The Failed Promise of User Fees: Empirical Evidence from the United States Patent and Trademark Office

11 J. Empirical Legal Stud. ___ (*forthcoming* 2014)

By MICHAEL D. FRAKES AND MELISSA F. WASSERMAN*

In an attempt to shed light on the impact of user-fee financing structures on the behavior of administrative agencies, we explore the relationship between the funding structure of the Patent and Trademark Office (PTO) and its examination practices. We suggest that the PTO's reliance on prior grantees to subsidize current applicants exposes the Agency to a risk that its obligatory costs will surpass incoming fee collections. When such risks materialize, we hypothesize, and thereafter document, that the PTO will restore financial balance by extending preferential examination treatment—i.e., higher granting propensities and/or shorter wait times—to some technologies over others.

I. INTRODUCTION

User fees have become ubiquitous. In the past thirty years, Congress has increasingly relied upon fees paid by users, rather than general tax revenue, to fund administrative agencies. Today, agencies ranging from the Food and Drug Administration, the Patent and Trademark Office, the Securities and Exchange Commission, to the Federal Trade Commission are funded, to a significant extent, through fees paid by their constituents. The rise of user-fee financing of administrative agencies has coincided with a growing literature on user fees themselves. To date, this literature has largely focused on the impact of user fees on associated user behaviors (de Rassenfosse and van Pottelsberghe de la Potterie 2013). For example, some scholars have noted that user fees may curb moral hazard on the part of agency beneficiaries (Griffin 1987; Gillette and Hopkins 1987), while others have drawn upon fee variations to estimate the price elasticity of demand for agency services (de Rassenfosse and van Pottelsberghe de la Potterie 2007; Moser 2007). In contrast, to the best of our knowledge, scholars have paid significantly less attention to the relationship between user fees and the behaviors of the agencies themselves; largely overlooking distortions that user fees, and particularly the structure of user fees, may place on agency behavior.[1] This oversight is of potential significance to the extent that the fees employed by agencies cause them to engage, at times, in undesirable practices.

In a partial attempt to fill this gap, this paper explores the relationship between agency fee structure and agency behavior in the context of the U.S. Patent and Trademark Office (PTO or Agency), a fully user-fee-funded body whose primary

---

[1] One exception is the Food and Drug Administration, as several studies have explored the impact of the Agency's user-fee funded review of new drug products on its behavior. Carpenter, Zucker, and Avorn (2008), for example, finds that medications approved immediately before user-fee imposed deadlines were more likely to exhibit safety problems during clinical use than medications approved at other times. Similarly, Olson (2004) finds that reduction in agency review times of new drug approvals cannot be explained by the increase in agency resources from user-fees.

task is examining patent applications.[2] We evaluate the impact of the PTO's user fee system in two steps. First, we suggest that the PTO's historical fee schedule exposes the Agency to a fundamental risk that its obligatory costs will come to surpass its incoming fee collections. Because the examination fees assessed on applicants cover less than one third of the cost to review applications, the PTO must subsidize patent examination by assessing fees on other activities—largely, the issuance of patents, which is associated with an issuance fee, and the renewal of patents, which is associated with a maintenance fee. Together, these "post-allowance" fees comprise over half of the PTO's budget. We contend that this cross-subsidized structure threatens the financial stability of the PTO insofar as it imposes a risk that the growth rate in patent grantees from which the PTO derives post-allowance fees lags behind the growth rate in incoming applications (among other risks). As such, we suggest that the PTO's fee schedule frustrates one of the very rationales that has, in part, driven the broad-based shift toward financing agencies through user fees—i.e., the ability to better accommodate unexpected demand shocks relative to fixed congressional budgets.[3]

Second, and at the heart of our analysis, we theorize the steps that the PTO may take in instances of binding budget constraints in order to restore financial balance while nonetheless satisfying as many of its obligations as it can. These steps themselves are facilitated by additional features of the PTO's fee and cost structures. In broad terms, we contend that a budget constrained PTO will distort its examination practices in an effort to increase the average fee income generated per application processed and/or to decrease the average examination costs incurred per application processed. Importantly, we propose that the PTO

---

[2] There is an established and growing literature bearing on the relationship between patent office fees and user behavior, including both positive studies, such determining the elasticity of renewal fees on the share of patents maintained (Pakes 1986; Schankerman and Pakes 1986) and normative analyses, such as utilizing PTO fees to diminish patent troll litigation (Magliocca 2007; Bessen and Love 2013). For a survey of this literature see de Rassenfosse and van Pottelsberghe de la Potterie (2013).

[3] A number of commentators have noted that a key benefit of user fee structures comes in their revenue-raising abilities and in their ability to ensure enhanced self sufficiency of agencies (Sepehri and Chernomas 2001; Gillette 1988).

achieves this financial rebalancing through the extension of preferential examination treatment—i.e., higher granting propensities and/or shorter wait times—to some technologies over others, a response that is potentially concerning to the extent that it stems solely from particularities of the agency's fee structure and not from legitimate social interests in intervening in certain industries.

The PTO may be able to increase average fee income per application examined by taking advantage of the fact that the majority of its fees are collected *only in the event that it grants patents*. As a result of this back-ended fee structure, the Agency may elect to grant patent applications at a higher rate to generate additional revenues. That is, a PTO facing a binding budget constraint may issue a patent on an application that it would have declined to grant when it was financially unconstrained. We explored this initial possibility in greater depth in our previous research (Frakes and Wasserman 2013). Drawing on the fact that certain types of patents are associated with higher fees relative to other types, this prior research predicts, and empirically supports, that a financially constrained PTO will target its inflationary granting proclivities on those technologies / applicant types that tend to garner higher post-allowance fees.

In the present analysis, we focus on the second manner in which a budget-constrained PTO may distort its examination practices—i.e., decreasing the wait times of certain application types in an effort to decrease the Agency's average cost per application processed. Under this approach, the PTO may extend preferential treatment in the Agency's substantial examination queue to those technologies that are the least complex and that thus cost the Agency the least to examine relative to those that carry the highest examination-processing costs. By reducing average processing costs through this examination prioritization, the Agency may increase the number of applications that it will be able to process using its limited incoming funds.

To be sure, each of these distortionary examination practices carries undesirable consequences. A grant-biased PTO is likely issuing patents on inventions that were either already known or represent only a trivial advancement over the existing scientific knowledge, thus subjecting society to unnecessary monopoly prices.[4] Moreover, the PTO's preferential examination-queuing treatment of certain technologies may distort, among other things, the allocation of resources across industries in society. Leaving some technologies with longer examination wait times relative to other technologies is no insignificant by-product of the PTO's funding process. After all, examination delays in general are associated with their own harmful consequences, including postponing the deployment of valuable inventions to the marketplace and increasing the uncertainty surrounding the rights of potential patents—the latter of which may limit a company's ability to license or otherwise engage in socially valuable activity.[5]

Finally, we theorize that certain congressional budgetary practices may bear on the PTO's choice as between these two policy levels—that is, inflationary granting practices versus examination prioritization. More particularly, we predict that the PTO would rely more heavily on the latter tool in the pre-2004 period, at which time Congress regularly diverted fee revenues from the Agency. We predict that the Agency would switch gears and prefer distortionary granting practices as a revenue-balancing tool following 2004, at which time the Agency was given greater ability to retain its anticipated revenue stream.

We test the predictions of the model using a rich database of previously-unavailable patent processing measures, which we amassed through the filing of Freedom of Information Act Requests to the PTO. We find evidence consistent

---

[4] The patent literature has long acknowledged the negative social welfare consequences of over-granting patents; moreover, there is widespread belief that the PTO is routinely issuing invalid patents (Lemley 2001; Lichtman and Lemley 2007; FTC 2003; Thomas 2001; Merges 1999; Lemley and Sampat 2008; Jaffe and Lerner 2004).

[5] The PTO itself has identified growing patent pendencies and the Agency's growing backlog of patent applications as its single "biggest challenge" (PTO 2008).

with the above predictions and suggesting that upon the occurrence of a negative financial shock to the PTO's financial health, the PTO begins to extend preferential examination-queuing treatment to those technologies that cost the Agency the least to examine. These results supplement Frakes and Wasserman (2013), which had found an increase in the grant rates of high-fee-generating patent types and technologies (relative to their low-fee-generating counterparts) in response to the same financial shocks. Moreover, in the present analysis, we find that the PTO's initial inclination to target low-cost examination prioritization over inflationary granting practices as a mechanism to deal with binding budget constraints appears to reverse course—that is, the PTO becomes more heavily reliant on over-granting as the relevant mechanism—upon the shift in congressional fee diversion practices in 2004.

Ultimately, this analysis suggests that the mere manner in which PTO fees are structured carries a potentially significant bearing on two aspects of patent policy that have received considerable attention in recent decades—that is (1) possible over-granting of patents and the associated dilution of patent quality and (2) the massive backlog of patent applications and the harms associated with the corresponding examination delays.

This paper proceeds as follows. Section II provides a background on the PTO's funding process and on the financial risks posed by this process. Section III theorizes how negative shocks to the PTO's financial health induce various distortions in the PTO's examination practices. Section IV describes the dataset and methodology utilized. Section V presents the results of our empirical analysis. Finally, Section VI discusses the implications of our findings.

II. BACKGROUND

The PTO became almost entirely user-fee financed in 1991.[6] At this time, however, the Agency was not granted the ability to set the levels of its fees nor was it given the right to automatically spend its fee collections. Instead, the PTO must receive congressional approval through annual appropriations to utilize its fee revenue. Historically, the PTO's appropriated budget—i.e., the amount it has been authorized to spend—has been set at or below its anticipated fee collections.[7] When the Agency's fee revenue falls below its appropriated budget, Congress does not provide the Agency with the difference. If the PTO's fee income surpasses its allocated budget, the Agency has typically not had access to these excess fees.[8] The model of PTO decision-making outlined in Section III initially approaches this structure in more simplistic terms and assumes that the Agency's operational funds derive purely from the fees it collects. It then relaxes this assumption to more accurately model some aspects of this nuanced interaction with Congress.

Since 1991, roughly 30% of the PTO's patent operating budget has been garnered through filing, search, and examination fees (collectively, "examination fees") (PTO 2012a, 1994). As such, while patent processing comprises the vast majority of the PTO's operational expenses (PTO 2012b), examination fees fail to cover one third of the cost incurred by the PTO to review applications. The PTO covers this funding gap through the imposition of a fee upon the issuance of a patent and a set of maintenance fees assessed over time in order for grantees to maintain the enforceability of their already issued patents. While the expenses to the PTO associated with issuing and maintaining a patent are minimal, these

---

[6] Omnibus Budget Reconciliation Act of 1990, Pub. L. No. 101-508, § 10101, 104 Stat. 1388, 1388-391 (1990).

[7] In other words, as discussed briefly in Section III, there are times when Congress has not allowed the PTO to spend all of its fee collections in a given year.

[8] Frakes and Wasserman (2013) provide a richer account (with corresponding support) of this budgetary process.

