strongly positive. This indicates that large firms tend to patent at a significantly higher rate than individual inventors during the last 6-7 days of the month. In columns 2 and 3, we replace the dummy for Days 24+ with a dummy for the last

[24] A separate plot of the daily share for individual assignees (Appendix Figure A.1) shows almost no month-end clustering, except at the end of September, which appears to be in response to potential fee hikes at the USPTO, which take effect on October 1 each year.

[25] Note that $\sum \overline{s}_d^F = \sum \overline{s}_d^I = 1$ where the sum is taken over all days of the year.

day of the month and a dummy for the last 3 days of the month, respectively. The coefficients on these dummies are also strongly positive and higher in magnitude than the coefficient on Days 24+ in column 1. If patent filing patterns were uniform throughout the year and identical across firms, we would expect 0.27% of annual patents in a single day for both individual inventors and firms. However, based on column 2, firms file about 0.54% (computed as the sum of the coefficient on the last-day dummy and the constant term) of annual patents on the last day of a month, which is twice the expected daily rate.[26] Column 3 yields similar results when we use a dummy for the last 3 days of the month.

5.3 Impact of Firm Size

Next, we examine the end-of-month behavior of large firms using patent-level regressions of the following form:

$$D^{n}_{pjtm} = \alpha . y_{jt} + \tau_t + \nu_m + \epsilon_{pjtm} \tag{7}$$

where D^{n}_{pjtm} is a dummy variable that is 1 if patent (or application) p belonging to firm j was applied in the last n working days of the month m in year t, y is a measure of firm size defined as the number of patents (or applications) belonging to firm j that were applied in year t and τ_t and ν_m are year and month fixed effects, respectively.

Table 2 presents results for three different end-of-month periods (1-working day, 3-working days and 5-working days). The coefficient on firm size is consistently positive and significant at the 1% level. The magnitude of the coefficients increases as the length of the end-of-month period increases. Based on column 1, a 10% increase in firm-size is associated with an increase in the probability of its patents being applied on the last day of the month by 2.842 X 10^{-4}. This amounts to a 7.1% increase in the probability of filing on the last day of the month.[27]. Similarly, based on columns 2 and 3, a 10% increase in firm-size is associated with an increase in the probability of patent applications in the last 3 working days by 4.6% and the probability of patent applications in the last 5 working days by 3.3%. The results on published applications show comparable effects (columns 4 through 6). Hence, end-of-month clustering is stronger for larger firms, consistent with the expectation that deadlines related to large firm routines are driving work flow patterns.

[26] Since the dependent variable is a difference, it is possible that the coefficients reflect a *lower* than expected rate of filing by individual assignees at month-ends. We ruled this out by running a separate regression similar to Equation 6 with the daily share for individual assignees as the dependent variable. The magnitude of the coefficient on Days 24+ was slightly higher (significant at the 10% level) than the coefficient on Days 1-7. Thus, individual assignees do not appear to patent less towards the end of the month. This is also confirmed by Appendix Figure A.1. We also checked the robustness of the OLS regression results in Table 1 to using multinomial logit regressions, where we simultaneously examine the probability of a patent being applied in the first 3 working days and that of a patent being applied in the last 3 working days of a month, relative to that of a patent being applied in the other working days of the month. These results are available on request.

[27] Assuming that a year has 250 working-days and that applications occur uniformly over time and across firms, the expected probability of a patent being applied on the last working day of the month is 4 X 10^{-3}

As a robustness check, we undertook an analysis similar to that in Table 1, except comparing 'large' corporates to 'small' corporates, with 'large' defined as assignees with more than 100 patents filed over our data period, and 'small' defined as those with only one patent over the period. Results (presented in Appendix Table A.1) confirm that the month-end clustering patterns are more striking for large corporates.

5.4 Impact of Fiscal Year Differences

Public firms are required to report their financial accounts once every year. The need for such periodic reporting usually creates deadlines internal to the firm and imposes them on some of the firm's external suppliers. For instance, firms may align their internal management reporting cycles and performance targets with the fiscal year (Oyer, 1998) or insist that their suppliers submit their payment requests by a certain date for it to be recognized as payable for that year. Firms are free to choose their own year-long period as the fiscal year. Though a majority of U.S. firms choose the calendar year (January 1 to December 31) as their fiscal year, some do not. Because the choice of the fiscal year is unlikely to be influenced by patent filing patterns, cross-sectional differences in fiscal year-ends across firms provides potentially exogenous variation to test for the presence of clustering patterns related to financial reporting deadlines. We find that, across firms, clustering is indeed correlated with fiscal year-ends. Specifically, Table 3 presents the monthly volumes of patent filings for each month by the fiscal year-end. The data show that, within a fiscal year, the share of filings that occur in the last month of the fiscal year is significantly larger than that for any other month. For example, January accounts for 16.2% of all patents filed by firms with a January-ending fiscal year; the corresponding figure for February is 22.1%. In every case, this share is significantly higher than the share that would be expected were patenting uniform throughout the year (8.3%).

For a variety of reasons unrelated to patenting, firms occasionally shift their fiscal years. Following Oyer (1998), we exploit changes in fiscal year-ends to see if clustering patterns within firms shift in the expected direction. This addresses the potential concern that omitted firm-specific factors may be driving clustering patterns in the cross-sectional regressions. Specifically, we examine regressions of the following form:

$$D^o_{pjtm} = \gamma.D_{or} + [\alpha.y_{jt}] + \tau_t + \nu_m + \eta_j + \epsilon_{pjtm} \tag{8}$$

$$D^n_{pjtm} = \grave{\gamma}.D_{or} + [\alpha.y_{jt}] + \tau_t + \nu_m + \eta_j + \epsilon_{pjtm} \tag{9}$$

$$D^n_{pjtm} - D^o_{pjtm} = \bar{\gamma}.D_{or} + [\alpha.y_{jt}] + \tau_t + \nu_m + \eta_j + \epsilon_{pjtm} \tag{10}$$

where D^o_{pjtm} is a dummy variable that is 1 if the patent was applied in the month in which the old fiscal year ended, D^n_{pjtm} is a dummy variable that is 1 if the patent was applied in the month in which the new fiscal year ended, and D_{or} is an "old regime" dummy variable equal to 1 for the time period prior to the fiscal year switch. As before, y_{jt} is log firm size measured in annual patent volume, and τ_t, ν_m and η_j are year, month and firm fixed effects, respectively.

We estimate the above equations for a sample of "switchers," i.e., firms that switch fiscal years. In column 1 (specification 8) of Table 4, the coefficient on the old fiscal-year dummy is strongly positive, implying that the propensity to file at the old fiscal year-end diminishes after the switch. The magnitude indicates that, after the switch, the propensity to file a patent at the *old* fiscal year-end decreases by 3.6%. This is very substantial when compared to the expected uniform rate of filing (8.3%) in a month. In column 2 (specification 9), the coefficient on the old fiscal-year dummy indicates that, after the switch, the propensity to file a patent at the *new* fiscal year-end increases by 1.97%. Again, this is substantial relative to the expected uniform rate of filing. Column 3 (specification 10 without the firm size control) confirms the net change (5.6%). Adding the firm size control (column 4) yields very similar results. The results using published applications (columns 5-8) also suggest a clear shift in the clustering patterns around fiscal year changes. Together, these regressions provide robust evidence that fiscal year-end routines of firms drive clustering patterns in patent filings.

5.5 Evidence from Interviews with Practitioners

Our interviews with practitioners also suggest a causal association between routine-driven deadlines and work flow patterns in patent filing. In these interviews, we presented the findings on the basic clustering patterns (Figure 2) to a number of corporate and external patent attorneys. The respondents pointed to two types of routines for the observed patterns of clustering.

First, almost all the internal and external legal staff immediately pointed to planning and reporting routines within firms as the likely driver of the quarterly and year-end clustering pattern. Interviewees explained that patent-intensive firms often set annual targets for patent volumes, which were then broken down into quarterly targets. A particular reason for breaking annual targets into quarterly goals was for financial planning, as expenses incurred in applications had to be reflected on quarterly accounting statements for listed firms.[28] These targets and time-lines then impact both arrangements and incentives for lawyers as well as R&D personnel. The respondents also mentioned that most large firms entered into long-term contractual arrangements with external lawyers, with specific targets for quarterly and annual patent filings. Failure to achieve these targets within the indicated time period was typically viewed poorly, and could potentially lead to loss of business for the law firm.[29]

[28]Our interviewees suggest that the costs including legal fees for each application typically ranged from US$8,000 to US$12,000, so that these fees for patent-intensive firms could run to several hundred thousand dollars every quarter.

[29]Almost all the firms we spoke to also provide specific incentives to the R&D personnel to encourage patent

Second, many of the lawyers suggested that billing routines coupled with contractual arrangements between law firms and large corporate clients were the likely driver of the month-end clustering patterns. Large firms typically enter into annual or quarterly contracts with external lawyers, often for prosecuting a targeted volume of patents, and the term of the contracts usually coincide with the annual or quarterly reporting cycles for the firms. In contrast, firms with smaller volumes hire external lawyers on a case-by-case basis and usually for a limited term length. Payments appear to be made to the external lawyers only after the application has been filed. Because external law firms often have monthly billing cycles, the lawyers have an incentive to complete applications before the end of the month. Further, attorneys in the law firms are evaluated on the amount of hours billed each period and are monitored monthly for the number of hours yet to be billed. This creates an additional incentive for the external lawyers to meet the end-of-month billing deadlines, as does the need to manage cash flows.

5.6 Other Robustness Checks

We perform additional tests to check the robustness of the results reported above. First, we use data on lawyers collated by Lai et al. (2009) to analyze if the firm-size effect documented in Section 5.3 is driven by arrangements between large firms and external lawyers. Results are presented in Appendix Table A.2. In columns 1 to 3, the specifications are similar to Equation 7 (columns 1 to 3 of Table 2), except that we add lawyer-year fixed effects to control for lawyer size. The results confirm that, within a law firm in the same year, filings for larger clients are more likely to be clustered at the month-end.[30]

Second, in the firm size analysis (Table 2), we further include fixed effects for patent class as well as those for individual inventors. The former is to rule out potential sector-specific omitted variables that may be correlated with both clustering pattern and firm size (e.g., if larger firms are more active in certain patent classes). The latter rules out a role for inventor-level heterogeneity, e.g., a concern that patterns of behavior of individual inventors may be driving the patterns in application. The results, presented in Appendix Table A.3, show that the baseline results are

filings; often inventors are provided monetary incentives on filing of the patent as well as on successful grant of the patent. Further, in most cases the number of patents applied/granted figure formally or informally in performance evaluations of the R&D personnel. Performance evaluations tend to take place only periodically, normally coinciding with the fiscal year reporting cycles. Thus, R&D personnel also have incentives to facilitate filing of patents within the performance evaluation periods. However, given the fairly long lead time, often 3 to 6 months between idea disclosure and patent application, and given that the last phases of application are handled by the lawyers, the incentives for R&D personnel do not appear to be the primary driver of the clustering patterns. Consistent with their lack of involvement and control over the actual application filing, and in contrast to the lawyers, the inventors we interviewed were unaware of any clustering patterns.

[30] *Within-law firm* correlation between month-end clustering and client firm size could be due to work-flow on larger clients being processed in batches to economize fixed costs in filing applications. However, this explanation is incomplete - it is unclear why this batch processing happens precisely at the end of the month. Moreover, the effect of firm size survives a further control of lawyer-firm pair monthly volume. Hence, the month-end clustering observed in large firms is unlikely to be driven by batch processing of larger volume work by lawyers.

robust to the inclusion of these additional controls of sector- or inventor-level heterogeneity.

6 Deadlines and Task Sorting

An important implication of deadlines is on task sorting. Our model predicts that deadline penalties will influence task sorting; in particular, it predicts that if the penalties for not completing tasks are sufficiently increasing in complexity, then there will be a clustering of higher complexity, higher value work around the deadline. In our context, patent attorneys involved in filings likely face a variety of different applications. In particular, different applications may involve different amount of complexity, which is likely to be positively correlated with the importance of the applications. In other words, some applications may be simple and straightforward, while other more important patents are likely to be more complex. Given this variety, it is plausible that attorneys make different decisions with regard to shifting of work effort for simpler versus more complex applications.