"post-allowance" fees comprise over 50% of the Agency's fee collections (PTO 2012a, 1994). Traditionally, this fee schedule has been justified by the belief that 1) low examination fees subsidize innovative activities and 2) post-allowance fees, especially maintenance fees, effectively shorten the lifetime of a patent, thereby reducing associated monopoly costs.[9] We do not dispute the merits of these arguments. Rather, we contend that this fee schedule, and its heavy reliance on past patent grantees to subsidize current patent applicants, frustrates the financial sustainability of the Agency.

Generally, there are a range of factors beyond the PTO's control under this funding structure that may disrupt the Agency's financial balance at any given point in time, either increasing its operational costs or decreasing its fee income. Primarily, aggregate Agency costs may increase if patent examinations presently demanded of the PTO (which may grow at a stochastic rate) increase relative to the existing stock of patents from which the PTO may collect post-allowance fees.[10] Aggregate Agency fee collections may fall for several reasons: (1) the quality of the stream of incoming patents may deteriorate, leaving the PTO otherwise inclined to grant patents less frequently; (2) patentees may elect to pay their maintenance fees at a lower rate; or (3) the aggregate incidence of small-entity applicants may rise.[11] In each such case, the indicated development will decrease the ratio between fees collected by the PTO and the obligatory

---

[9] The PTO itself reinforces these rationales in a recent Final Rule published in the Federal Register (PTO 2013), noting that (1) by "keeping front-end fees below the cost of application processing . . . the final fee schedule continues to . . . ease access to the patent system . . . . and (2) "higher back-end fees also foster innovation" as the "expiration of a patent [which occurs when renewal fees lapse] makes the subject matter of the patent available in the public domain for subsequent commercialization.".

[10] The PTO is dependent on renewal fees from patents that were issued 3.5, 7.5, and 11.5 years ago to sustain its processing of patent applications today. However, the PTO is processing many more applications today than it was even four, eight, or twelve years ago. As the ratio of the PTO's fee levels has not dramatically changed over time, it is unlikely that they were set to allow for such a dramatic growth in the volume of processed patent applications. Thus, the PTO's financial health may be threatened solely by the fact that the Agency, which is under continued pressure to expand its capacity to examine patent applications in order to decrease its growing backlog, must fund its expansion in processing capacity based on previously issued patents. As discussed below, aggregate Agency costs may also increase if incoming patent applications shift towards more complex technologies to which the PTO allocates more examination hours (and thus greater salary expenses).

[11] Small-entity examination fees are set at half the level paid by large entities (Frakes and Wasserman 2013).

operational costs. Thus, all else equal, these developments increase the likelihood that the Agency's fee collections will fail to cover its examination expenses. We next theorize the steps the PTO may take in such instances.

III. THEORY

A. Basic Framework

In this section, we introduce a simple conceptual framework of PTO decisionmaking under financial constraints. In Online Appendix A, we set forth a relatively more formal treatment for this framework. This section effectively summarizes the key points of that model. Primarily, we view the PTO as using fees collected from patent applicants and previous patent grantees to fund its examination practices.[12] Subject to a constraint that these fee collections at least match the costs associated with those patents that it examines each period, the PTO maximizes its objective function through its determination of (1) when to examine each incoming application and (2) at what rate to grant patent protection to applicants. We also allow these determinations to vary across different technologies. The PTO's objectives are primarily enhanced in this model by the timely processing of applications.[13] On the other hand, such objectives are diminished when the Agency deviates from the proper determination regarding

---

[12] Note at the outset that we take as given the congressionally-determined fee structure itself. That is, we make no attempt to model Congress' decisionmaking process.

[13] Garnering utility from processing applications, the PTO receiving an application at time *t* discounts any utility obtained from processing such application to the extent that it elects to examine this application at a later time period—e.g., *t+2*. The PTO's own rhetoric supports this characterization of its objectives, insofar as the Agency has emphasized a substantial concern with their growing backlog and has expressed a strong desire to minimize this backlog and to maintain timely processing of applications (PTO 2008). We note that one can arguably view such objectives as contemporaneous with maximizing social welfare. After all, much has been said in the literature about the welfare costs associated with patent examination delays (London Economics 2010). We acknowledge, of course, that some applicants may place value on delayed examinations to the extent that it provides them with an option value to pursue other avenues of protection. Under this alternative perspective on the Agency's objectives—i.e., viewing the PTO as maximizing social welfare—one can largely abstract away from such option-value considerations in light of the default costs associated with delays in all other contexts.

the extension of patent protection to each applicant and when the Agency extends disparate examination treatment across different types of applicants.[14]

B. Unconstrained Optimum

With these preferences, we first contend that, if unconstrained financially, the PTO would choose to examine all applications as they arrive. In other words, the PTO otherwise desires to promote the throughput of its applications and thus to avoid running a backlog of applications. The PTO could achieve this unconstrained maximum as long as the growth in incoming fees keeps pace with the growth of incoming patent applications, an outcome that could be guaranteed as long as user fees are assessed only on applicants and at a level equal to the full marginal cost (to the Agency) of examination. However, as discussed in Section II, with examination fees set below cost and with the differential covered through the imposition of post-allowance fees, the PTO is subjected to a risk that its budget constraint will bind.

C. Agency Response to Binding Budget Constraint: Low-Cost Prioritization

Once the PTO's budget constraint is triggered, the basic prediction of our model is that the PTO will attempt to maintain as much application throughput as it can under its limited resources by either taking steps to increase its average revenues per application processed—i.e., by increasing its inclinations to grant certain patents—or taking steps to decrease its average costs per application processed—i.e., by focusing its processing efforts on lower-cost applications.

---

[14] The model's assumptions regarding the PTO's preferences—its value function is increased when patent processing times are decreased and its value function is diminished upon the issuing of invalid patents—are supported by the PTO's own mission statement that it "[f]oster[s] innovation, competitiveness and economic growth, domestically and abroad by delivering high quality and timely examination of patent [] applications." (PTO 2012a)

To demonstrate this claim, we proceed in steps and first operate under the crude assumption that the PTO's only tool in its choice set is the timing by which it elects to examine its applications, where this timing decision is applied uniformly across all technologies.[15] With a limited choice set of this nature, the natural response for the Agency to take in the event that its incoming fee collections are insufficient to finance its awaiting examination costs is simply to focus on that proportion of its applications that it can afford and consequently grow a backlog of awaiting examinations. As a baseline, therefore, we associate the presence of a significant application backlog with a PTO in financial imbalance. This outcome leaves welfare lower than it would be under the unconstrained optimum insofar as some applications are delayed.

We suggest that a PTO whose objective function is increased with greater application throughput may be dissatisfied with the need for a large backlog and will be motivated to look for ways to open up additional resources for its examination efforts. Consistent with our understanding of actual PTO practices and of its examiner compensation structure, we assume here that the PTO cannot uncover such resources by simply cutting per-unit processing costs through a decision to spend less time on each application that it processes.[16] Nonetheless, we do predict that the PTO may attempt to lower the per-application cost *on average* through its selection of applications to review.

To demonstrate the selection mechanism, consider the PTO's response once we relax our previous assumption and allow the Agency to alter its examination-

---

[15] Consistent with our modeling of the PTO as holding such a limited choice set, in recent decades, the Agency likely had a limited number of ways it could reestablish financial equilibrium. During the time frame of this study the PTO did not have the authority to set the level of its fees. Moreover, when the Agency's fee revenue fell short of its allocated budget, Congress did not supply the PTO with this difference, nor did the Agency have the ability to borrow the shortfall from a third party.

[16] As part of its determination of examiner compensation, the PTO allocates a specified number of hours to applications in different technology groups, with further adjustments to this allocation based on examiner seniority. Based on information we received pursuant to a FOIA request, this examiner-hour schedule has remained completely fixed over the entirety of our sample period. As such, the PTO does not seem to have diluted its per-application work expectations in an effort to reduce costs.

timing decisions across different technologies. In this instance and during times of binding budget constraints, we predict that the PTO will elect to cut its average examination costs by focusing the applications that they elect to process, at least for some period of time, on those that cost the Agency the least to examine.[17] For a given stream of incoming post-allowance fees, prioritizing low-cost patent applications will afford the Agency the ability to process a greater number of applications presently—relative to the benchmark whereby it could not engage in preferential examination queuing across different patent types—while nonetheless satisfying its budget constraint.

As suggested, this claim effectively takes as given the aggregate amount of incoming fees. However, it is worth mentioning that by allowing the PTO to process more applications in this manner (relative to the no-prioritization benchmark where it would be forced to grow its backlog across the board), this selective examination approach will, if anything, only cause the Agency to collect more issuance fees in the near term. All told, prioritization of low-cost applications during times of binding budget constraints will enhance the profitability of those examinations selected by the Agency.

Facilitating the ability to engage in such cost-reducing selection/prioritization is the heterogeneity in examination costs to the Agency across different types of patent applications. The main source of such heterogeneity is the technology covered by the application. The technology addressed in an application is identified by the PTO before it enters examination so that the Agency can route applications to examiners—those individuals at the PTO whose primary responsibility is to examine patent applications—who will have sufficient

---

[17] Importantly, patent applicants pay the same examination fees regardless of the examination cost (to the PTO) associated with their application,, s*ee* 37 C.F.R. § 1.16(a)(1) (2011) (basic filing fee is $330 and $165 for small entity); 37 C.F.R. § 1.16(k) (2011) (utility search fee is $540 and $270 for small entity); 37 C.F.R. § 1.16(o) (2011) (utility examination fee is $220 and $110 for small entity), in which event variations in examination costs across application types translate into variations in profitability to the Agency.

scientific background to review the application. PTO examiners are allocated a pre-determined number of hours to review each application and are effectively compensated under the assumption that they spend that amount of time on the application. Considering the cost per hour of an employee's time, the per-application examination cost is largely a function of this examination-hour allotment. Acknowledging that technologies involve varying degree of examination complexity, the PTO allocates different processing times (or pre-determined number of hours) to applications of different technologies.[18]

While the present study is not designed to explore the mechanism behind such practices (we are pursuing such questions with follow-on work), there may be at least two different channels for prioritizing certain types of patents for review. In the first channel, the PTO may simply cut resources directed towards the examination of its most costly patent applications—e.g., by enacting a hiring freeze in technologies that are the most expensive to review, thereby reducing resources for those areas as their associated examiners resign. In the second channel, the PTO may simply shift resources away from its most costly applications towards its least costly applications—e.g., by reassigning current patent examiners.[19]

On a final note, while prioritization of low-cost applications will allow the PTO to enhance its objective function through the consequent increase in timely reviews, such prioritization also comes at a cost. We model the PTO as placing some disutility on extending preferential treatment to certain technologies. As

---

[18] For instance, for patent applications involving artificial intelligence, an examiner with experience corresponding to a General Schedule (GS level) 12 is apportioned 31.6 hours. In contrast, examiners with the same experience in the compound tools area (i.e., hammers, wrenches, etc.), are allotted only 13.8 hours. In Online Appendix A, we provide greater detail regarding those factors bearing on PTO examination expenses across technologies (including secondary factors beyond technological complexity).