Because directly measuring the quality or importance of the idea underlying each patent application is difficult, we examine three different observed variables that are likely to be correlated with task application complexity or importance. First, we use the number of claims as a measure of both the complexity of the application and the quality of the underlying idea (e.g., Lanjouw and Schankerman, 2004). Fortunately, information on number of claims is available also in the application data, allowing us to examine all filed applications. Second, we look at the number of citations to a patent within the 5 years from its application date. Though citations have been used in the literature as a proxy for patent quality (e.g., Hall, Jaffe and Trajtenberg, 2005), one drawback of this measure is that it is applicable only to approved patent applications. Nevertheless, if poorer (higher) quality ideas are systematically sorted towards month-ends, then we may expect the quality of month-end patents to be lower (higher) *on average*. Our third measure is the probability of renewal of the patent. It is likely that the patents that the assignee firm decides not to pay renewal (i.e., maintenance) fees are of lower value (Pakes, 1986).

The results using all the available data are presented in Figure 4 and Table 5.[31] In Figure 4, across all three measures of idea quality, there is a sharp increase in the measure towards the end of the month; the patent or application filed close to month-ends contain more claims, are cited more frequently and are more likely to be renewed. Regressions in Table 5 confirm this pattern. In columns 1 to 3, the number of claims is significantly larger for month-end applications: 0.9% larger for applications filed in the last 3 days of the month, and 1.7% larger for those filed on the last working day. Similarly, in columns 4 to 6, the 5-year citation count is greater for month-end patents: citation count is about 0.6% higher for patents filed in the last 3 working days of the month. Finally, in columns 7 to 9, month-end patents are more likely to be renewed, by about

[31] We checked and found the results in Table 5 robust to using data for only 2001-04 (see Appendix Table A.4).

0.2%. Together, these tests strongly suggest that applications of higher compexity/quality ideas are more prone to be completed towards month-ends. Further, these tests also rule out the possibility of lawyers being less careful or more hurried with lower-value applications.

7 Deadlines and Work Quality

Beyond task sorting, another under-studied implication of deadlines is on the quality of work. Our model predicts that if complex tasks are accelerated, deadline-driven clustering in work flows will likely be associated with lower quality of the work product.[32] In this section, we examine work quality using three different measures: the probability of receiving an 'Application Incompete' notice from the USPTO, the probability of patent approval, and the duration of application review.

The first measure is an indicator variable for a notice from the USPTO that the submitted application is incomplete. This 'Application Incomplete' notice is sent by the USPTO if the minimum requirements of completeness discussed in Section 4 are not met, and implies that the submitted materials are not yet entitled to a filing date (McWha and Ryan, 2008). Usually, these notices are due to the applicant's failure to conform to administrative formalities such as omitting some of the required documents, missing signatures of some inventors, or using formats against the USPTO guidelines. Therefore, they are a good indicator of deficiencies during the final stage of the filing process. Though the applicants may be able to contest this notice and claim the original filing date as the priority date, these notices require rework and re-filing by the lawyers, and hence impose costs on the filing firm. The extra time and effort due to refiling could be a sizeable fraction of the cost of the initial application.[33] Furthermore, the additional processing delays patent approval. Thus, in general, it is in the best interests of the applicant to ensure that the application meets the minimum requirements for completeness.

To examine whether month-end applications are more likely to be declared incomplete, we use a sample of transaction history data from Google Patents and estimate the following regression model:

$$AI_{pjtm} = \beta_D . D^k_{pjltm} + \alpha . y_{jt} + \tau_t + \nu_m + f_j + \widetilde{\epsilon}_{pjtm} \tag{11}$$

where D^k_{pjltm} is a dummy defined earlier for the last k working days of the month, AI_{pjtm} is 1 if the application received an 'Application Incomplete' notice and 0 otherwise, τ_t and ν_m are application year and application month fixed effects, and f_j denotes firm fixed effects.

[32] To be precise, the prediction requires another assumption that the penalties are not high enough that certain tasks are always accelerated. This assumption is likely to hold in reality since firms are unlikely to set up an incentive structure that results in certain tasks being always accelerated.

[33] Our interviews with attorneys indicate that external attorneys charge client firms the costs associated with refiling unless errors are clearly their fault. One respondent indicated that their firm charged around $3,000-5,000 for each re-filing involving minor changes, and from $5,000-10,000 for those that involve more complicated responses. This is very significant compared to the $10,000-15,000 the firm charges for the original application filing.

The first three columns of Table 6 present the results. From column 1, applications filed on the last working day of the month are 4.6% more likely to receive an incomplete notice from the USPTO. The corresponding numbers for the last 3 and 5 working days of the month are 3.6% and 3.0%, respectively. Given that only 29% of applications in the sample receive such a notice, these differences are substantial (a 15.8%, 12.4% and 10.3% higher probability relative to the baseline mean rate).

Because a range of different types of errors induced by rushing of applications could result in them being deemed incomplete, the 'Application Incomplete' dummy is our preferred measure of work quality. While the rework related to 'Application Incomplete' notices could entail considerable costs for the attorneys, the more serious concern for the client firm is that the application is so sloppy that it gets rejected. Hence, as a second measure we examine the probability of approval of a patent. Because it takes several years for an application to be approved (or rejected), we limit our analyses to applications in the years 2001 to 2004; for similar reasons and for comparability, we use the same time period for the other two measures examined in this section. The results from regressions similar to Equation 11 are presented in columns 4 to 6 of Table 6. In all three regressions, coefficients on the end-of-the-month dummies are negative with the latter two columns showing strong statistical significance. Based on column 5, the probability of approval for applications filed in the last 3 working days of a month is about 0.46% lower than that for those filed on other days. Because the average failure rate in the sample is only about 19%, this translates to a 2.4% higher rate of rejection. The lower magnitude for this result is unsurprising since final approval is more likely to be driven by the quality (and novelty) of the underlying idea, rather than the quality of the drafting of the document itself.

Errors or sloppiness in the preparation of the application could cause delays in the review process. Thus, as a third measure of work quality, we examine the log of the duration (in days) between filing date and grant date for approved patent applications. Though the review length is not associated with any direct monetary costs at the USPTO, longer reviews are costly for two reasons. First of all, applications with longer reviews are more likely to receive multiple notices that require re-work, which in turn substantially increases cost. Moreover, a firm can file a patent infringement lawsuit only if the patent has been granted. Thus, delayed approval reduces the potential bargaining power of the firm in competitive situations. We expect review length of a patent application to be negatively related to the quality of the application draft, conditional on firm fixed effects that control for unobserved firm-spepcific heterogeneity relating to technology and quality of idea, and annual firm portfolio size that controls for within-firm innovation shocks. That is, all else equal, poorly drafted applications are likely to require additional work, leading to longer review duration.The results in columns 7 to 9 of Table 6 indicate that the review duration is indeed longer for month-end patents by 1.0 to 2.4% .

Figure 5 presents a variant of these results in a graphical format. Specifically, the figures present a predicted fractional polynomial fit of the residuals of the dependent variable (de-meaned of firm fixed effects) on working day of the month. It is clear from the figures that there is a discernible change in work quality as we move from the first working day of a month to the last. The magnitude of the differences are similar to that documented in Table 6. Together, consistent with our model predictions, these results strongly suggest that the work quality of the applications completed on the last few days of the month is distinctly lower.

To ensure the robustness of these results, we perform a series of tests on the statistical association between month-end clustering and work quality. First, we rerun specifications in Table 6 including firm-claims fixed effects. This is to examine if the work quality results are entirely a consequence of the sorting of more complex jobs to the month-end (as suggested by results in Section 6), or if completion close to deadline has a negative effective after conditioning flexibly on complexity. The results (panel A of Table A.5) are similar to Table 6, with somewhat larger magnitudes for most coefficients (suggesting that work quality is worse when controlling carefully for the number of claims). Thus, consistent with our model, the reduction in work quality holds even when comparing work output with the *same number of claims within the same firm.*

Second, we re-estimate models in Table 6 including firm-volume fixed effects. This is to check if the applications at the end of the month are poorer in quality mainly because of higher volumes or if there is an additional month-end effect induced by the deadlines themselves. To allow the effect of volume to be completely flexible, we define firm-volume as the number of patents filed by a firm in a given month. Hence, this compares differences in work quality measures between month-ends and other days *with the same volume of applications within the same firm.* The results (panel B of Appendix Table A.5) show that, while there is a decline in magnitude of the month-end effects, this decline is not very large (about 25%, 28%, and 35% decline in columns 3, 6 and 9 relative to Table 6), and the effects remain statistically significant. Thus, though higher volume accounts for a part of the month-end effect, a larger part of the negative effect on work quality appears to be an independent month-end effect.[34] We interpret this as suggesting that work completed close to deadlines suffer additionally from being accelerated, even relative to other workdays with equally high volumes of work.

Third, we repeat the analysis in Table 6 using data for years 2001 and 2002 only (providing us with at least 8 years of post-application data). This mitigates potential bias from right censoring

[34]Higher volume does lower work quality; in untabulated regressions with log daily volume as a control, we found higher volume is associated with significantly lower likelihood of approval, higher probability of receiving application incomplete notices, and longer review duration.

in the dependent variables.[35] The results presented in Appendix Table A.6 are similar to those based on 2001-04 data in Table 6, suggesting that right censoring does not significantly affect the baseline results.

Fourth, we interact the dummies of last working days of the month with firm size. Because routines and associated deadlines are likely to be stronger at larger firms, if the decline in quality of month-end applications is indeed related to deadlines, we should expect month-end effects for work quality to increase with firm size. In Appendix Table A.7, though statistically insignificant, the signs of the interaction terms indicate that the month-end effects are correlated with firm size in the manner we would expect if lower quality were related to routine-driven deadlines *and* routines are more prevalent at larger firms.

Finally, we re-ran Table 6 excluding the month of the fiscal year-end. One of our interviewees noted that in their firm, borderline ideas in the first three quarters of the year were re-reviewed in the last quarter. If annual targets had not been met, some of these marginal ideas were filed towards the end of the year. We believe the results in Section 6 rules out negative selection as an explanation for the work quality results, and that the inclusion of month fixed effects control for general month-specific variations in work quality. Nevertheless, we explicitly address this possibility by removing from the analysis the final months of the fiscal year. The baseline results were robust to this exclusion (Appendix Table A.8).

8 Linking the Model and Empirics: A Simulation

Our model presented earlier analyzes how deadline-related penalties influence work flows, task sorting and task quality. In this section, we examine if the model can reasonably replicate our empirical findings, particularly figures 2, 4 and 5. In so doing, we hope to provide both a reality check of the model as well as to illustrate the linkage between the model and the empirics.

We simulate data for 27 months of 25 days for 50 firms, with patent application jobs arriving at a rate of one per day. To abstract from initial transitional dynamics, we drop the first 3 months. We assume $a(x) = c(x) = x^2$, and $b(x) = x^3$, so that $t^* = x$ and $f(t^*) = t^*$. We set $\underline{T} = 25$ and $\overline{T} = 126$ so that the range of durations extends from one month to five months. We set γ=15 X 10^{-6} to yield a last day share of 10%. In the benchmark "No Deadline" regime there is no

[35] For example, application success is defined as 0 for applications where there is no evidence of approval as of November 2011. As noted in Section 7, this was why we restricted the data period to 2001-2004, allowing for 6 years of time for processing of applications. In general, we may not expect this right censoring to bias results for month-end applications relative to others; however, the finding in Table 5 that the duration of review is longer for month-end applications raises some concern that month-end applications coded as unsuccessful may in fact be still under review. Concerns relating to right censoring also apply to the duration of review. Right censoring is less of an issue for the 'Application Incomplete' variable because, conditional on censoring, the mean observed duration between application date and notification date is much lower (only about 2 months) than that between application date and grant date (about 3.25 years).

month-end penalty. In both the “No Deadline” and “Deadline” scenarios, we added a stochastic component to the filing date, to capture un-modeled factors (e.g., disruptions due to health issues or personal vacations). Specifically, we allowed for each application to finish up to 5 days before or after a scheduled date (including application initially rescheduled to be filed at month-end), with a 5% probability for each of the 10 alternative days.

The simulated data confirm the model propositions and replicate key empirical results (see Appendix Figure A.2 and Appendix Table A.9). In particular, we find that in the deadline regime: (i) there is significant clustering at the month-end; (ii) complexity is higher at the month-end; and (iii) the error rate is higher at the month-end, both conditional and unconditional on complexity. Furthermore, the magnitude of the effects are reasonably comparable to our empirical results. We checked and confirmed robustness of these results to a range of alternative parameter values consistent with the above assumptions, and to alternative penalty schemes such as making the penalty dependent on the degree of completion.