[19] The ability of the PTO to partake in this second channel will be limited by the extent that the most costly and the least costly patent applications to examine require the same scientific training or knowledge to review. We have submitted additional FOIA requests with the PTO bearing on their staffing practices that we hope will facilitate some direct tests of these mechanism considerations.

such, in determining its examination-timing decisions for each technology during times of binding budget constraints, the PTO is effectively forced to balance the enhanced utility afforded by greater throughput of applications with the disutility from driving a wedge in practices across technologies. In the event that an interior solution is met, we predict that the Agency will nonetheless engage in some limited degree of prioritization of low-cost technologies in its examination queuing. Having addressed the first manner in which parties often respond to budgetary concerns—reducing costs—we now address the second general approach to budget balancing—raising additional revenues.

D. Agency Response to Binding Budget Constraint: Inflationary Granting

In the model thus far, we have assumed that the PTO will decide to grant or reject any given patent application in accordance with optimal innovation policy—e.g., to reject applications covering "obvious" innovations. We now expand the PTO's choice set to allow the Agency to deviate its actual grant rate from this otherwise appropriate grant rate. As specified in Frakes and Wasserman (2013), the PTO may elect to grant at an otherwise elevated rate in order to raise additional revenues during times of binding budget constraints, taking advantage of the fact that the PTO's fee structure provides the Agency with nearly half of its budget from funds that arise only in the event that it grants patents (through the collection of issuance and maintenance fees).[20] That is, a financially constrained PTO may grant a patent that an unconstrained PTO would have rejected in an effort to generate additional revenue through post-allowance fees.

---

[20] We contend that this practice of excessive granting provides the Agency with the most ready source of additional funds in the event of a constrained budget. To begin, over our sample period, the PTO did not possess the ability to change its fee levels themselves in response to a budgetary shortfall. Furthermore, we do not model the possibility of the PTO lobbying Congress for additional funds at such times. While the Agency has enjoyed some success in lobbying for larger fees, its success has been far from universal (Watson 1992).

Furthermore, balancing its desire to increase application throughput with the disutility it places on deviating from its otherwise optimal grant rate, we predict that the Agency may limit the degree of any such deviation by targeting these distortionary over-granting proclivities on patent types that generate the highest fees for the Agency—i.e., large entity applicants (which pay double the fees relative to small-entities) and applicants within technologies that consistently renew their patents at higher rates.[21]   Frakes and Wasserman (2013) present evidence consistent with such predictions. We replicate these findings, in part, below.

<center>E. The Choice between Policy Levers</center>

As the previous two subsections have illustrated, the PTO has two policy levers by which it can attempt to establish financial equilibrium while nonetheless facilitating sufficient aggregate throughput of examinations—i.e., prioritizing low-examination-cost technologies and extending preferential granting rates to high-fee-generating patent types / technologies. While acknowledging that the PTO may, at times, turn to both such practices, the focus of this last subsection is to examine when the PTO may preference one mechanism over the other.

As discussed in Section II, the PTO historically has not had the authority to spend its entire fee collections. Instead, Congress routinely utilized a portion of the PTO's fees to fund other government operations, a practice referred to as "fee diversion." The PTO's ability to retain its fee income, however, has varied over time. Consider first the period of time between 1991 and 2003, an era which we refer to as "strong-fee diversion" because the Agency's spending authority was effectively set *below* its anticipated fee collections (Frakes and Wasserman 2013). By effectively capping the PTO's fee income in this manner, Congress likely

---

[21] At the same time, the PTO will be inclined to limit the degree to which it treats technologies / patent types differently in this manner as a result of the disutility it places on distortions in PTO practices across applicant types.

<center></center>

blunted the PTO's incentives to increase its grant rates in response to budgetary pressures because the Agency was less likely to retain the fruits of such efforts.[22] Accordingly, during this period, we predict that the PTO will be relatively more inclined to respond to a binding resource constraint by decreasing its average per-application examination costs through its prioritization practices.

On the other hand, we characterize 2003 onwards as a period as "weak-fee diversion" because, at this time, Congress effectively lifted the most obstructive aspects of such practices, leaving the PTO with the newly acquired ability to spend *all* of its anticipated fee collections.[23] This alteration of practices, we contend, enhanced the PTO's incentive to seek an expanded grant rate (and a correspondingly higher anticipated revenue amount) in response to a binding budget constraint. As such, we predict that during the post-2003 period, relative to the prior period, the PTO will be more likely to turn to distortionary granting practices than to low-cost-prioritization practices during times of financial stress. Further, we suggest that the PTO may have even taken advantage of the additional revenue it could raise during this latter period to unwind some of the backlog discrepancies that it had amassed across technologies in the preceding period, acknowledging, after all, that the Agency, all else equal, places value on minimizing processing disparities across technologies. In other words, our analysis suggests that following 2004, the PTO may have been inclined to reverse course and now extend preferential queuing treatment to high-cost (as opposed to

---

[22] Consider the year 2002, for instance. While the PTO projected it would collect roughly $1.35 billion over that year, Congress only authorized a budget of $1.15 billion. (Office of Management and Budget 2002, PTO 2002).

[23] We label this period as "weak" given that some vestiges of fee diversion remained, mainly that fee collections beyond the PTO's anticipated revenue stream for that year may not be made immediately available to the PTO. Nonetheless, during this period, if the PTO wished to commit itself to a higher grant rate and thus set a higher anticipated revenue target, it would be able to retain all fee income that fell within this target. Further details on such fee-diversion practices are provided in Frakes and Wasserman (2013)

low-cost) technologies in order to bring its examination queuing practices back into balance.[24]

IV. DATA AND METHODOLOGY

A. Data

Following the filing of a Freedom of Information Act (FOIA) request with the PTO, we received data on technology-specific examination delays over time. More specifically, for each patent class and for every year from 1991 to 2010, we collected information on the average length of time between receipt of an application and the completion of the "first office action" by the PTO—that is, the first substantive communication from the PTO setting forth the Agency's arguments regarding the patentability of the applicant's invention. The resulting dataset is organized at the technology-year level. We categorize technologies using the 37 sub-categories specified by Hall, Jaffe and Trajtenberg (2002).[25]

---

[24] We note that we are focusing on the behavior of the Agency, as distinct from the behavior of the individual examiners within the PTO acting in accordance with their own objectives. While very little empirical work has been done focusing on understanding the determinants of PTO decisionmaking, some empirical research has been done exploring the role of examiner heterogeneity in explaining patent processing outcomes (Cockburn et al. 2003, Lemley and Sampat 2012, Mann 2013). In one recent, Agency-specific empirical study, Hedge (2012) discusses how the PTO's appropriations process with Congress—including the diversion of funds from the Agency—may leave the PTO subject to a risk that it will have insufficient funds to accommodate the examination of an unanticipated inflow of examinations. Hedge nicely motivates this concern through a graphical time series analysis demonstrating that the Agency's growing backlog appears to emerge more strongly during periods of time when the PTO's appropriated budget falls short of what it needs to examine the incoming applications for that year. Hegde's analysis arguably lends support to the first part of our analysis in demonstrating that the nature of the PTO's funding process exposes it to a risk that it will be unable to generate enough funds to examine all incoming applications (though emphasizing separate aspects of the funding process). Hegde's analysis, however, does not address the heart of this paper—that is, how the Agency's technology-specific examination practices may be impacted in the Agency's attempt to address any resulting funding gaps.

[25] The raw data on examination durations that we received from the PTO is at the level of a PTO "class" (of which there are nearly 500), whereas the raw data that we received from the PTO allowing us to estimate examination costs is at the level of an "Art Unit" (of which there are nearly 300). Accordingly, the lowest level of disaggregation that we could conceivably use in approaching our analysis is the relatively coarser Art Unit level. However, approaching the analysis at the Art Unit level would still be problematic in light of the difficulties involved in mapping PTO classes to Art Units, primarily the fact that, even though Art Units are coarser, classes are not exactly perfect subunits of Art Units—i.e., there are times when given classes cover more than one Art Unit. As such, to alleviate these mapping difficulties, and to alleviate concerns over slight changes in class and Art Unit classifications over time (variations that are necessarily dampened as one conceptualizes technologies at broader and broader levels), we simply elect to approach this analysis using the 37 technology sub-category groups developed by Hall, Jaffe and Trajtenberg (2002) and frequently used by the literature. To the extent that the PTO is actually implementing any examination prioritization practices at the Art Unit level, this choice of aggregation may lead to some attenuation of the results. Of course, to the extent that the PTO

The second major dimension of data required for this analysis bears on the variability in average examination costs facing the PTO across its various technology groups. Guided by our understanding of the examination process and using rich, technology-specific data on examiner workload schedules (likewise obtained pursuant to a FOIA request), we estimate the average cost of a patent examination (from the perspective of the PTO) within each technology classification. Further details on this calculation are provided in Online Appendix A.[26]

Table 1 presents descriptive statistics for the key variables used in the regression analysis. To focus on demonstrating the variability in examination costs, Table 1 standardizes the estimated costs per average application such that the mean cost across technologies equals 1. This measure's standard deviation equals 0.14, with technology-specific costs ranging from a low that is roughly 33% below the mean amount to a high cost that is 44% above the mean amount. In Online Appendix A, we present a breakdown of the standardized cost estimates for each of the 37 technology categories. This breakdown evidences meaningful variation in examination costs across technologies.

TABLE 1: MEANS AND STANDARD DEVIATIONS OF SELECTED VARIABLES

First-Office-Action Pendency (Months)	17.02
	(8.16)
Estimated Cost Per Application (Normalized to Mean 1)	1.00
	(0.14)

organizes any prioritization efforts at a coarser technological basis, the chosen specification may better capture the dynamics under investigation.

[26] As demonstrated in Online Appendix B, our results are not sensitive to the precise calculation undertaken and are robust to alternative allocations of examination costs across technologies, including one simply based on complexity-based variations in the baseline number of hours allotted to given technologies.

Sustainability Score (Normalized to Mean 1)	1.00
	(0.29)

Standard deviations are in parentheses. Descriptive statistics are from a sample of 740 technology-year cells from 1991 to 2010. The statistics are weighted by the number of first office actions undertaken in the relevant category-year cell. Patent processing and cost data is from the PTO.

On average over our sample period, the PTO took 17.0 months after receiving an application before completing its first office action. For those technology categories with below-average examination costs to the PTO, the Agency took an average of 15.1 months to complete its first office action. For those technologies with above-average costs, this duration was, on average, 20.1 months. As such, the PTO does in fact take longer on average to examine applications within technologies that carry high examination costs for the Agency. Of course, our story cannot simply end there. Examination times may differ across technologies for a host of reasons. It becomes necessary for us to separate the influence of such other factors from the financial motivations emphasized by the above theoretical predictions, an exercise to which we now turn.