9 Discussion and Conclusion

Personal experiences and considerable anecdotal evidence suggest that deadlines influence work flows. Further, whether it be students re-arranging their study schedules or salespeople modifying their sales patterns, deadlines have been shown in the literature to change work flows in many contexts (e.g., Oyer 1998, Ariely and Wertenbroch 2002). This study adds to that body of empirical evidence by examining large scale, high frequency data on a single, relatively uniform process across thousands of firms over a long period of time, and documenting striking patterns of month-end, quarter-end and year-end clustering in patent applications. The magnitudes of these clustering patterns are significant, e.g., the daily share of corporate patents is about 17.7% higher in the last week of the month than in the first week of the month.

Several pieces of evidence strongly suggest a role for routine-related deadlines at large firms in generating these patterns. Applications by individual assignees do not show sharp spikes near month-ends, except before October 1 when any increases in USPTO fees take effect. In contrast, relative to the baseline pattern of patenting by individual assignees, the share of corporate patents in the last week is about 7.5% higher and about 3.8% lower in the first week. Further, regression results show that the propensity to file during month-ends is significantly higher among firms with high patent volume. Changes in fiscal year-ends are systematically correlated with changes in the clustering patterns of patent filings. Finally, interviews with patent attorneys and inventors indicate that routine-generated deadlines are the likely cause of the clustering patterns.

Prior studies have suggested that these changes in work flows may be caused by time-inconsistent preferences (or procrastination) when prioritizing various tasks to be completed before the deadline (e.g., O'Donaghue and Rabin 2001). We diverge from this line of thought by offering a simple model that does not invoke time-inconsistent preferences as an explanation for these clustering patterns. In particular, we develop a simple model of rational responses to discontinuous incentives around deadline thresholds. The intuition captured by the model is straightforward: faced with a discontinuous penalty at the deadline, agents respond by shifting completion of some work to the deadline. This explains the clustering of applications around deadlines. Our model also makes related predictions on task sorting and work quality. In particular, it predicts that in a likely scenario where deadline penalties do not decrease with task complexity, more complex tasks are more likely to be accelerated for completion. Further, our model predicts a higher error rate on tasks completed near deadlines if more complex tasks are accelerated and deadline penalties are low enough that at least some fraction of tasks are completed without rushing.[36]

Results from additional empirical analyses are consistent with these latter predictions. We find that applications filed near month-ends are of higher importance and complexity. The number of claims is 0.8-1.7% larger for applications filed near month-ends. Conditional on approval, the 5-year citation count for patents filed near month-ends is 0.4-1.5% higher as is the probability of renewal (0.2-0.3%). Thus, this suggests that the penalty structure is not decreasing in complexity.[37] This is not unreasonable since more complex filings are likely to be more valuable to the firm, and hence delays in such filings are likely to attract greater penalties. Beyond shedding light on the systematic association between task sorting and deadlines, these results also rule out the possibility that, in our context, deadline-related reductions in work quality are driven by agents choosing to rush tasks of lower importance. To our knowledge, ours is the first study to document clustering of more complex tasks close to deadlines; future work in other contexts could shed light on the robustness of this finding.

Finally, deadline-driven clustering may be interesting but is relevant for workplace productivity only if such clustering affects the quality of work performed around deadlines. We find that, consistent with our model, output quality is indeed lower around month-ends. The effect is stronger when using a shorter-term measure that likely reflects errors in the last stage of the filing process: applications filed closer to month-ends are about 12.5% more likely to be considered incomplete by the USPTO. The impact on the eventual success of the applications is smaller; they tend to be rejected 2.5% more than those filed during other days, and take 1.5% longer to review. The modest magnitudes for longer-term measures should be interpreted in keeping in mind that

[36]Our first and second predictions are also consistent with the time-inconsistent preferences model in O'Donaghue and Rabin (2001); the key novelty in their model is that it generates extra procrastination on more important tasks. Our prediction on work quality could potentially be obtained from their model by including a decline in work quality as an additional cost to accelerating work close to the deadline.

[37]To be precise, the results indicate that $f'(t^*) < \frac{2f(t^*)}{t^*}$.

the underlying idea quality or novelty, rather than documentation errors, is likely to play a much greater role in determining in eventual application approval and total review duration.

Our work is closely related to two recent papers that examine the effect of deadlines on work quality. Carpenter et al. (2012) find that an administrative tool that introduced deadlines at the U.S. FDA led to a piling up of decisions before deadlines, and these "just-before-deadline" approvals were linked to significantly higher rates of postmarket safety problems - the probability of a new black-box warning was 3.3 times greater and the probability of a safety-based withdrawal was 6.9 times greater for a drug approved two months before the deadline. Cadena et al. (2011) find significant clustering of credit placement at the end of the month in the Colombian bank they study. A randomized intervention (using prizes coupled with reminders and public recognition) reduced the clustering and evened out credit placement, reduced work-related stress levels, and increased job satisfaction. However, there was no significant impact on overall productivity (measured in the number of loans placed per period), or delinquency rates (which could be viewed as a measure of work quality). Our finding of modest effects on work quality, particularly on the longer term, is consistent with the results in Cadena et al. (2011). The contrast with Carpenter et al. (2012), who find strikingly large work quality effects, could be because of two important differences in the context: (i) they examine a government agency whereas we examine behavior of firms in a competitive marketplace; and (ii) they study one program that introduced new deadlines - these may have been suboptimal relative to the optimal duration required for a careful review. In our context, both the discipline imposed by market competition, as well as refinement of deadlines and associated incentives over time, could have led to a system with less severe distorting effects, particularly on long-term measures of work quality.

Together, our empirical results are consistent with our model predictions, suggesting that invoking time-inconsistent preferences is not necessary to explain the robust empirical regularities we document. However, two caveats are worth noting. First, as discussed earlier, unlike other models such as O'Donaghue and Rabin (2001), our model does not endogenize deadlines. It focuses on behavior, *once* deadlines are in place; it is agnostic about why deadlines exist. Thus, formal or informal penalties for missing deadlines may still be aimed at controlling for procrastinating behavior arising from time-inconsistent preferences. A second important caveat is that our data do not allow us to definitely answer whether the observed patterns are indeed driven by pre-planned acceleration (as in our model) or by ex-post acceleration prompted by procrastination. While it is possible that our results are driven by procrastination, it is less likely for two reasons. One, delays in our context can be very costly.[38] Thus, firms and their lawyers have strong incentives to avoid

[38]For instance, the importance of filing first is often discussed in the well-known Elisha Gray-Alexander Graham Bell controversy, though accounts of this controversy vary. In brief, on February 14, 1876, Alexander Graham Bell, through his attorney, filed for an important patent related to the telephone. Even though Elisha Gray had developed

delays. Moreover, no practitioner we interviewed with suggested controlling for procrastination as a motivation for quotas and deadlines; rather, these appeared to be motivated by planning and coordination objectives. Nonetheless, it is possible that procrastination has a role in these results. We leave a more thorough investigation of this question to future work.

References

[1] Akerlof, G.A. 1991. "Procrastination and obedience," *American Economic Review*, 81 (2): 1-19.

[2] Ariely, D. and Wertenbroch, K. 2002. "Procrastination, deadlines, and performance: Self-control by precommitment," *Psychological Science*, 13 (3): 219-224.

[3] Asch, B.J. 1990. "Navy recruiter productivity and the freeman plan," *DTIC Document*. RAND Corporation, Santa Monica, California.

[4] Becker, M.C. 2004. "Organizational routines: A review of the literature," *Industrial and Corporate Change*, 13 (4): 643-677.

[5] Cadena, X., Schoar, A., Cristea, A. and Delgado-Medrano, H.M. 2011. "Fighting procrastination in the workplace: An experiment," *NBER Working Paper* No 16944.

[6] Carpenter, D., Chattopadhyay, J., Moffitt, S. and Nall, C. 2012. "The complications of controlling agency time discretion: FDA review deadlines and postmarket drug safety," *American Journal of Political Science*, 56 (1): 98114.

[7] Clark, G. 1994. "Factory discipline," *Journal of Economic History*, 54 (1): 128-163.

[8] Courty, P. and Marschke, G. 1997. "Measuring government performance: Lessons from a federal job-training program," *American Economic Review*, 87 (2): 383-388.

[9] Hall, B.H., Jaffe, A.B. and Trajtenberg, M. 2001. "The NBER patent citation data file: Lessons, insights and methodological tools," *NBER Working Paper* No. 8498.

[10] Hall, B.H., Jaffe, A.B. and Trajtenberg, M. 2005. "Market value and patent citations," *RAND Journal of Economics*, 36 (1): 16-38.

[11] Hayek, F.A. 1952. *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. The University of Chicago Press: Chicago.

[12] Hounshell, D.A. 1975. "Elisha Gray and the telephone: On the disadvantages of being an expert," *Technology and Culture*, 16 (2): 133-161.

[13] Kaur, S., Kremer, M. and Mullainathan, S.. 2010. "Self-control and the development of work arrangements," *American Economic Review*, 100 (2): 1-8.

[14] Knott, A.M and B. McKelvey. 1999. "Nirvana efficiency: a comparative test of residual claims and routines," *Journal of Economic Behavior & Organization*, 38 (4): 365-383.

the same idea earlier, he did not file for a patent, and, on February 14, 1876, only filed a 'caveat,' a procedure available then to inventors to give an advance notice of their idea to the USPTO. Since Bell had already filed an application, Gray was required to file a formal application within the next three months. Given the high time and monetary costs of interference proceedings, Gray was advised not to file a formal patent application (Hounshell, 1975).

[15] Lai, R., D'Amour, A. and Fleming, L. 2009. "The careers and co-authorship networks of U.S. patent-holders, since 1975," *Working Paper*. Harvard Business School and Harvard Institute for Quantitative Social Science.

[16] Lanjouw, J.O. and Schankerman, M. 2004. "Patent quality and research productivity: Measuring innovation with multiple indicators," *Economic Journal*, 114: 441-465.

[17] Langlois, R.N. 1992. "Transaction-cost economics in real time," *Industrial and Corporate Change*, 1 (1): 99-127.

[18] March, J.G. and Olsen, J.P. 1989. *Rediscovering Institutions*. The Free Press: New York, NY.

[19] McWha, K. and Ryan, T., 2008. "Filing patent applications with the bare minimum of everything and associated pitfalls," *AIPLA Annual Meeting*, October 2008. Available at http://www.aipla.org/learningcenter /library/papers/am/AM08Materials/ Documents/McWha-paper.pdf.

[20] Merrill, S.A., Levin, R.A. and Myers, M.B., (Eds). 2004. *A Patent System for the 21st Century*. Committee on Intellectual Property Rights in the Knowledge-Based Economy, National Research Council, National Academies Press, Washinton D.C..

[21] Nelson, R.R. and Winter, S.G. 1982. *An Evolutionary Theory of Economic Change*. Belknap Press/Harvard University Press: Cambridge.