B. Methodology

Sources of variation—In order to identify the relationship between the PTO's financial incentives and its examination prioritization practices, we draw upon two fundamental forms of variation along this financial dimension. First, we explore how the PTO's examination-timing practices fluctuate in connection with changes in the PTO's aggregate financial health, as captured by the indicators discussed below. That is, we explore whether the PTO begins to extend examination-queuing priority to its least costly examinations upon the occurrence of events that leave the PTO with insufficient funds to examine all awaiting applications. Second, we explore how the PTO modified its practices in connection with the alteration of congressional policy in 2004 that left the PTO

with greater ability to retain any fees that it collects. As explained above, we contend that this reform may have left the PTO more inclined to turn to distortionary granting practices (as opposed to distortionary prioritization practices) as a mechanism to satisfy its binding budget constraint while nonetheless processing as many applications as it can.

We essentially treat each of these sources of variation as exogenous in nature, at least with respect to how these aggregate developments subsequently impact *technology-specific* decisions. We further discuss the validity of this assumption below and estimate alternative specifications that are designed to achieve greater independence between the markers of the PTO's financial health and its examination-queuing decisions.

Difference-in-Difference Specifications—First, we consider those specifications that draw upon a negative financial shock to the PTO's health. A development that leaves the PTO with inadequate fees to cover its obligatory costs will, absent any alteration of practices, necessarily result in some need for an expanded backlog—i.e., added processing delays. Our hypothesis is that the PTO will alter its practices such that this necessary growth in examination delays will be targeted on high-cost technologies, leaving it less pronounced in low-cost technologies. The resulting decrease in average examination expenses will allow it to process more applications with its limited source of income. To capture this prioritization response, we estimate the association between an aggregate financial shock to the PTO and the differential examination pendencies across technologies that carry different examination costs to the PTO. In other words, we test for a heterogeneous response across technologies to this overall financial shock.

One can effectively view this as akin to a difference-in-difference design, whereby the high-cost technologies may be seen as the "treatment" groups, insofar as the necessary examination delays stemming from the negative shocks to

the PTO's aggregate financial health will be targeted on these groups.[27] Among other things, this specification carries the benefit of accounting flexibly for unobservable determinants of patent examination durations (e.g., labor market developments that impact the ability of the Agency to hire its examiner corps), under the assumption that such unobservable factors impact PTO practices with respect to its treatment of high-cost and low-cost technologies alike.

Following other studies that have tested for heterogeneous impacts of national reforms/shocks (Acemoglu and Finkelstein 2008), we estimate the following:

(1) $PENDENCY_{ct} = \alpha + \gamma_c + \lambda_t + \beta_1 (SUSTAINABILITY_t * COST_c) + \beta_2 X_{ct} + \varepsilon_{ct}$

where $PENDENCY_{ct}$ is the average length of time between the filing of an application and the completion of the first office action in reviewing that application, where this average is taken within each technology category, *c,* and year, *t,* cell.[28] Category fixed effects and year fixed effects are specified by γ_c and λ_t, accounting for fixed differences in patent pendency lengths across technologies and years. $SUSTAINABILITY_t$ captures the annual sustainability score described below, a metric indicative of the PTO's financial health, with a high score representing a more favorable financial balance. Technology-specific estimates of average examination costs are captured by $COST_c$. X_{ct} includes certain time-varying covariates specific to technology categories, allowing us to account for temporal fluctuations in certain features of the underlying

---

[27] The specification is similar to a typical difference-in-difference specification to the extent that it captures a heterogenous response to some external shocks across difference groups predicted to be more or less exposed to that shock. Of course, it naturally differs from the typical difference-in-difference specification to the extent that the "shock" variable in this specification is changing continuously over time.

[28] As demonstrated by Online Appendix B, the results presented below are nearly identical when we also consider alternative pendency lengths based on the average total time between the filing of an application and the ultimate issuance of a patent, out of the full universe of patent issuances over the sample period.

applications.[29] Included in X_{ct} are also controls for the share of new application filings arising from the respective technology to address concerns that any change in the differential pendency between high and low cost technologies is driven by differential trends in filings across technologies.

   The coefficient of interest is represented by β1, capturing the degree to which a deterioration in the PTO's financial state, as captured by a decrease in the sustainability score, is associated with an increase in the differential examination pendency length between high-cost and low-cost technologies.[30] An increase of this nature would correspond to an estimated *negative coefficient* of the interaction term between the sustainability score and technology-specific cost variable—i.e., the difference-in-difference variable of interest—and would be consistent with the theoretical predictions set forth in Section III. In this initial specification, we focus on the period of strong fee diversion prior to 2004. During this time, the PTO faced a consistent fee-diversion environment in which the Agency was, if anything, perhaps more inclined to turn to low-cost prioritization rather than inflationary granting practices as a way to ensure that it examines as many applications as it can while still satisfying its budget constraint.

Dynamic Specifications—Next, we track how the differential examination pendency lengths between high- and low-cost technologies trend on a year-by-year basis. To allow us to also observe how this differential trend evolves after the shift towards a weak fee-diversion regime in 2004, we estimate this specification over the full 1991-2010 period. As such, we estimate the following:

---

[29] Such covariates include the average number of patent claims, the average number of citations to the relevant patents, and the percentage of patentees representing various inventor types (e.g., individual, corporate, government, etc.). Data on claims, assignee types, and citations by technology category were obtained from the National Bureau of Economic Research Patent Data Project, available at https://sites.google.com/site/patentdataproject/Home. Technically, these averages are taken out of those patent applications that ultimately issue.

[30] Perhaps more specifically, this coefficient captures the degree to which the PTO's examination-pendency response to a negative shock to the PTO's financial state increases as we move along the distribution of examination costs across technologies.

(2) $PENDENCY_{ct} = \alpha + \gamma_c + \lambda_t + \sum_{t=1991}^{2010}(\beta_t(d_t * COST_c)) + \beta_2 X_{ct} + \varepsilon_{ct}$

where d_t is an indicator variable for year t.[31]

Fee-Diversion Interaction Specifications—In our final specification, we explore the impact of the 2004 congressional fee-diversion reform. Given that the PTO may take advantage of this reform to rely more heavily on altering its granting practices to deal with financial difficulties as opposed to the low-cost prioritization approach, we test for a weakening (or reversal) in the relationship between sustainability fluctuations and the differential examination pendency patterns. To execute this final specification, we now consider the full sample period[32] and modify specification (1) above to interact the main difference-in-difference variable (SUSTAINABILITY*COST) with an indicator variable for being in the weak fee-diversion period following 2003, FEE_DIVERSION:[33]

(3) $PENDENCY_{ct} = \alpha + \gamma_c + \lambda_t + \beta_1 (SUSTAINABILITY_t * COST_c) + \beta_2(SUSTAINABILITY_t * FEE\_DIVERSION_t) + \beta_3(COST_c *$

---

[31] Since the technology-year covariates are only available in the pre-2006 period, our primary specifications consider the longer sample period discussed below. Nonetheless, we also present results with such measures included.

[32] We predict that the PTO will respond differently following 2003 to a negative financial shock to its aggregate health. That is, it will no longer continue to respond to a negative shock by prioritizing low cost applications as it did previously (at least we predict that it will not do so with the same vigor). In fact, the Agency's aggregate health continues to deteriorate until 2008, giving empirical relevancy to this discussion. In 2009 and 2010, however, the Agency's sustainability score begins to slowly move back up. This fee-diversion analysis does not predict that the Agency's response to a positive financial shock will similarly reverse in the post-2003 period. That is, we do not predict that the Agency will respond to a positive financial shock in this period by shifting towards prioritization of low-cost applications. If anything, we predict that a positive financial shock during this time period will only reinforce its ability to bring its technology-specific backlogs back into balance through a shift towards prioritization of high-cost groups. As such, to bring the empirical specification inline with our theoretical predictions, we limit the sample period for this fee-diversion interaction specification to the period prior to 2009, thus only considering post-2003 years in which the Agency's health is declining. In any event, we note that the results persist when nonetheless estimating this specification over the full 1991-2010 period.

[33] Since technology-year covariates are not available over this full period, we do not include them in this specification, except for the initial application filings control (which is fully available). We also include all constitutive elements of this interaction.

$$FEE\_DIVERSION_t) + \beta_4(SUSTAINABILITY_t * COST_c *$$
$$FEE\_DIVERSION_t) + \beta_5 X_{ct} + \varepsilon_{ct}$$

The coefficient of interest is captured by β_4, which can effectively be interpreted as the degree to which the main difference-in-difference finding from specification (1) is more pronounced during periods of weak fee diversion relative to periods of strong fee diversion. Our theory predicts that sustainability score reductions—i.e., negative financial shocks to the PTO—lead to increases in the differential examination pendency length between high and low cost technologies—i.e., a negative coefficient of β_1. Moreover, our theory predicts that this negative relationship will be smaller (or even turn positive) in the post-2003 period. As such, our theory predicts a positive value for the coefficient of the interaction variable, β_4.

Financial Sustainability Calculations—Vital to this analysis is an assessment of when the PTO is likely in financial peril and thus in need of taking the hypothesized actions to ensure that it can break even while financing as many examinations as it can. To facilitate this assessment, we construct empirical markers indicative of a PTO in danger of failing to break even. In developing these markers, we draw upon the theory set forth in Section II, whereby we predicted that the PTO would be more likely to trigger its budget constraint upon a change in various factors, including, a decrease in its annual renewal fee collections (decreasing aggregate revenues), an increase in the average complexity of its examinations (increasing aggregate costs), a decrease in the proportion of large-entity patentees (decreasing aggregate revenues) and an increase in the number of applications presently demanded of the PTO and awaiting examination relative to the existing stock of patents from which the PTO collects post-allowance fees (increasing aggregate costs relative to revenues).

In any given year, all of these factors are changing at the same time, even though the PTO is only subject to a single sustainability constraint. If the PTO elects to distort its practices that year, it will only do so in response to some overall assessment as to its likelihood of facing financial friction. Accordingly, in our primary approach, we follow Frakes and Wasserman (2013) and combine these factors into a composite sustainability measure.[34] This "sustainability score" is constructed to simulate the impact of these various developments on the PTO's financial balance in a manner consistent with the empirically relevant influence of each such factor. To capture this balance, we use data on the annual fluctuations in the above factors (e.g., annual maintenance rates), along with information on the parameters of the PTO's fee schedule, to simulate an annual measure equal to the ratio between (1) the issuance and post-issuance fees generated by the existing stock of patents at a given point in time and (2) the net costs associated with the examinations demanded of the PTO at that time. A higher simulated sustainability score is suggestive of fewer financial pressures and thus less need of additional funds, while a decline in this score suggests a negative shock to the PTO's financial health. The score is meant to provide a sense of the ease with which the PTO may use its stream of incoming funds to satisfy the substantial costs associated with all of those examinations presently awaiting the PTO. Online Appendix A provides greater detail regarding this calculation.

---

[34] To complement this primary approach based on an aggregated measure, in Online Appendix B, we essentially treat each of these factors as separate markers of a PTO more likely to face a binding budget constraint and estimate regressions that explore how fluctuations over time in these factors are associated with differential examination pendencies across high- and low-cost technologies.
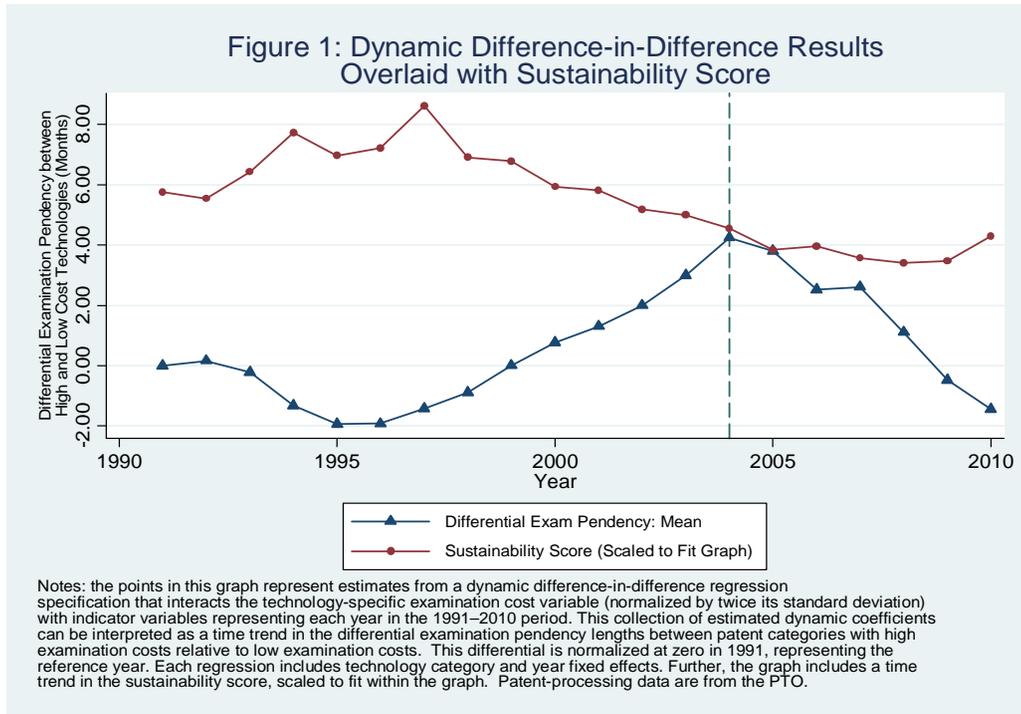
V. RESULTS

A. Dynamic Difference-in-Difference Results: Graphical Analysis

Overview—We begin our empirical analysis with a visual depiction of our key findings. In Figure 1, we plot a year-to-year time trend in the differential examination pendency lengths between high- and low-cost technologies, spanning the 1991 to 2010 period, where this differential is normalized to 0 in 1991.[35] The points along this trend—marked by the triangles within the graph—represent coefficients (β_t) from the dynamic difference-in-difference specification indicated in equation (2).[36]

---

[35] In subsection V(B) below, we explain in greater detail how to interpret this allocation of technologies into high- and low-cost bins.

[36] Considering that our theoretical foundation begins with the risks of financial unsustainability posed by the PTO's *particular* fee structure, and considering that our precise empirical inquiries focus on the distortionary examination practices used when those risks materialize, we naturally select 1991 as the beginning of our sample period, as this was the time when the PTO implemented this particular fee structure. Structuring a natural-experiment methodology drawing on this 1991 development itself would be complicated by the fact that prior to that time, when the Agency was heavily financed through general tax revenues, the Agency may have still faced cost pressures and may have still desired to defray high examination expenses in order to stay within fixed congressional budgets. After all, as suggested previously, a basic premise of retreating from general revenues was to avoid this very development in the first place. An ideal counterfactual analysis for our purposes would be one that explored the impact of moving from a user-fee structure in which examination fees covered all examination costs (eliminating the risks at issue) to the present structure in which examination fees are quite insufficient to cover such costs, as opposed to a reform of the actual 1991 variety. Nonetheless, to present a complete picture of PTO examination practices over the entire period in which we have available data—1986 to 2010—Figure B7 of Online Appendix B plots a corresponding time trend in the differential examination pendency over this longer period. As this figure demonstrates, this more complete picture does not alter our conclusions—if anything, it strengthens them. The 1991 reform was arguably more relevant for our previous research into distortionary granting practices (Frakes and Wasserman 2013), considering that the ability to deal with financial difficulties through fee-generating practices was naturally limited prior to the adoption of the near fully user-fee funded structure in 1991.

Figure 1: Dynamic Difference-in-Difference Results Overlaid with Sustainability Score

Notes: the points in this graph represent estimates from a dynamic difference-in-difference regression specification that interacts the technology-specific examination cost variable (normalized by twice its standard deviation) with indicator variables representing each year in the 1991–2010 period. This collection of estimated dynamic coefficients can be interpreted as a time trend in the differential examination pendency lengths between patent categories with high examination costs relative to low examination costs. This differential is normalized at zero in 1991, representing the reference year. Each regression includes technology category and year fixed effects. Further, the graph includes a time trend in the sustainability score, scaled to fit within the graph. Patent-processing data are from the PTO.

Importantly, for our purposes, the absolute value of these coefficients is not crucial. Instead, our primary concern is how the estimates of these coefficients trend upwards (i.e., the patent pendency of high-cost technologies grows at a higher rate than low-cost technologies) or downwards and how such movements correspond with the trend of the PTO's financial health. To facilitate a visual representation of this correspondence, we overlay the differential examination pendency time trend with a time trend in the sustainability score, as represented by the circular points in Figure 1. Furthermore, we include a dashed vertical line in 2004 to mark the time at which fee diversion policies changed, leaving the PTO with arguably greater incentives to turn to fee-related distortions in order to cover its budgetary shortfalls. In Figure 2, we replicate Figure 1, except that we calculate an alternative sustainability score that estimates the present examination demand facing the PTO using the aggregate backlog of applications awaiting a

first office action, as opposed to the total number of applications awaiting completion of examination. Overall, the patterns presented in Figures 1 and 2 are consistent with each of the theoretical predictions set forth in Section III.

Figure 2: Dynamic Difference-in-Difference Results Overlaid with Alternative Sustainability Score

Notes: this figure replicates Figure 1 except that it overlays an alternative sustainability score that uses the aggregate inventory of applications awaiting a first office action as the aggregate examination demand facing the Agency.
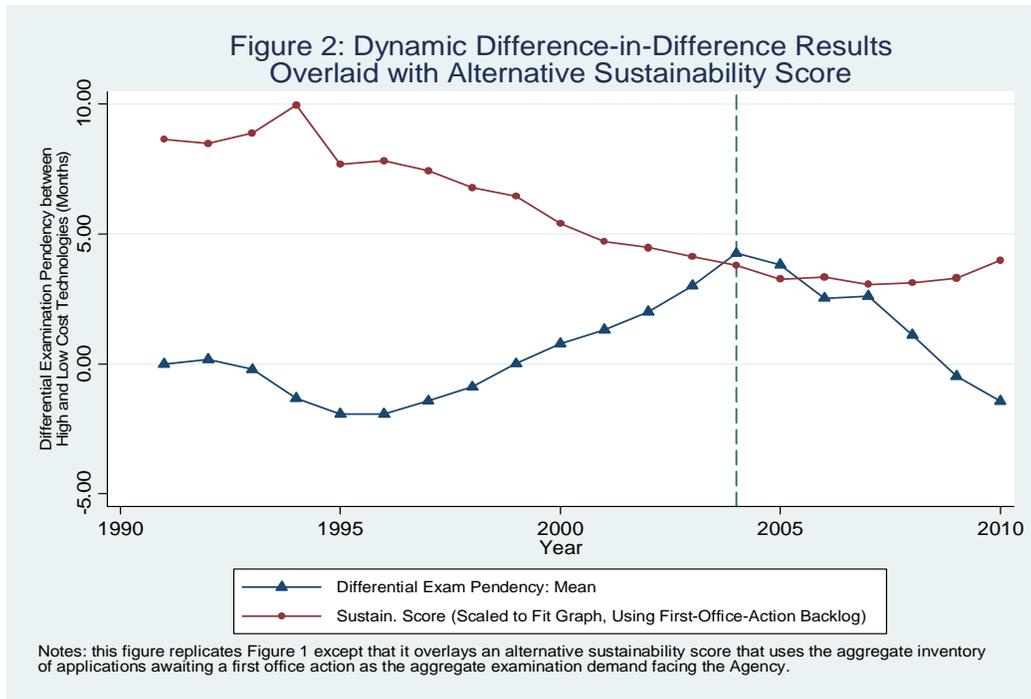
Summary of Graphical Analysis—To begin, in the early-mid 1990s, the sustainability score is rising, suggesting that the PTO's financial position may have been improving and thus that the Agency likely found itself less in need of taking any distortionary actions to either increase or conserve resources.[37] This period thus presents us with an important falsification exercise: given that PTO financial conditions were positive at this time, do we observe the PTO taking measures that we only predict it taking during times of financial instability?

---

[37] The PTO may have experienced a financial boost over this short time period, as it began to collect the substantial 12-year maintenance fees for the first time during these years. (PTO 1995). The PTO saw a substantial jump in renewal fee income in the fiscal year 1994. *Compare* (PTO 1995, p. 59) (noting that 32% of patent fee collections resulted from maintenance fees) *with* (PTO 1994, p. 29 and fig.9 (noting that 26% of patent fee collections resulted from maintenance fees).

Frakes and Wasserman (2013) previously demonstrated that the PTO, over this time period, did not appear to engage in distortionary granting practices (i.e., granting at higher rates to high fee-generating patent types), a finding that is replicated in Figure 4 below. Likewise, in the present analysis, the results do not suggest that the Agency engaged in the resource-conserving tactic of prioritizing its low-cost examinations during this period. If anything, it appears that the Agency may have *decreased* its relative examination delay between high- and low-cost technologies over these early years—that is, it may have given preferential examination-queuing treatment to high-cost areas. Because the initiation of our sample period marked a point in which the Agency had been maintaining a significantly longer queue for high-cost technologies, this development is nonetheless consistent with our theory in that the PTO may have taken advantage of these favorable financial conditions to undo some of the backlog differential that it had amassed up to that time (perhaps as a response to bugetary pressures during the 1980s).[38]

In the period following 1997 (or in the period following 1995 as in the case of the alternative sustainability score presented in Figure 2), the Agency's sustainability score begins to fall steadily, suggesting a series of negative shocks to its ability to satisfy its current obligations with the incoming stream of post-allowance fees. Consistent with the above theoretical predictions, it is indeed at this time when the PTO begins to maintain longer and longer examination delays for high-cost relative to low-cost technologies, as signified by the similarly smooth upward trend in the coefficient estimates in the 1997-2004 period. In