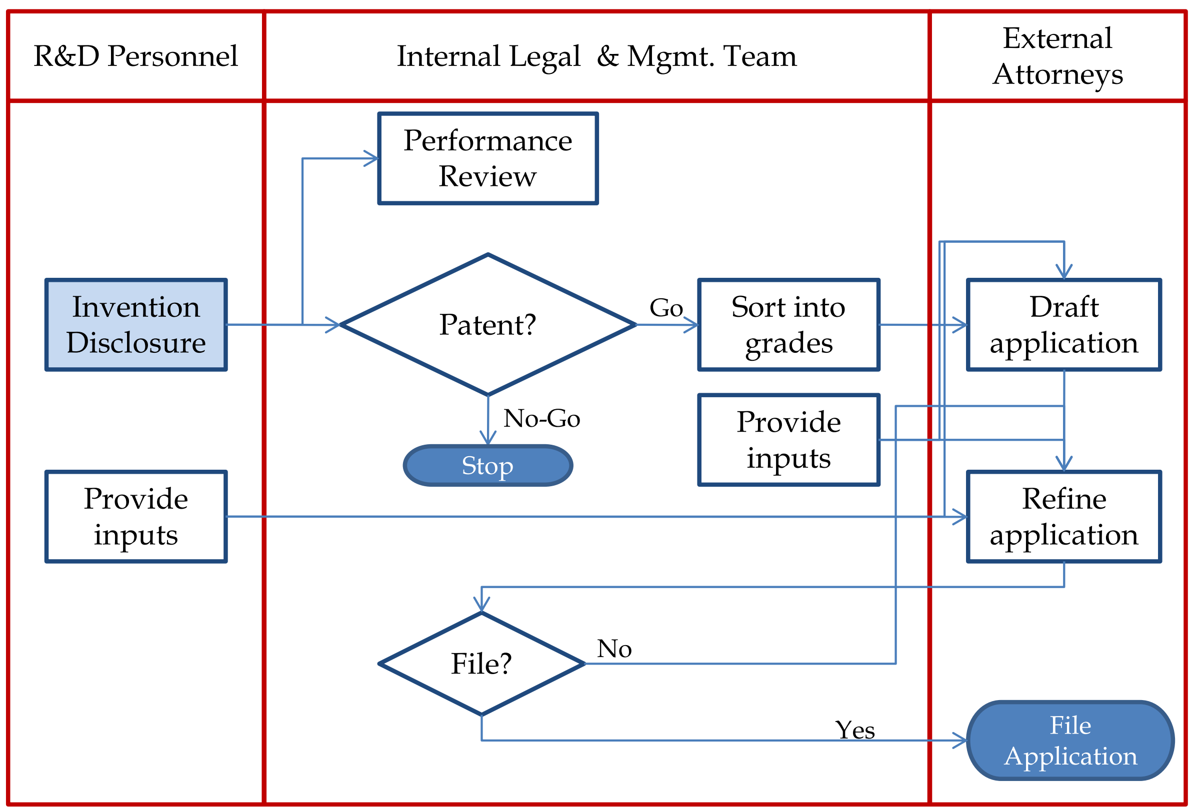
[22] O'Donaghue, T., and Rabin, M. 2001. "Choice and procrastination." *Quarterly Journal of Economics*, 116 (1): 121-160.

[23] Oyer, P. 1998. "Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality," *Quarterly Journal of Economics*, 113 (1): 149-185.

[24] Pakes, A. 1986. "Patents as options: some estimates of the value of holding European patent stocks," *Econometrica*, 54 (4): 755-784.

[25] Simon, H.A. 1947. *Administrative Behavior*. Macmillan: New York.

Figure 1: Patent Application Process at a Large, Patent-intensive Firm



Notes:
(1) Disclosure to filing time: 1 to 2 months
(2) Two deadlines per year for invention disclosure

Figure 2: Patent Filings Day Share of Year Total

Dotted line represents applications and solid line represents granted patents based on data for the period 2001-2009. Letters denote month-ends.

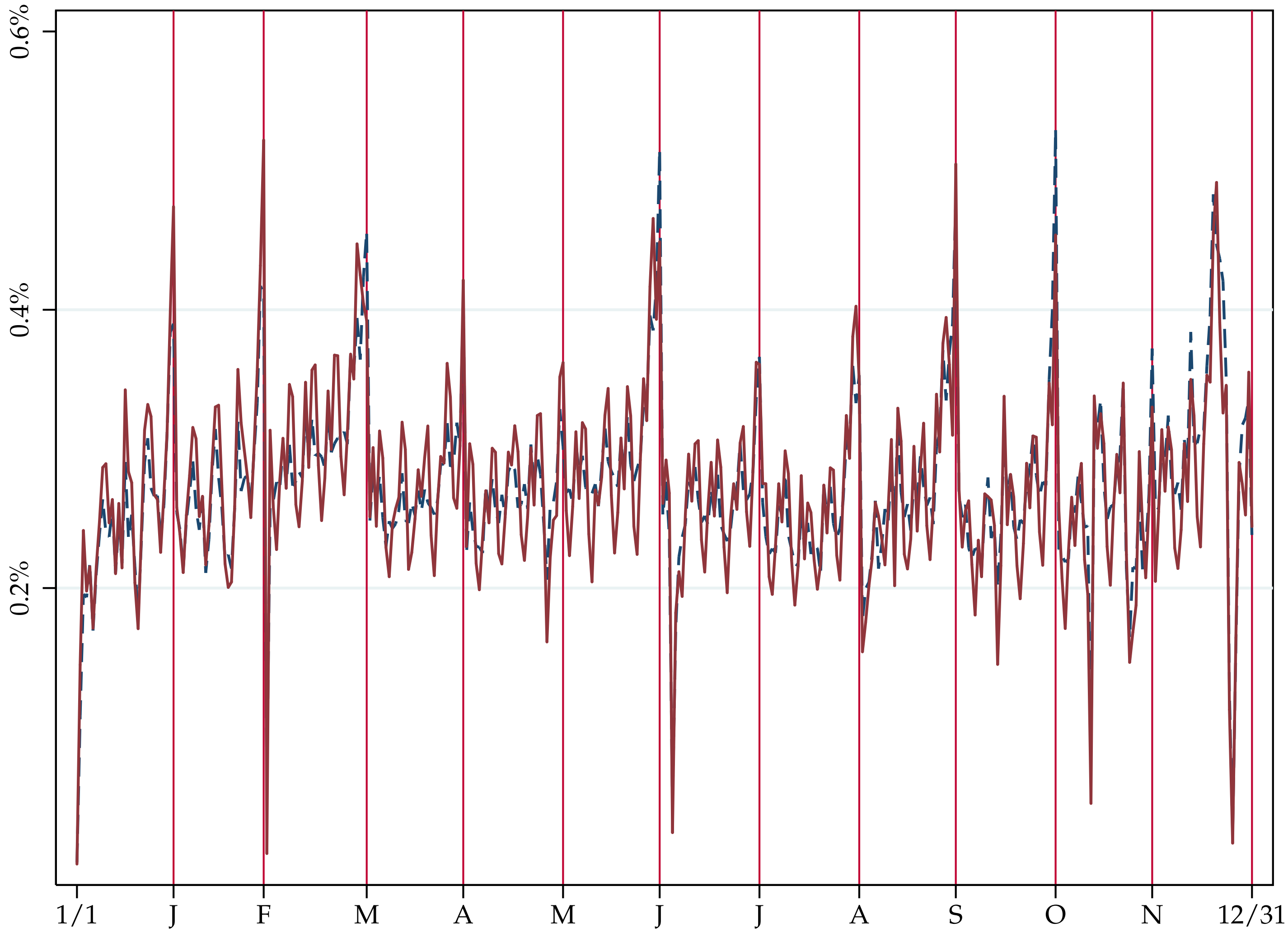


Figure 3: Corporate vs. Individual Assignees

Graph presents difference between day-share of filings for corporate and individual assignees, based on patent data for 1976-2009. Individual assignees are defined as those patents for whom there is no assignee (i.e., the inventors are the owners) in the USPTO data. Letters denote month-end.

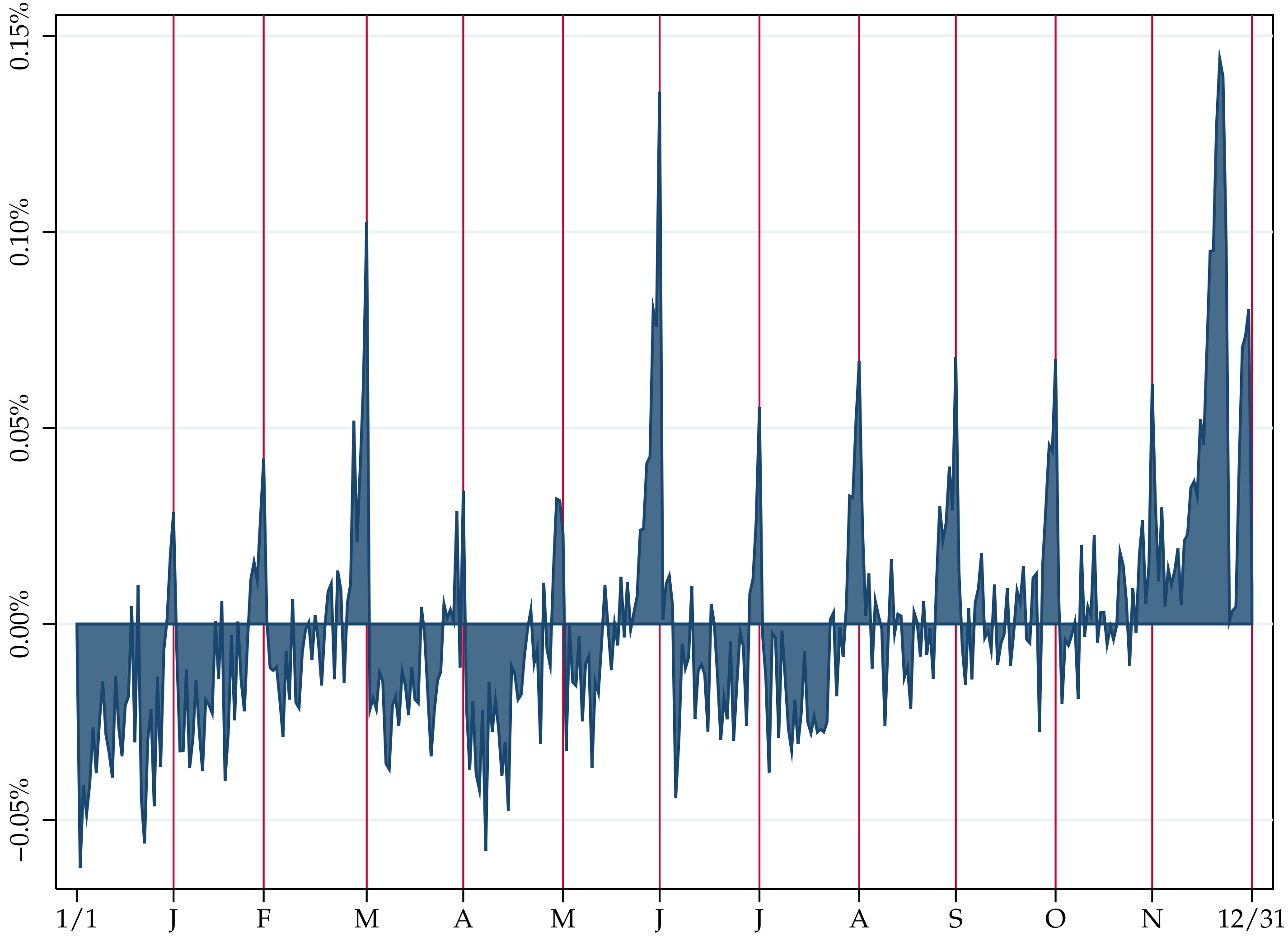


Figure 4: Month-ends and Application Complexity/Importance

In each of the figures, the line denotes predicted fractional polynomial fit; the area denotes 95% confidence interval; the horizontal axis indicates working day of the month. The dependent variable, from the first figure to the last, is log number of claims in the application, log number of citations to the patent within 5 years of application, and a dummy variable equal to 1 if fees for renewal of patent was paid 3.5 years after grant of the patent and 0 otherwise, respectively. For log claims, data on applications published in 2001-2010 are used. For log cites, data on granted patents in 1976 to 2004 are used, and for 3.5 year renewal dummy, patent data in 1976-2005 are used. All dependent variables are residuals de-meaned of firm, application year and grant year fixed effects.

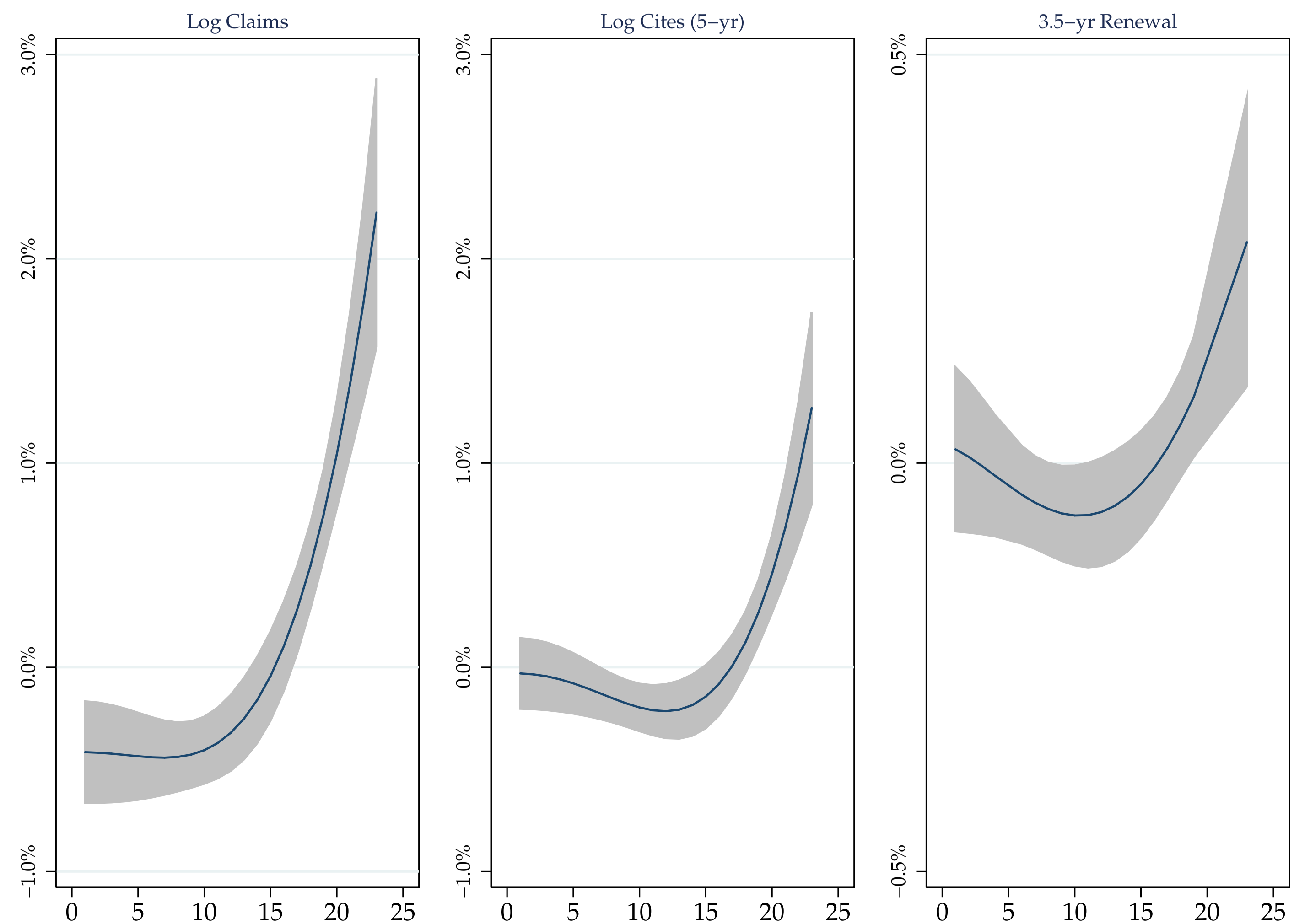


Figure 5: Month-ends and Application Quality

In each of the figures, the line denotes predicted fractional polynomial fit; the area denotes 95% confidence interval; the horizontal axis indicates working day of the month. The dependent variable, from the first figure to the last, is a dummy variable equal to 1 if the application received an 'Application Incomplete' notice from the USPTO and 0 otherwise; a dummy variable that is 1 if the application was approved by Nov 2011 and 0 otherwise; and log number of days between application date and grant date for successful patent applications, respectively. Only data on published applications in 2001-04 are used. All dependent variables are de-meaned of firm fixed effects.

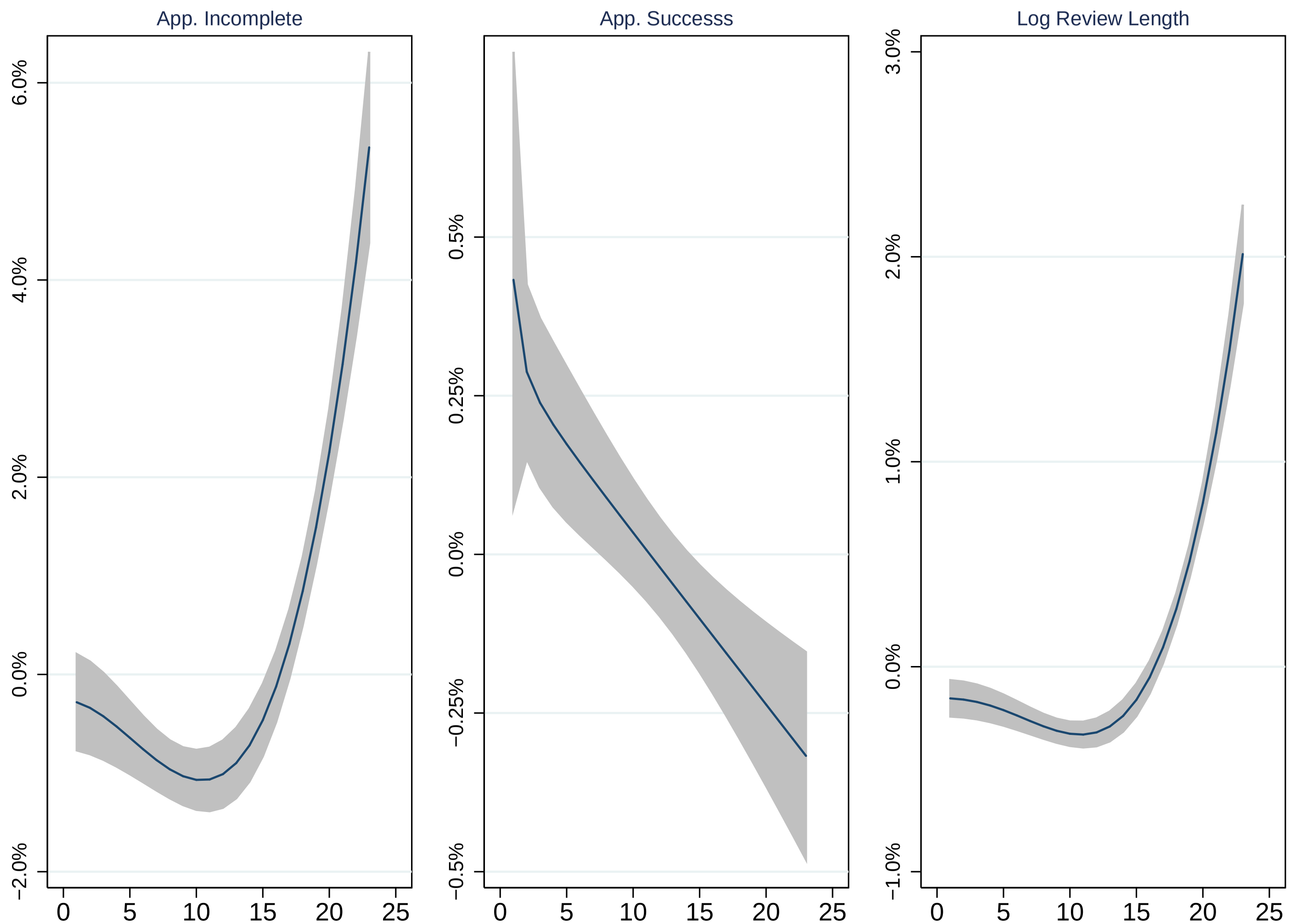


Table 1: Heterogeneity in Patenting Patterns: Individuals versus Corporates

The table presents coefficients from OLS regressions. The dependent variable is the difference between the day-share of annual patent applications for firms and the corresponding share for individual assignees. The independent variables are dummy variables corresponding to the period mentioned in the rows. These regressions use data on granted patents from 1976-2009. The day-share for a given day is computed as the ratio of the number of successful patents applied on that day to the total number of successful patents over the sample period.

	(1)	(2)	(3)
Dep. Var.	Difference in day share		
Days 1-7	-0.103**	-0.256**	-0.166**
	(0.025)	(0.039)	(0.039)
Days 8-15	-0.101**	-0.254**	-0.164**
	(0.019)	(0.035)	(0.035)
Days 16-23	-0.002	-0.155**	-0.066
	(0.036)	(0.046)	(0.046)
Days 24 and Beyond	0.205**		
	(0.032)		
Last Day of Month		0.389**	
		(0.108)	
Last 3 Days of Month			0.353**
			(0.058)
Constant		0.153**	0.064*
		(0.029)	(0.029)
R^2	0.170	0.217	0.251
N	366	366	366

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Table 2: Effect of Firm Size on End-of-Month Clustering – Fixed Effects OLS

Columns 1-3 use granted patents in 1976-2009 (N=2,878,229) while columns 4-6 use published applications filed in 2001-2010 (N=1,856,934). The dependent variable (Dn) in each column is a dummy variable that is 1 if the patent (or application) was applied on the last n working day of a month, and 0 otherwise. Assignee-year size is defined as the number of successful patents applied by an assignee during a given year.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	D1	D3	D5	D1	D3	D5
	(Patent data)			(Application data)		
Log Firm-Year Size	2.842**	5.533**	6.542**	2.662**	5.350**	6.001**
	(0.582)	(0.702)	(0.676)	(0.700)	(1.075)	(1.195)
R^2	0.003	0.003	0.003	0.003	0.003	0.005
Month Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Table 3: Monthly Volumes and Fiscal Year End

Each cell indicates the fraction of all granted patents applied in a given month for a given fiscal-year end. Thus, the first two cells in column 1 indicate that, of all patents of firms with fiscal years ending in January, 16.2% were filed in January and 6.1% were filed in February.

	Fiscal Year Month-End											
Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Jan	**16.2**	**5.9**	**6.5**	**6.9**	**6.4**	**6.7**	**7.2**	**6.6**	**6.7**	**7.9**	**5.9**	**6.7**
Feb	**6.1**	**22.1**	**7.7**	**7.4**	8.5	**7.2**	**7.2**	**9**	**7.4**	**7**	**7.0**	**6.8**
Mar	**6.4**	8.0	**14.1**	**9.7**	8.3	**9.4**	8.3	**8.1**	**9.1**	**8.7**	8.3	**8.5**
Apr	**7.3**	7.1	8.4	**18.7**	8.7	8.2	**7.2**	**7.6**	**7.5**	**9.6**	9.2	**7.8**
May	7.6	8.2	**7.5**	**5.8**	**13.7**	8.2	8.7	**7.4**	**8.8**	**7.9**	8.2	**8.2**
Jun	**7.0**	9.2	**9.2**	**6.9**	8.5	**14.6**	**10.2**	8.4	**9.3**	8.4	8.4	**9.6**
Jul	8.7	**5.7**	**7.4**	**5.6**	**6.8**	**7.5**	**11.6**	8.6	**7.6**	8.4	**6.6**	**7.3**
Aug	7.6	**5.5**	8.5	**6.7**	8.2	**6.9**	**7.3**	**18.3**	8.5	8.2	8.4	**7.9**
Sep	**6.7**	7.0	8.4	7.8	8.0	**7.4**	**7.2**	**6.9**	**12.6**	**7.9**	8.0	**8.4**
Oct	8.4	7.2	**7.9**	**9.7**	8.3	**7.9**	8.5	**6.4**	**7.9**	**12.3**	7.8	8.4
Nov	7.6	7.1	**6.8**	**6.5**	**7.2**	**7.7**	**7.3**	**6.1**	**6.8**	**6.8**	**13.2**	**7.9**
Dec	**10.5**	7.2	**7.6**	8.3	**7.7**	8.3	**9.4**	**6.7**	**7.7**	**6.9**	9.0	**12.5**
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Cells highlighted in grey represent maximum within the respective columns.

Numbers highlighted in bold font indicate they are different from 8.3% (i.e. 1/12) at or below the 5% level.