---

[38] After all, in our model, the PTO does place a disutility on differential treatments across technologies, all else equal. The average technology-specific backlog, normalized by the number of initial filings specific to that technology, for the 4 highest-cost technologies in the 1985-1990 period was roughly double that of the same average for the 4 lowest-cost technologies. While the PTO did not establish its near fully user-fee funded structure until 1991, it may have nonetheless been cost conscious in the time prior to that point. After all, even with budgets fixed and provided by Congress, agencies may be concerned with costs unexpectedly exceeding those budgets, leaving them unable to fully satisfy their obligations. As such, prior to 1991, cost considerations may have similarly induced the Agency to build up examination delays in its costliest technologies, just as we demonstrate may have occurred in the latter part of the 1990s into the 2000s.

other words, in the face of a growing and growing financial imbalance, the PTO appeared to have conserved resources by extending more and more preferential queuing treatment to its cheapest examinations, allowing it to process as many of its overall obligations as it could.
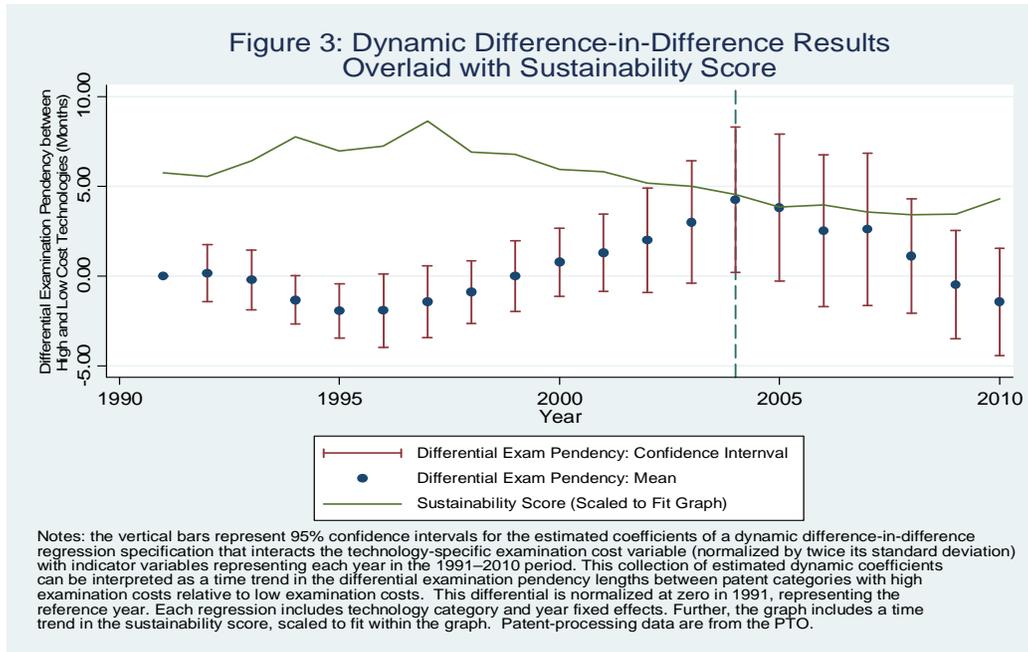
In the immediate years following the 2004 congressional reform, the financial condition of the PTO remains strained. We predicted that the PTO would at least slow down its distortionary prioritization practices at this point given the shift in congressional fee diversion practices leaving the Agency with greater ability to rely upon fee-generating practices as a tool to deal with binding budget constraints. In actuality, not only did these prioritization practices slow down, they began to reverse course. As Figures 1 and 2 demonstrate, this trend towards more and more prioritization of low-cost applications comes to a relatively abrupt end in 2004, as estimates of the coefficients begin to trend downwards. That is, it is at this time that the PTO begins to undo these prior tendencies and bring its technology-specific backlogs into balance by decreasing the relative examination pendency between high- and low-cost technologies—i.e., by swapping the target of its preferential queuing treatment in favor of high-cost technologies. The additional revenues collected / retained through distortionary granting practices— now made easier by Congress's reform—may have been beneficial enough that the Agency found itself in a position where it could begin to bring its differential examination-timing practices back into balance, as Figure 1 demonstrates the Agency may have done (after all, all else equal, the PTO disfavors such imbalance under our model). Reinforcing this ability to stabilize its prioritization practices, the PTO's financial sustainability also appears to have started to swing back in a favorable direction at the very end of the sample.

Inference and Graphical Analysis—In Figure 3, we attempt to demonstrate the statistical significance of the various swings in the diferential examination

pendencies between high- and low-cost technologies just discussed. As above, over the first half of the 1990s, this differential examination pendency trend is falling. By 1995, the confidence interval for the estimated differential examination delay is bounded below 0, the reference point marking the beginning of the sample in 1991, suggesting that the trend over this period is not likely to be a reflection of pure chance. Similarly, demonstrating the significance of the observed increase in the differential examination pendency between 1997 and 2004, Figure 3 illustrates that by 2000, the point estimates for these coefficients have bounded above the confidence interval for the 1997 coefficient estimate. Finally, demonstrating the statistical significance of the decline in the differential examination pendency lengths that occurred following 2004, Figure 3 further illustrates that by the end of the decade, the confidence intervals for the coefficient estimates have bounded away from the point estimate for the differential examination pendency at 2004.

Distortionary Granting Mechanisms—During times in which the Agency's sustainability score was declining, Frakes and Wasserman (2013) hypothesized, and indeed found evidence suggesting, that the Agency would begin to grant patents at incrementally higher rates (thus rejecting them at lower rates) within those patent classes that garner higher post-allowance fees. Such patent types include large entity applicants, which pay double the fees relative to small-entities, along with technologies that consistently renew their patents at high rates, thus generating greater renewal fees for the PTO.[39]

---

[39] We note that high maintenance-rate technologies are not necessarily one and the same as high-examination-cost technologies. While maintenance fees do not vary across technology classifications, patentees elect to pay these fees at dramatically different rates across such classifications. Accordingly, the PTO stands to gain more financially by granting patents in technologies that are likely to be renewed at a higher rate relative to those likely to be renewed a lower rate. The PTO also stands to earn more fee income by granting a patent to a large entity rather than a small entity, as large entities pay twice the amount of post-allowance fees than small entities (Frakes and Wasserman 2013)

Figure 3: Dynamic Difference-in-Difference Results Overlaid with Sustainability Score

Notes: the vertical bars represent 95% confidence intervals for the estimated coefficients of a dynamic difference-in-difference regression specification that interacts the technology-specific examination cost variable (normalized by twice its standard deviation) with indicator variables representing each year in the 1991–2010 period. This collection of estimated dynamic coefficients can be interpreted as a time trend in the differential examination pendency lengths between patent categories with high examination costs relative to low examination costs. This differential is normalized at zero in 1991, representing the reference year. Each regression includes technology category and year fixed effects. Further, the graph includes a time trend in the sustainability score, scaled to fit within the graph. Patent-processing data are from the PTO.

Additionally, Frakes and Wasserman (2013) demonstrated that this fee-based distortion would be more pronounced following the move towards weak-fee diversion policies in 2004. We demonstrate / replicate these findings in Figures 4 and 5, further supporting the suggestion from Figures 1-3 that the PTO may have turned away from its low-cost prioritization tactics to dealing with budgetary woes toward distortinary-granting tactices following the congressional fee-diversion reform in 2004. More specifically, Figures 4 and 5 plot a time trend in the differential grant rate between high fee-generating patent types and low fee-generating patent types. In Figure 4, we specify high fee-generating patent types as those technologies with high maintenance rates, whereas, in Figure 5, we specify high fee-generating types as those patent applicants with large entity status. As suggested by these graphs, the PTO's distortionary tendencies to grant more heavily to certain patent types appears to take off right at 2004.

Figure 4: Time Trend in the Differential Grant Rate Between Technologies with High and Low Maintenance Rates
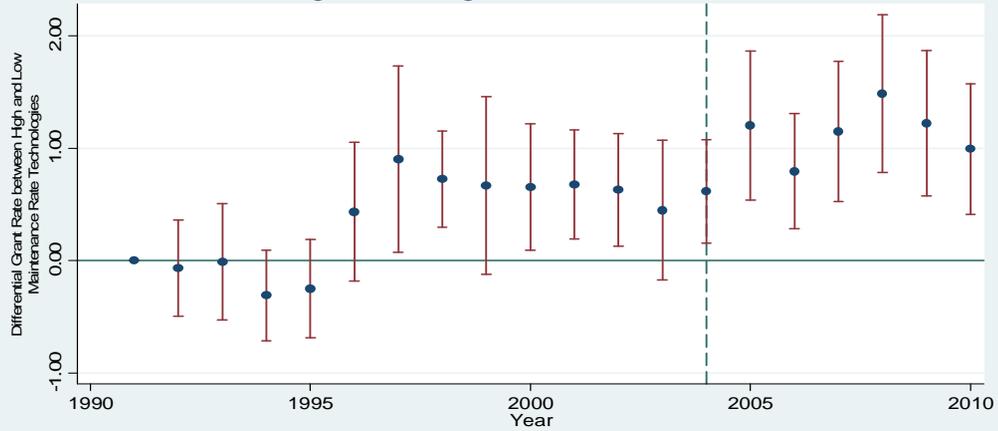
Notes: the vertical bars represent 95% confidence intervals for the estimated coefficients of a dynamic difference-in-difference regression specification that interacts the technology-category maintenance rate (4 year) with indicator variables representing each year in the 1991–2010 period. This collection of estimated dynamic coefficients can be interpreted as a time trend in the differential grant rate between high and low maintenance rate technologies. This differential is normalized to zero in 1991, representing the reference year. Each regression includes technology category and year fixed effects. Patent-processing data are from the PTO.

Figure 5: Time Trend in the Differential Grant Rate Between Large and Small Entities
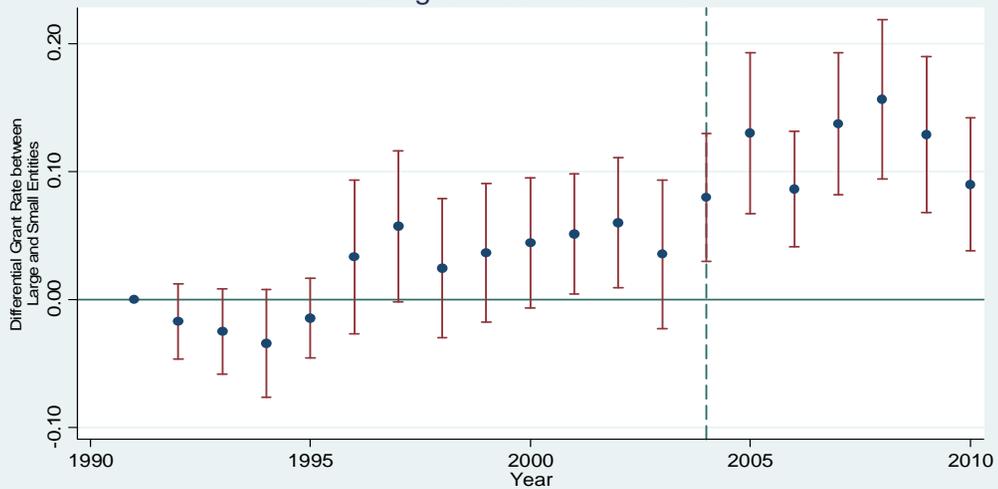
Notes: the vertical bars represent 95% confidence intervals for the estimated coefficients of a dynamic difference-in-difference regression specification that interacts the large-entity-status variable with indicator variables representing each year in the 1991–2010 period. This collection of estimated dynamic coefficients can be interpreted as a time trend in the differential grant rate between large and small entities. This differential is normalized to zero in 1991, representing the reference year. Each regression includes technology category and year fixed effects. Patent-processing data are from the PTO.