Table 4: Effect of Accounting-Year Shifts

This table presents coefficients from OLS regressions with firm-fixed effects. The first four columns (N=59,457) use data on granted patents from 1976-2009 while the next four (N=105,530) use the sample of published applications that were applied in 2001-2010. Data in all columns are based on the sample of 'switchers' - firms that change fiscal years. Among the dependent variables, D_o is a dummy that is 1 if a patent is applied in the month corresponding to firm's *old* fiscal-year end, and 0 otherwise; and D_n is a dummy that equals 1 if a patent is applied in the month corresponding to firm's *new* fiscal-year end, and 0 otherwise. The main independent variable of interest, 'Old Regime Dummy', is 1 before the firm switches to its new fiscal-year, and 0 thereafter.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	D_o	D_n	$D_n - D_o$	$D_n - D_o$	D_o	D_n	$D_n - D_o$	$D_n - D_o$
Old Regime Dummy	36.417**	-19.724**	-56.141**	-56.927**	19.986**	-2.502	-22.488*	-22.560*
	(6.896)	(5.780)	(11.140)	(11.508)	(6.036)	(6.364)	(10.669)	(10.648)
Log Firm-Year Size				-4.896				0.731
				(4.768)				(2.357)
R^2	0.110	0.410	0.119	0.119	0.136	0.387	0.254	0.254
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Table 5: Nature of Task-Sorting to Month-ends

This table presents coefficients from fixed-effects OLS regressions. The dependent variable in the first three columns (N=1,801,602) is log number of claims in the application; in columns 4 to 6 (N=2,633,488) it is log number of citations to the patent within 5 years of application; and in columns 7 to 9 (N=1,938,296) it is a dummy variable equal to 1 if fees for renewal of patent was paid 3.5 years after grant of the patent, and 0 otherwise. For log claims, data on applications published in 2001-2010 are used. For log cites, data on granted patents in 1976 to 2004 are used, and for 3.5 year renewal dummy, patent data in 1976-2005 are used.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Log(number of claims)			Log number of cites (5 years)			Renewal (3.5 years)		
Last W. Day of Month	1.735**			1.502**			0.282*		
	(0.316)			(0.302)			(0.125)		
Last 3 W. Days of Month		0.943**			0.612**			0.205*	
		(0.226)			(0.197)			(0.086)	
Last 5 W. Days of Month			0.795**			0.390*			0.216**
			(0.215)			(0.168)			(0.077)
Log Firm-Year Size	0.361	0.36	0.362	-3.643**	-3.643**	-3.642**	-2.158**	-2.159**	-2.159**
	(0.836)	(0.836)	(0.836)	(0.673)	(0.673)	(0.673)	(0.341)	(0.341)	(0.341)
R^2	0.338	0.338	0.338	0.384	0.384	0.384	0.324	0.324	0.324
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Application Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Grant Year Fixed Effects				YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 100.

Table 6: Work Flow Clustering and Application Work Quality

This table presents coefficients from fixed-effects OLS regressions. Only data on published applications in 2001-04 are used. The dependent variable in columns 1-3 (N=92,533) is a dummy variable that is 1 if the application received an 'Application Incomplete' notice from the USPTO, and 0 otherwise; in columns 4-6 (N=785,051) it is a dummy variable that equals 1 if the application was approved and 0 otherwise; and in columns 7 to 9 (N=629,355) it is log number of days between application date and grant date for successful patent applications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Application Incomplete			Application Approved			Log Review Duration		
Last W. Day of Month	45.903**			-1.441			24.315**		
	(10.735)			(2.239)			(3.596)		
Last 3 W. Days of Month		36.167**			-4.578**			14.698**	
		(7.063)			(1.497)			(2.526)	
Last 5 W. Days of Month			30.029**			-4.456**			10.166**
			(5.549)			(1.153)			(2.227)
Log Firm-Year Size	-3.257	-3.226	-3.377	-19.786**	-19.772**	-19.762**	-3.582	-3.581	-3.588
	-8.031	-8.035	-8.045	-3.474	-3.474	-3.475	-8.026	-8.017	-8.017
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** $p<0.01$, * $p<0.05$, + $p<0.1$; Coefficients and standard errors multiplied by 1000.

Online (Web) Appendix

I Properties of z^*

Setting inequality (5) to γ and solving for z^* yields

$$z^* = t^* \left[\frac{t^*\mu}{1+t^*\mu}\right] \tag{12}$$

where $\mu = \sqrt{\frac{2\gamma}{f(t^*)}}$.

Since the left hand side of (5) is increasing in z, z^* represents the upper bound beyond which it is not optimal to accelerate tasks of a given complexity.

I.A z^* is increasing in γ for any given t^*

Taking the partial derivative of z^* w.r.t. γ, we get $\frac{\partial z^*}{\partial \gamma} = \frac{\mu(t^*)^2}{2\gamma(1+t^*\mu)^2}$.

I.B $\frac{z^*}{t^*}$ is increasing in γ for any given t^*

It follows from equation (12) that $\frac{z^*}{t^*} = \left[\frac{t^*\mu}{1+t^*\mu}\right]$. Taking the partial derivative of this w.r.t. γ, we get $\frac{\partial(z^*/t^*)}{\partial\gamma} = \frac{t^*\mu}{2\gamma(1+t^*\mu)^2} > 0$.

I.C Signs of the second derivatives

We begin with $\left[\frac{z^*}{t^*}\right]$. Using the result from above for the first partial derivative w.r.t γ, and then taking its partial derivative w.r.t t^*, we get $\left[\frac{\partial^2(z^*/t^*)}{\partial t^*\partial\gamma}\right] = \frac{1}{2\gamma}\frac{\partial\phi}{\partial t^*}\frac{1-\phi}{1+\phi}$ where $\phi = t^*\mu$. Therefore, at very low values of γ, $(1-\phi) > 0$ so that the sign of the second derivative is the same as the sign of $\left[\frac{\partial\phi}{\partial t^*}\right]$. Similarly, for high enough γ, $(1-\phi) < 0$ so that the sign of the second derivative is the opposite of $\left[\frac{\partial\phi}{\partial t^*}\right]$. The result follows from the fact the partial derivative of $\left[\frac{z^*}{t^*}\right]$ w.r.t t^* has the same sign as $\left[\frac{\partial\phi}{\partial t^*}\right]$.

Turning to z^*, $z^* = t^*\left[\frac{z^*}{t^*}\right]$. Therefore, $\frac{\partial z^*}{\partial\gamma} = t^*\left[\frac{\partial(z^*/t^*)}{\partial\gamma}\right]$. Differentiating this w.r.t t^*, we get $t^*\left[\frac{\partial^2(z^*/t^*)}{\partial t^*\partial\gamma}\right] + \left[\frac{\partial(z^*/t^*)}{\partial\gamma}\right]$. Using the results from above, this can be simplified to $\frac{1}{2\gamma}\frac{\partial\phi}{\partial t^*}\frac{1-\phi}{1+\phi} + \frac{\phi}{2\gamma(1+\phi)^2}$, which can be further simplified to $\frac{1}{2\gamma(1+\phi)}\left[(1-\phi)\left[\frac{\partial\phi}{\partial t^*}\right] + \frac{\phi}{1+\phi}\right]$. As $\gamma \to 0$, this tends to $\frac{1}{2\gamma}\left[\frac{\partial\phi}{\partial t^*}\right]$. Now, $\frac{\partial z^*}{\partial t^*} = \frac{1}{2\gamma(1+\phi)}\left[(1-\phi)\frac{\partial\phi}{\partial t^*} + \frac{\phi}{1+\phi}\right]$ which has the same sign as $\left[\frac{\partial\phi}{\partial t^*}\right]$ for very low values of γ.

II Clustering and Complexity

First, consider the derivative of $\left[\frac{z^*}{t^*}\right]$ w.r.t t^*. The sign of the derivative depends on the sign on $t^*\left[\frac{\partial\mu}{\partial t^*}\right] + \mu$. Substituting $\mu = \sqrt{\frac{2\gamma}{f(t^*)}}$, and simplifying, we get $\sqrt{\frac{2\gamma}{f(t^*)}}\left[1 - \frac{t^*f'(t^*)}{2f(t^*)}\right]$. Thus, when $f'(t^*) < \frac{2f(t^*)}{t^*}$ for all t^*, $\left[\frac{z^*}{t^*}\right]$ is increasing in t^*.

Turning to z^*, $z^* = t^*\left[\frac{z^*}{t^*}\right]$. Taking the derivative w.r.t t^*, we get $\frac{z^*}{t^*} + t^*\frac{\partial(z^*/t^*)}{\partial t^*}$, which simplifies to $\frac{2t^*\mu + (t^*)^2\mu^2 + (t^*)^2\frac{\partial\mu}{\partial t^*}}{(1+t^*\mu)^2}$. This is positive if the numerator is positive. Applying this, substituting for μ and simplifying, we obtain the inequality: $f'(t^*) < 4\frac{f(t^*)}{t^*} + 2\sqrt{2\gamma f(t^*)}$, which is true if $f'(t^*) < 4\frac{f(t^*)}{t^*} + 2\sqrt{2\gamma\underline{F}}$. Similarly, the numerator is negative if $f'(t^*) > 4\frac{f(t^*)}{t^*} + 2\sqrt{2\gamma f(t^*)}$, which is true if $f'(t^*) > 4\frac{f(t^*)}{t^*} + 2\sqrt{2\gamma\overline{F}}$.

III Other Analyses

III.A $\overline{T} < D$ and $\underline{T} > D$

In the first instance, the duration for the most complex task is less than the time between deadlines. Therefore, only (z^*/t^*) and its properties is relevant in determining behavior. In the second case, the duration for the least complex task is more than the time between deadlines. Therefore, all tasks span across at least two deadlines. Further, in this case, all deadline behavior is determined by z^*, and not by (z^*/t^*). A summary of the model's predictions is presented in the table following this appendix.

III.B Impact of decreasing D

Decreasing D is equivalent to increasing the frequency of deadlines. Suppose, in the beginning, D_0 is so large that $\overline{T} < D_0$. Then, initially, there is no effect of a decrease in D on work flows, task sorting or task quality. As D declines enough so that $\overline{T} > D_1$, say, then under some conditions, D becomes relevant in determining deadline behavior. In particular, if γ is high enough that $z^* > D_1$ for tasks of some complexities, then decreasing D below D_1, to say D_2 increases the extent of task acceleration. This occurs since for some tasks, z^* in the first regime (when $D = D_1$) will be s.t. $D_2 < z^* < D_1$. While these tasks will only be partially accelerated when $D = D_1$, they will be completely accelerated irrespective of the instant of arrival when $D = D_2$. The impact on complexity will depend on the rate of increase of $f(t^*)$. If it is s.t. both (z^*/t^*) and z^* are increasing, then average complexity of accelerated tasks will increase as D decreases. If both are decreasing, then the average complexity of accelerated tasks will decrease. Finally, decreasing D when leads to a decrease in the mean error rate. This occurs because the error rate will be lower for tasks that are completely accelerated irrespective of the day of arrival.

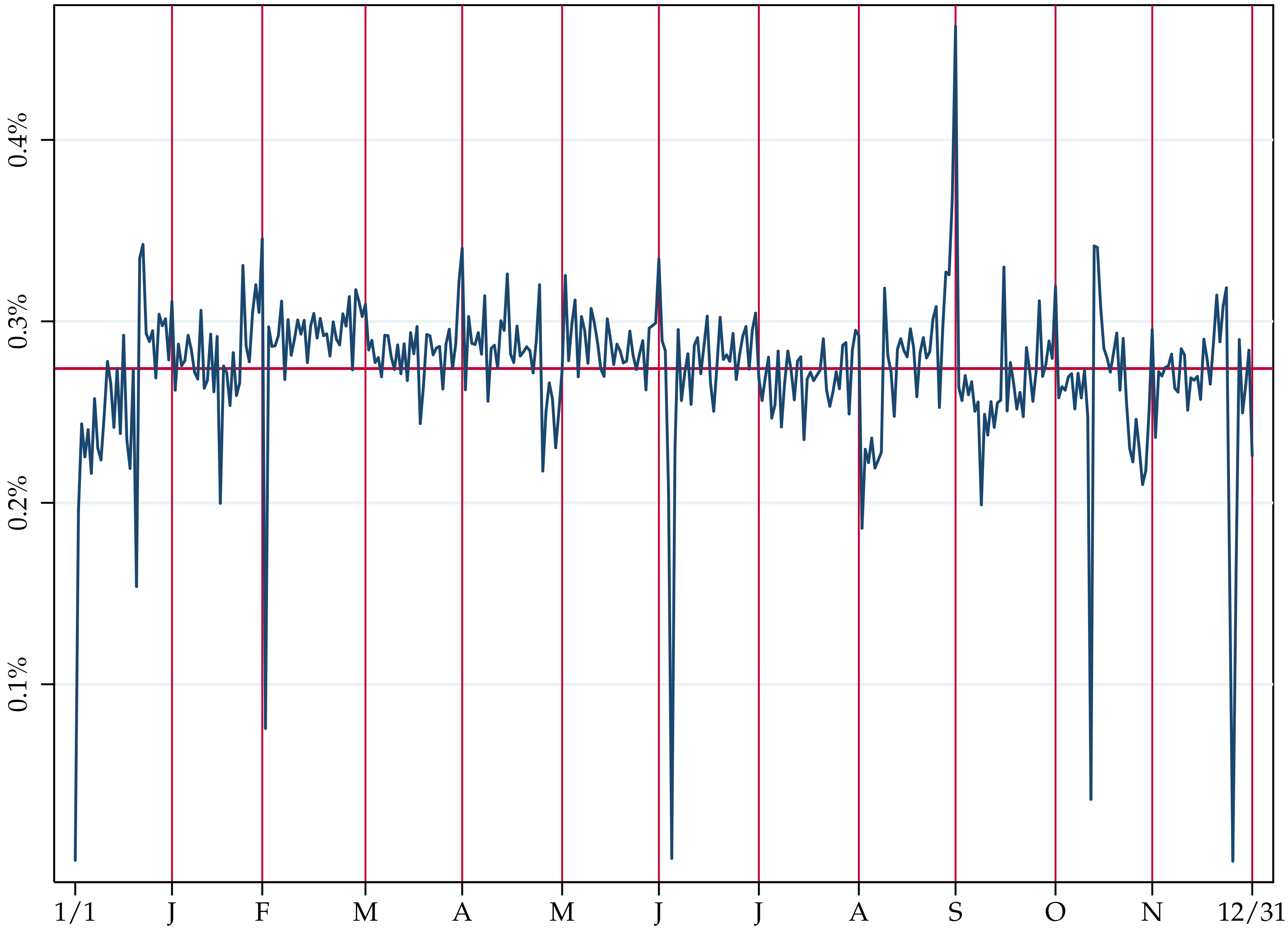
SUMMARY OF MODEL PREDICTIONS

	$\overline{T} < D$	$\underline{T} < D < \overline{T}$	$D < \underline{T}$
Work Flows			
Mass of tasks accelerated	$\frac{M}{2}\int_{\underline{T}}^{\overline{T}} \frac{z^*}{t^*}\tau(t^*)dt^*; z \epsilon [0, t^*)$	$\frac{M}{2}\int_{\underline{T}}^{D} \frac{z^*}{t^*}\tau(t^*)dt^* + \frac{M}{2}\int_{D}^{\overline{T}} \frac{z^*}{D}\tau(t^*)dt^*$	$\frac{M}{2}\int_{\underline{T}}^{\overline{T}} \frac{z^*}{D}\tau(t^*)dt^*; z \epsilon [0, D]$
Change in above with γ	←--Increases--→		
Acceleration irrespective of arrival instant	No task	Possibly some but not all tasks	Possibly all tasks at high enough γ
Task Sorting			
Mean complexity of accelerated tasks (τ_{acc}) relative to that of non-accelerated tasks ($\bar{\tau}$).	$f'(t^*) < \frac{2f(t^*)}{t^*}: \tau_{acc} > \bar{\tau}$ $f'(t^*) = \frac{2f(t^*)}{t^*}: \tau_{acc} = \bar{\tau}$ $f'(t^*) > \frac{2f(t^*)}{t^*}: \tau_{acc} < \bar{\tau}$	$f'(t^*) < \frac{2f(t^*)}{t^*}: \tau_{acc} > \bar{\tau}$ $f'(t^*) > \frac{4f(t^*)}{t^*} + \sqrt{8\gamma\overline{F}}: \tau_{acc} < \bar{\tau}$	$f'(t^*) < \frac{4f(t^*)}{t^*} + \sqrt{8\gamma\underline{F}}: \tau_{acc} > \bar{\tau}$ $f'(t^*) > \frac{4f(t^*)}{t^*} + \sqrt{8\gamma\overline{F}}: \tau_{acc} < \bar{\tau}$
Change in mean complexity of accelerated tasks with γ	←-----------For low values of γ, $\mid \tau_{acc} - \bar{\tau} \mid$ increases; for high values of γ, moves towards $\bar{\tau}$ as long as $\tau_{acc} \neq \bar{\tau}$------------→		
Task Quality			
Mean error rate conditional on complexity	←---Higher than baseline mean error rate---→		
Change in above with γ	Increases, unbounded	Increases, unbounded for $t^* \leq D$, bounded for $t^* > D$	Increases, bounded
Mean error rate unconditional on complexity	$f'(t^*) < \frac{2f(t^*)}{t^*}$: Higher, increasing in γ $f'(t^*) > \frac{2f(t^*)}{t^*}$: Ambiguous at low γ; higher than mean baseline rate above high enough γ	For γ low enough s.t. $z^* < D \; \forall \; z^*$, and $f'(t^*) < \frac{2f(t^*)}{t^*}$, higher than mean baseline rate	For γ low enough s.t. $z^* < D \; \forall \; z^*$, and $f'(t^*) < \frac{2f(t^*)}{t^*}$, higher than mean baseline rate

,

Appendix Figure A.1: Individual Assignees – Patent Filings Day Share of Year Total

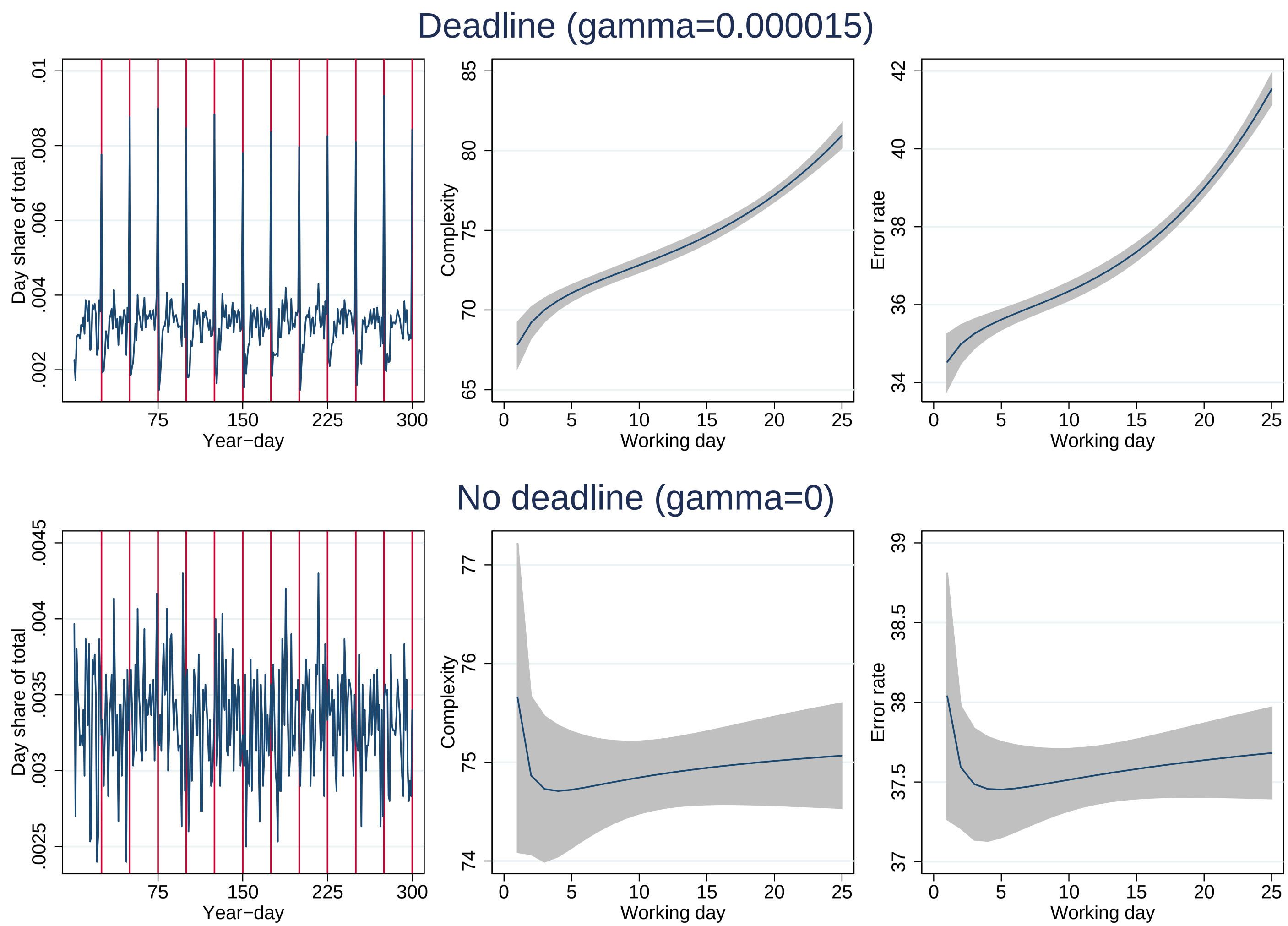
Solid line represents day share of granted patents for individual assignees (aggregating data for 1976-2009). Individual assignees are defined as patents for whom there is no assignee (i.e., the inventors are the owners) in the USPTO data. Letters denote month-ends.



,

Appendix Figure A.2: Simulation: Month-end Clustering, Complexity and Work Quality

The first row corresponds to a deadline punishment parameter $\gamma = 15X10^{-6}$ and the second row of figures correspond to simulation without a penalty for crossing the month-end threshold. In the second and third figures of each row, the line denotes predicted fractional polynomial fit; the area denotes 95% confidence interval; the horizontal axis indicates working day of the month. The dependent variable in the first figure in each row is the day share of total patents, in the second figure it is complexity (x in the model), and in the third figure it is error rate (as defined in the model). A month is defined as 25 working days.



Appendix Table A.1: Inter-firm Heterogeneity in Patenting Patterns: Small vs Large Corporates

The table presents coefficients from OLS regressions. The dependent variable is the difference between the day-share of annual patenting (or applications) for "Large" corporate assignees and the corresponding share for "Small" corporate assignees. "Large" assignees are defined as those with more than 100 patents filed over the data period. "Small" assignees are defined as those with just a single patent over the data period. The independent variables are dummy variables corresponding to the period indicated in the rows. The first three columns use data on granted patents from 1976-2009 while the next three use data on published applications that were applied between 2001 and 2010. For granted patents, the day-share for a given day (e.g. for Jun 30 for large assignees) is computed as the ratio of the number of successful patents applied on that day (6,864 for large assignees for Jun 30) to the total number of successful patents over the sample period (1,043,628 for large assignees, thus arriving at a share of 0.00658). Hence, each observation corresponds to one day of the year. The calculations for applications is similar.

	(1)	(2)	(3)	(4)	(5)	(6)
	(Granted Patents Data)			(Application Data)		
Days 1-7	-0.117**	-0.369**	-0.282**	-0.167**	-0.469**	-0.377**
	(0.031)	(0.053)	(0.054)	(0.033)	(0.059)	(0.054)
Days 8-15	-0.148**	-0.401**	-0.314**	-0.182**	-0.484**	-0.392**
	(0.024)	(0.049)	(0.050)	(0.033)	(0.058)	(0.053)
Days 16-23	-0.044	-0.296**	-0.209**	-0.037	-0.339**	-0.248**
	(0.039)	(0.059)	(0.059)	(0.039)	(0.062)	(0.057)
Days 24 and Beyond	0.314**			0.390**		
	(0.044)			(0.056)		
Last Day of Month		0.461**			0.659**	
		(0.137)			(0.214)	
Last 3 Days of Month			0.371**			0.449**
			(0.087)			(0.117)
Constant		0.253**	0.165**		0.302**	0.210**
		(0.043)	(0.044)		(0.048)	(0.042)
R^2	0.226	0.267	0.281	0.256	0.316	0.314
N	366	366	366	366	366	366

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.2: Lawyers and End-of-Month Clustering

The table presents coefficients from OLS regressions. These regressions use data on granted patents from 1995-2009 (N=1,792,838). Earlier data are not used due to lack of data on lawyers. The dependent variable (Dn) in each column is a dummy variable that is 1 if the patent was applied on the last n working day of a month, and 0 otherwise. Assignee-year size is defined as the number of successful patents applied by an assignee during a given year. Lawyer-firm-year-month volume is defined as the number of patents applied by a lawyer for a given firm in a given month.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	D1	D3	D5	D1	D3	D5
Log Firm-Year Size	2.620**	5.559**	6.473**	0.945*	2.297**	2.987**
	(0.438)	(0.580)	(0.601)	(0.373)	(0.551)	(0.555)
Log Lawyer-Firm-Year-Month Volume				5.930**	11.552**	12.342**
				(1.246)	(2.049)	(2.138)
R^2	0.117	0.109	0.107	0.118	0.110	0.108
Month Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Lawyer-Year Fixed Effects	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.3: Firm Size and End-of-Month Clustering – Including Patent Class and Inventor Fixed Effects

All columns use data on granted patents from 1976-2009. The dependent variable (Dn) in each column is a dummy variable that is 1 if the patent (or application) was filed on the last n working day of a month, and 0 otherwise. Assignee-year size is defined as the number of successful patents applied by an assignee during a given year. In columns 1-3, each observation corresponds to an inventor-patent observation (N=6,386,845), while in columns 4-6, each observation corresponds to a patent (N=2,878,229).