B. Difference-in-Difference Results

Basic Findings—To complement the above graphical analysis, we now explain our key hypotheses by presenting regression results for the basic difference-in-difference specification set forth in equation (1). As previously stated, we focus this initial inquiry on the period of time in which the PTO experienced strong-fee diversion and during which we predicted that the Agency would be especially inclined to turn to cost-minimizing measures to deal with budgetary shortfalls. The reported coefficients in Table 2 are that of the key difference-in-difference variable (β_1), capturing the extent to which the differential examination pendency length between high-cost and low-cost technologies will itself change in connection with an increase in the sustainability score—i.e., an improvement in the PTO's financial state. Considering that the hypothesis under consideration predicts an expansion in this differential upon a *decline* in the sustainability score, the above analysis predicts a negative estimate of this key coefficient. As demonstrated by Table 2, we indeed estimate coefficients of the expected sign.[40]

The regressions underlying Table 2 are structured to determine how the differential examination delays between high- and low-cost technologies will grow upon a 2 standard-deviation decline in the sustainability score. We chose to scale and interpret our results with respect to a decline in the financial health of the Agency of this magnitude in order to ensure that we compare how the differential examination delays between high- and low-cost technologies look during the PTO's "good times"—i.e., its lucrative periods—relative to its "bad times"—i.e., its financially strained periods.[41] To clarify further, the regression is structured in a manner that technically does not group technologies into simple

---

[40] Standard errors in Table 2 and throughout this analysis are clustered at the technology level to account for auto-correlation in examination pendency within technologies over time.

[41] In this instance, a shift of this nature represents a nearly 58% decline relative to the mean financial health of the Agency.

high-cost and low-cost bins; rather, it allows each technology to carry a unique cost level and asks how the impact of this negative financial shock to the Agency causes a divergence in examination delays as we move along this distribution of examination costs. We quantify this distributional movement by reference to a 2-standard-deviation shift in such costs. For similar reasons, a shift of this magnitude is meant to capture a meaningful level of separation between the costs of different technologies, facilitating our interpretation of the results as capturing the differential examination experiences of "high-cost" versus "low-cost" technologies.[42]

TABLE 2. THE IMPACT OF PTO FINANCIAL HEALTH ON DIFFERENTIAL EXAM DURATIONS (MONTHS) BETWEEN HIGH- AND LOW-COST TECHNOLOGIES (1991-2003)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SUSTAINABILITY * COST	-3.5*** (1.3)	-2.4** (1.1)	-3.2** (1.2)	-3.3** (1.2)	-2.7** (1.0)	-2.5** (1.1)	-2.2** (1.0)
N	481	481	481	481	481	481	481
Sustainability Score Construction	Version 1	Version 2	Version 3	Version 4	Version 5	Version 1	Version 1
Technology-Year Filing Rates?	NO	NO	NO	NO	NO	YES	YES
Other Technology-Year Covariates?	NO	NO	NO	NO	NO	NO	YES

* significant at 10% ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within technologies over time. All regressions include technology fixed effects and year fixed effects. The sustainability score and technology-specific cost measures are each normalized such that an increase in "1" of each normalized variable represents an increase of 2 standard deviations of the respective measure. Reported coefficients are for the difference-in-difference variable of interest and reflect the degree to which a 2-standard-deviation improvement in the sustainability score is associated with a change in the differential exam pendency between a technology with examination cost X and a technology with an examination cost equal to 2 standard deviations below X. Regressions are weighted by the number of first office actions associated with each technology-year cell. Version 1 of the sustainability score bases the examination demand facing the Agency on the total backlog of applications awaiting review. Version 2 calculates the score applying to a given technology-year cell by excluding contributions to the score from that technology-year cell. Version 3 bases the examination demand facing the Agency on the total backlog of applications awaiting review among the low-cost technologies only. Version 4 of the sustainability score excludes the average examination cost from the underlying calculation. Finally, Version 5 bases the examination demand facing the Agency on the total backlog of applications awaiting a first office action. Patent processing data is from the PTO.

[42] After all, a 2-standard deviation shift likely allows us to compare one side of the cost distribution with the other. In this instance, a 2-standard-deviation shift in examination costs represents a difference in costs of roughly 28% of the mean cost

With this guidance in mind, the estimated coefficient of -3.5 reported in Column 1 suggests that a 2 standard-deviation decrease in the sustainability score is associated with a 3.5 months-longer examination delay for technologies with cost *X* relative to those with a cost equal to 2 standard deviations below *X* (statistically significant at 1% level). In other words, when the Agency finds itself in a situation where it is more likely to struggle in breaking even, it will indeed begin to maintain longer examination delays in its costliest technologies, effectively prioritizing its less costly technologies.[43]

Alternative Sustainability-Score Formulations—For the sake of maintaining baseline consistency, we begin this analysis with the same sustainability score construction used in Frakes and Wasserman (2013), identified as "Version 1" in Table 2. In this subsection, we demonstrate the robustness of the above result to the construction of various alternative sustainability-score measures that are designed to address certain concerns specific to the present project.

At first blush, one may be concerned with an arguably mechanical connection between our dependent variable—i.e., examination pendencies—and our metric of PTO financial sustainability, considering that one key component of the sustainability-score is the aggregate number of applications awaiting PTO examination—i.e., backlog—a measure that may necessarily go hand in hand with longer examination delays. To be clear, however, our analysis is not simply observing the correlation between the sustainability score and overall examination pendencies, an exercise that would raise concerns over a mechanical relationship. Rather, our interest is in correlating the sustainability score with the *differential* in the examination pendency across technologies carrying different examination

---

[43] Note that the dynamic regressions underlying Figures 1 and 2 likewise capture "high cost" and "low cost" with reference to a 2-standard-deviation movement along the cost distribution across technologies.

costs. A negative association between these measures (as estimated) does not at all arise by construction. After all, a shock that leads to longer durations in low-cost areas would contribute to the total backlog (and thus to the sustainability score) just as much as a shock that leads to longer durations in high-cost areas.

To be clear, the fact that the Agency's backlog is a component of our metric of PTO financial health is rather fundamental to this analysis. As suggested in Section III, we view a PTO with a large aggregate backlog as one and the same as a PTO likely facing a binding budget constraint—i.e., it ostensibly lacks the resources to satisfy its obligations. Taking this as an input into our analysis, we then predict that the PTO will take the hypothesized actions in order to make this necessary backlog smaller than it otherwise needs to be. Nonetheless, even if this exercise does not pose a concern over a mechanical relationship that immediately explains the results, we may be concerned that the connection that nonetheless exists between examination pendencies and the sustainability score calculation threatens the contended exogeneity of the sustainability score measure. After all, consider a shock that leads to relatively longer examination delays in a high-cost technology—e.g., a labor market development that impacts the ability to hire examiners in that field.[44] Absent any offsetting developments, this shock would

---

[44] Consider for instance a general economic boom. With enhanced markets generally, the PTO may find itself in a position where it has trouble competing with outside entities for labor—i.e., for its examiners. Economic booms may be particularly stronger for some technologies relative to others—e.g., for our set of high-cost technologies, which happen to be particularly complex technologies. Whether we think of the unobservable shock of concern as something in general that could lead to a differential difficulty in hiring examiners in such complex fields or more specifically as a macroeconomic shock to the U.S. economy that leads to this differential hiring outcome, the analysis to follow will discuss various ways in which we attempt to rule out that the findings can be explained by such developments. At the outset, however, we should note that macroeconomic trends, whether measured by gross domestic product or unemployment rate, do not perfectly match up with the trends depicted in Figures 1—3, so as to heavily motivate this more specific concern in the first place. Following a recession in the early 1990's, the U.S. unemployment rate began to fall in 1993. Though this is still several years before the observed increase in the examination pendencies of high relative to low-cost technologies, this trend initially does line up relatively closely with the deterioration of interest in the PTO's financial health, serving as a potentially spurious explanation. However, the unemployment rate for the overall economy begins to rise in 2001. Under this spurious story of concern, one would predict that the differential trend in examination-pendencies across technologies that emerged in the middle-to-late 1990s would reverse in the early 2000s—i.e., now, it would be easier under the macroeconomic environment for the PTO to compete with outside markets for talent, especially in complex technologies. One would then predict that things would swing back in the other direction by the mid-2000s, as the economy begins to improve again. As demonstrated by Figures 1—3, however, the observed story of interest, correlating swings in the PTO's

likely lead to an increase in the aggregate backlog of examinations awaiting review and thus contribute to a lower sustainability score, potentially explaining our observed relationship—that is, our finding of an increase in high-cost examination delays relative to low-cost delays following sustainability declines. In other words, one may be concerned that the sustainability score is not independent from some of the unobserved determinants of our outcomes of interest, confounding identification.

Momentarily, we show that the above results are robust to various alternative approaches that are designed to remedy this connection. However, we first make two observations: (1) these considerations would not undercut the validity of our second natural experiment framework—i.e., observing the impact of the shift towards weak-fee diversion in 2004 (discussed below) and (2) any complications that arise from this connection are perhaps likely to be minor in the first place. After all, the sustainability score is an *aggregate* PTO measure and thus fluctuations in this score already abstract away from the influence of the more localized *technology-specific* examination decision under evaluation. As such, any substantial relationship observed between the aggregate measure and the technology-specific duration decision is perhaps more likely to originate from the influence of the sustainability score itself, as opposed to the other way around.

In our first attempt to alleviate the source of this possible feedback loop, we calculate a modified sustainability score ("Version 2"), which is applied to each technology-year observation in the regression in a manner that removes the influence of all factors specific to that technology-year cell from the sustainability score calculation. For instance, the score that applies to the Resins category in 1998 does not include the Resins-specific backlog count at the time. As this measure is still highly aggregate in nature, it continues to capture the PTO's

---

sustainability score with trends in the differential examination pendencies, does not demonstrate the pattern that one would predict if one thought that the findings attributed solely to fluctuations in the macroeconomic environment.

bottom line, while perhaps better ensuring the independence of the sustainability calculation from the technology-specific duration decision at hand. As demonstrated by Column 2 of Table 2, the results from this alternative approach do not alter the conclusions drawn above.