	(1)	(2)	(3)	(4)	(5)	(6)
Dep var.	D1	D3	D5	D1	D3	D5
	(Patent class fixed effects)			(Inventor fixed effects)		
Log Firm-Year Size	2.284**	4.560**	5.481**	1.464**	2.573**	2.670**
	(0.454)	(0.571)	(0.578)	(0.238)	(0.392)	(0.481)
R^2	0.004	0.004	0.004	0.302	0.299	0.298
Month Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Inventor Fixed Effects	NO	NO	NO	YES	YES	YES
Patent Class Fixed Effects	YES	YES	YES	NO	NO	NO

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.4: Testing for Sorting/Selection on Idea Quality – Using 2001-04 Data

This table presents coefficients from fixed-effects OLS regressions. Only published application data from 2001 to 2004 are used. The dependent variable in columns 1 to 3 is log number of claims in the application (N=785,051); in columns 4 to 6, it is log number of citations to the patent in next 5 years (N=642,252); and in columns 7 to 9, it is a dummy variable equal to 1 if fees for renewal of patent was paid 3.5 years after grant of the patent (N=224,262), and 0 otherwise. For 3.5 year renewal dummy, a restriction of grant year being before 2005 is imposed, which reduces the number of observations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Log(number of claims)			Log number of cites (5 years)			Renewal (3.5 years)		
Last W. Day of Month	2.763**			1.553**			0.132		
	(0.423)			(0.511)			(0.276)		
Last 3 W. Days of Month		1.338**			0.703*			0.090	
		(0.291)			(0.351)			(0.191)	
Last 5 W. Days of Month			1.140**			0.703*			0.116
			(0.362)			(0.310)			(0.164)
Log Firm-Year Size	0.332	0.334	0.332	-3.782**	-3.781**	-3.782**	-0.541	-0.541	-0.541
	(0.826)	(0.825)	(0.826)	-0.745	-0.745	-0.745	(0.367)	(0.367)	(0.367)
R^2	0.384	0.384	0.384	0.39	0.39	0.39	0.346	0.346	0.346
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Application Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Grant Year Fixed Effects				YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 100.

Appendix Table A.5: Clustering and Work Quality: Role of Volume and Complexity

This table presents coefficients from fixed-effects OLS regressions. Only data on published applications from 2001-04 are used. The dependent variable in columns 1-3 (N=92,533) is a dummy variable that is 1 if the application received an 'Application Incomplete' notice from the USPTO, and 0 otherwise; in columns 4-6 (N=785,051) it is a dummy variable that equals 1 if the application was approved and, 0 otherwise; and in columns 7 to 9 (N=629,355) it is log number of days between application date and grant date for successful patent applications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Application Incomplete			Application Approved			Log Review Duration		
	X=1	X=3	X=5	X=1	X=3	X=5	X=1	X=3	X=5
PANEL A: WITH FIRM-CLAIMS FIXED EFFECTS									
Last X W. Days of Month	50.458**	38.708**	34.335**	-3.423	-5.524**	-5.576**	21.022**	13.847**	10.160**
	(19.066)	(12.887)	(10.085)	(2.961)	(2.005)	(1.576)	(4.786)	(3.331)	(3.017)
Log Firm-Year Size	-18.187	-18.078	-18.483	-18.939**	-18.924**	-18.908**	-11.7	-11.733	-11.75
	(17.723)	(17.749)	(17.795)	(5.612)	(5.612)	(5.612)	(13.786)	(13.770)	(13.768)
R^2	0.665	0.665	0.665	0.501	0.501	0.501	0.502	0.502	0.502
PANEL B: WITH FIRM-VOLUME FIXED EFFECTS									
Last X W. Days of Month	31.503**	26.542**	22.514**	2.275	-2.735+	-3.202*	15.789**	7.553**	4.786*
	(8.091)	(6.074)	(4.764)	(2.221)	(1.622)	(1.278)	(3.239)	(2.222)	(1.934)
Log Number of Claims	19.564**	19.624**	19.681**	6.967**	6.985**	6.986**	100.836**	100.863**	100.875**
	(4.904)	(4.909)	(4.904)	(1.625)	(1.624)	(1.625)	(2.440)	(2.439)	(2.439)
Log Firm-Year Size	-2.259	-2.172	-2.323	-16.323**	-16.388**	-16.394**	-8	-8.019	-8.073
	(8.198)	(8.197)	(8.202)	(3.164)	(3.161)	(3.162)	(6.329)	(6.331)	(6.342)
R^2	0.403	0.403	0.403	0.338	0.338	0.338	0.336	0.336	0.336
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.6: Clustering and Work Quality – Using only 2001-02 Data to Reduce Right Censoring

This table presents coefficients from fixed-effects OLS regressions. Only data on published applications from 2001-02 are used, increasing the minimum post-application time frame to 8 years to reduce concerns from right censoring of data. The dependent variable in columns 1-3 (N=43,638) is a dummy variable that is 1 if the application received an 'Application Incomplete' notice from the USPTO, and 0 otherwise; in columns 4-6 (N=360,094) it is a dummy variable that equals 1 if the application was approved, and 0 otherwise; and in columns 7 to 9 (N=304,331) it is log number of days between application date and grant date for successful patent applications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Application Incomplete			Application Approved			Log Review Duration		
Last W. Day of Month	45.305**			-1.626			21.394**		
	(10.694)			(2.237)			(3.511)		
Last 3 W. Days of Month		35.907**			-4.672**			13.259**	
		(7.050)			(1.497)			(2.437)	
Last 5 W. Days of Month			29.775**			-4.549**			8.947**
			(5.524)			(1.151)			(2.179)
Log Number of claims	20.703**	20.784**	20.784**	6.507**	6.523**	6.525**	100.054**	100.084**	100.100**
	(4.950)	(4.963)	(4.932)	(1.586)	(1.585)	(1.586)	(2.792)	(2.790)	(2.801)
Log Firm-Year Size	-3.636	-3.609	-3.762	-19.813**	-19.799**	-19.789**	-3.795	-3.796	-3.801
	(8.002)	(8.006)	(8.015)	(3.477)	(3.478)	(3.478)	(7.511)	(7.503)	(7.503)
R^2	0.377	0.377	0.377	0.314	0.314	0.314	0.31	0.31	0.31
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.7: Sources of Reduced Work Quality – Interaction with Firm Size

This table presents coefficients from fixed-effects OLS regressions. Only data on published applications from 2001-04 are used. In columns 1-3 (N=92,533), the dependent variable is 1 if the application received an 'Application Incomplete' notice from the USPTO, and 0 otherwise; in columns 4-6 (N=785,051) it is a dummy variable that is 1 if the application was approved, and 0 otherwise; and in columns 7 to 9 (N=629,355) it is log number of days between application date and grant date for successful patent applications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Application Incomplete			Application Approved			Log Review Duration		
Last W. Day of Month	7.778			6.876			19.311*		
	(5.263)			(21.507)			(8.218)		
Last 3 W. Days of Month		-0.833			16.84			4.368	
		(3.477)			(14.617)			(5.356)	
Last 5 W. Days of Month			-0.571			10.472			-2.725
			(2.810)			(11.132)			(4.549)
Log Number of Claims	6.508**	6.525**	6.529**	20.701**	20.756**	20.751**	100.054**	100.077**	100.086**
	(1.586)	(1.585)	(1.586)	(4.951)	(4.964)	(4.933)	(2.792)	(2.789)	(2.801)
Log Firm-Year Size	-19.826**	-19.808**	-19.798**	-3.513	-3.571	-3.736	-3.791	-3.775	-3.774
	(3.478)	(3.478)	(3.478)	(8.003)	(8.006)	(8.019)	(7.509)	(7.501)	(7.502)
Log Size X Last W. Day of Month	**-1.519**$^{+}$			**8.400**			**0.334**		
	(0.877)			**(5.253)**			**(1.460)**		
Log Size X Last 3 W. Day of Month		**-0.628**			**4.168**			**1.441**	
		(0.599)			**(3.685)**			**(0.979)**	
Log Size X Last 5 W. Day of Month			**-0.658**			**4.256**			**1.913***
			(0.462)			**(2.709)**			**(0.848)**
R^2	0.314	0.314	0.314	0.377	0.377	0.377	0.31	0.31	0.31
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors clustered by firm in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.8: Sources of Reduced Work Quality – Excluding Last Month Fiscal Year

This table presents coefficients from fixed-effects OLS regressions. Only data on published applications from 2001-04 are used. The dependent variable in columns 1-3 (N=83,785) is a dummy variable that is 1 if the application received an 'Application Incomplete' notice from the USPTO, and 0 otherwise; in columns 4-6 (N=700,709) it is a dummy variable that equals 1 if the application was approved, and 0 otherwise; and in columns 7 to 9 (N=562,694) it is log number of days between application date and grant date for successful patent applications.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Application Incomplete			Application Approved			Log Review Duration		
Last W. Day of Month	36.803**			-2.729			22.147**		
	(10.599)			(2.422)			(3.564)		
Last 3 W. Days of Month		26.843**			-5.439**			14.515**	
		(6.668)			(1.665)			(2.569)	
Last 5 W. Days of Month			21.646**			-5.098**			8.867**
			(5.315)			(1.278)			(2.243)
Log Number of Claims	22.624**	22.670**	22.678**	6.472**	6.486**	6.489**	99.770**	99.795**	99.816**
	(5.095)	(5.104)	(5.080)	(1.555)	(1.554)	(1.555)	(2.873)	(2.871)	(2.883)
Log Assignee-Year Size	-6.746	-6.702	-6.803	-19.684**	-19.668**	-19.658**	-3.109	-3.122	-3.119
	(8.440)	(8.441)	(8.447)	(3.529)	(3.530)	(3.530)	(7.516)	(7.511)	(7.507)
R^2	0.386	0.386	0.386	0.32	0.32	0.32	0.314	0.314	0.314
Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses; ** p<0.01, * p<0.05, + p<0.1; Coefficients and standard errors multiplied by 1000.

Appendix Table A.9: Simulation Regressions – Day Share, Complexity and Error Rate

Columns 1-4 correspond to simulation without a penalty for crossing the month-end threshold, and columns 5-8 correspond to a deadline punishment parameter $\gamma = 15X10^{-6}$. The dependent variable in column 1 and 5 is the day share of total number of patent applications. The dependent variable in columns 2 and 5 is the complexity parameter (x in the model). The dependent variable in columns 3, 4, 7 and 8 is the error rate. Number of observations is 300 (corresponding to 300 working days of the every simulated year) in columns 1 and 5, and 30,000 for the rest of the columns (one patent a day for 24 months of 25 days for 50 firms).

	No Deadline ($\gamma = 0$)				Deadline ($\gamma = 15X10^{-6}$)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Day Share	Complexity	Error rate	Error rate	Day Share	Complexity	Error rate	Error rate
Last W. Day of Month	0.0000	0.7070	0.4070	0.0559	0.00530**	12.75**	8.003**	1.603**
	(0.0001)	(0.6370)	(0.3320)	(0.0612)	(0.0001)	(0.5260)	(0.2930)	(0.0506)
Constant	0.00333**	68.10**	34.27**	12.92**	0.00312**	67.17**	33.80**	12.90**
	(0.0000)	(0.2960)	(0.1450)	(0.1620)	(0.0000)	(0.2890)	(0.1450)	(0.1660)
R^2	0.000	0.023	0.022	0.974	0.794	0.038	0.045	0.973
Complexity FE	No	No	No	Yes	No	No	No	Yes

Notes: Robust standard errors in parentheses; ** $p<0.01$, * $p<0.05$, + $p<0.1$; Coefficients and standard errors multiplied by 1000.