However, we do acknowledge that the unobservable development of concern—e.g., hiring difficulties—could be a shock that is common to all high-cost categories, as opposed to being specific to a given technology. In this case, it may not be enough simply to remove the influence of the specific technology under consideration.[45] As such, to rule out this final concern, we calculate yet another sustainability score ("Version 3"), whereby we entirely remove the influence of *all* high-cost-technologies from the calculation of the aggregate PTO backlog used in the annual sustainability score. With this calculation, we simply sever the potentially troubling link between the backlog measure used in the PTO's financial health indicator and the possibility of some unrelated development that leads to longer examination times in all high-cost examination groups. To the extent that the Agency is experiencing negative financial developments across the board (as is evidenced by our data),[46] it is arguably valid to use the developments of the low-cost groups as a proxy for the general financial status of the Agency and, in turn, to use this proxy in exploring the link between Agency financial health and examination practices. Encouragingly, as demonstrated by Column 3 of Table 2, this final alternative approach likewise generates substantially similar findings.

All told, there are numerous reasons to believe that the various coefficients estimated in Table 2 are reflective of examination-queuing responses to *external*

---

[45] This common shock that leads to longer examination periods in all of the high-cost categories could lead to a higher aggregate backlog, even when we specify that aggregate by removing the influence of the technology group associated with each observation.

[46] That is, when we calculate this modified sustainability score using the aggregate backlog among the low-cost technologies, we generate a time trend that looks strikingly similar to those presented in Figures 1 and 2. The benefit of this modified score is that it removes any feedback that can come from developments in the high-cost categories and that could complicate the inferences that we draw from observed increases in prosecution time among those high-cost groups.

financial conditions, as opposed to being reflective of other factors that simultaneously determine the outcome of interest and the sustainability metric. These results are consistent with a story in which, in the face of a growing aggregate backlog—i.e., a growing financial imbalance—the PTO will *subsequently* elect to prioritize its low-cost examinations in the queue in an attempt to process as much of this aggregate backlog under its limited resources.

Differential Application Rates—Building on the more general discussion from the previous subsection, we now consider a specific pattern of events that could spuriously explain the above results: a relative increase in the application rates within those technologies that cost the PTO the most to examine during the 1997 to 2004 period (when the sustainability score happens to be falling). Giving the possibility of rigidity in the PTO's ability to accommodate this differential application influx (e.g, through staffing reallocations), these developments may contribute to the differential in examination pendencies across technologies observed over this period. As demonstrated by Column 6 of Table 2, however, the above findings persist with the inclusion of controls for the share of initial application filings attributable to each technology, easing such concerns.[47]

We should note that the above concern is perhaps more than one of a mere spurious correlation between the sustainability score and a differential trend in application filings of this sort. Similar to the discussion in the previous subsection, this context likewise implicates a possible relationship between the sustainability score calculation itself and any such differential trend across technologies in application rates, threatening the independence assumption. After all, a shock in application rates to a high-cost technology could, all else equal, lead to both a relatively higher examination queue for that high-cost group and a

---

[47] The results also persist when controlling for the level of application filings within the relevant technology-year cell (available upon request) . Disaggregated filings data were likewise obtained from the PTO pursuant to a FOIA request. Initial filings exclude continuation filings.

lower sustainability score as a result of its contribution to a higher aggregate backlog of applications awaiting review. However, this is but a mere example of the problem discussed in the previous subsection, in which case the sustainability-score modifications discussed previously should alleviate such concerns.

In any event, we bring up this independence concern again as it takes on another dimension in this differential application-trend context. A shock that leads to an increase in application filings in a high-cost technology may not only lead to a decrease in the sustainability score through its contribution to an overall larger backlog of applications awaiting review. It will also lead to an increase in the average cost of those examinations awaiting the agency, a factor likewise built into the construction of the baseline sustainability score introduced in Frakes and Wasserman (2013). As another attempt to achieve exogeneity with our sustainability score metric and to address this final concern, we construct yet another alternative score ("Version 4") that simply removes this average examination cost / complexity factor from the calculation, essentially focusing the score on all other factors that disrupt the balance between the Agency's incoming fees and the costs associated with all of its examination obligations awaiting. As demonstrated by Column 4 of Table 2, the estimated coefficient of interest remains virtually unchanged with this modification, suggesting perhaps little need for concern in the first instance.

Other—In Column 7 of Table 2, we include as additional controls information on various observable characteristics of the average patent issued in each technology-year cell, including the number of patent claims and the number of citations made to these patents (a possible indicator of patent quality (Trajtenberg 1990, Hedge and Sampat 2009)), along with the percentage of patents in the technology-year combination that fall into each of eight different classifications of patentee types (i.e., individual, corporate or government; foreign or domestic). The results

likewise persist with the inclusion of such variables. Finally, in Column 5, we demonstrate the robustness of these results to the use of the alternative sustainability measure discussed in the graphical analysis above ("Version 5"), in which we use the aggregate inventory of applications awaiting a first office action as an estimate of the examination demand currently awaiting the Agency.

Fee-Diversion Interaction Specifications—As set forth in Section III and as discussed in the graphical analysis above, we predict that the PTO took advantage of the 2004 fee-diversion reform to shift its profitability-enhancing tactics more to its fee-generating practices during times of financial difficulty and away from its per-disposal cost-reducing tactics (i.e., low-cost prioritization). More specifically, we predict that any association between a negative shock to the PTO's financial health and the differential examination pendency between high- and low-cost technologies—i.e., the finding from Table 2—is likely to be weaker following 2004 than it was between 1991 and 2003.

TABLE 3. THE RELATIONSHIP BETWEEN PTO FINANCIAL HEALTH AND THE DIFFERENTIAL EXAM DURATIONS (MONTHS) BETWEEN HIGH- AND LOW-COST TECHNOLOGIES, INTERACTED WITH THE 2004 CONGRESSIONAL FEE-DIVERSION REFORM

	(1)	(2)	(3)	(4)	(5)	(6)
SUSTAINABILITY * COST	-3.9*** (1.2)	-3.1** (1.1)	-3.6*** (1.2)	-3.7*** (1.2)	-3.1*** (1.0)	-3.0** (1.1)
SUSTAINABILITY * COST * FEE_DIVERSION	11.2*** (3.9)	5.3** (2.6)	13.2** (5.5)	10.4*** (3.6)	17.0** (8.2)	10.0*** (3.3)
N	666	666	666	666	666	666
Sustainability Score Construction	Version 1	Version 2	Version 3	Version 4	Version 5	Version 1
Technology-Year Filing Rates?	NO	NO	NO	NO	NO	YES

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within technology categories over time. All regressions include technology fixed effects and year fixed effects. The sustainability score and technology-specific cost measures are each normalized such that an increase in "1" of each normalized variable represents an increase of 2 standard deviations of the respective measure. Reported coefficients are for: (1) the difference-in-difference variable of interest and (2) the

interaction of that variable with an indicator for being in the weak fee-diversion period. While the coefficients are not reported, the underlying regressions also include the other constitutive elements of these interactions. Regressions are weighted by the number of first office actions associated with each technology-year cell. Version 1 of the sustainability score bases the examination demand facing the Agency on the total backlog of applications awaiting review. Version 2 calculates the score applying to a given technology-year cell by excluding contributions to the score from that technology-year cell. Version 3 bases the examination demand facing the Agency on the total backlog of applications awaiting review among the low-cost technologies only. Version 4 of the sustainability score excludes the average examination cost from the underlying calculation. Finally, Version 5 bases the examination demand facing the Agency on the total backlog of applications awaiting a first office action. Patent processing data is from the PTO.

To test this prediction in a basic regression framework, we estimate the specification set forth in equation (3) and focus our attention on β_4, representing the coefficient of the interaction between the main difference-in-difference variable (COST*SUSTAINABILITY) and a dummy variable for being in the post-2003 period. As discussed in Section III, the above prediction suggests a positive estimate for this coefficient. As presented in Column 1 of Table 3, we do indeed estimate a positive coefficient of this nature, confirming that the distortionary practice of delaying relatively costlier examinations during negative financial shocks either slowed down, halted, or reversed itself following a shift towards a weaker fee diversion regime.[48] As presented in Table 3, these findings are robust to the various alternative formulations of the sustainability score discussed above and to the inclusion of controls for the share of total initial applications attributable to each technology (for each year).[49]

Robustness Checks—In addition to those already mentioned, we subject these findings to a range of specification checks. To begin, we estimate the above

---

[48] In fact, the magnitude of the estimated coefficient suggests a trend reversal, consistent with the graphical analysis. The main 1991-2003 difference-in-difference findings that are implicated by Table 3 (*i.e.*, β_1) correspond closely in magnitude with those derived in Table 2 and suggest that a 2 standard-deviation decline in the sustainability score is associated with a 3.8 months-longer examination pendency for high-cost relative to low-cost technologies. For the post-2004 period, on the other hand, such a decline in the sustainability score is associated with a 7.3 months-*shorter* examination pendency for high-cost relative to low-cost technologies (*i.e.*, $\beta_1 + \beta_4$).

[49] Since the remaining covariates—e.g., number of claims—are available only in the pre-2006 period, we do not include them in this analysis.

pattern of results after dropping each technology category one by one.[50]   This pattern is virtually unaffected by this sequential exercise, confirming that our findings cannot be explained by particular developments in any one technology. In Online Appendix B, we further demonstrate the robustness of our findings to the following:

- the inclusion of technology-year controls capturing the intensity of usage of certain PTO practices bearing on prosecution times: (1) "restriction requirements" (2) "requests for continued examinations" / "continuing prosecution applications" and (3) traditional continuation applications;

- the specification of the examination-delay measure as the average length of time between the filing of an application and the ultimate issuance of a patent;

- the use of both the natural log of the examination duration and the growth rate in the examination duration (relative to the base year) as the operable dependent variable;

- less parametric treatment of examination costs across technologies—i.e., exploring how the PTO responds to more categorical increase in costs;

- the inclusion of technology-specific linear time trends;

- the use of a more flexible randomization-inference approach.

Moreover, Online Appendix B further demonstrates the robustness of the above graphical analysis to the inclusion of various technology-year control variables.

---

[50] Hegde (2012) notes that the PTO's biotechnology and business method units in the 1990s may have faced particular difficulties in retaining examiners. To address concerns that the results are attributable to these collective developments, we also note that our results are robust to the joint exclusion of both NBER subcategories covering such units.

VI. DISCUSSION AND CONCLUSION

A. General Implications

Two areas of patent policy that have received substantial attention by academics and commentators over the years involve (1) possibly excessive granting practices on the part of the PTO and the consequent watering down of the quality of patents issued and (2) the mounting backlog of examinations awaiting review by examiners and the harms associated with the consequent delays. Such matters are of significance whether we speak of them in the aggregate or in terms of heterogeneity across technologies. Building on Frakes and Wasserman (2013), the present analysis has suggested that the mere nature of the PTO's user-fee structure and its general examination cost structure may play an important role in the discussions surrounding these key patent-system outcomes.

B. Implications of the America Invents Act

The possible influence of the PTO's funding structure on its examination queuing practices is perhaps of even greater concern moving forward as a result of certain features of the recently-enacted America Invents Act (the AIA), arguably the most significant modification to the patent system since 1952. Among other things, the AIA grants the PTO, for the first time, explicit authority to prioritize the examination of inventions of "national importance or national competitiveness."[51] The open-ended policy language associated with this prioritization authority gives the PTO almost unfettered discretion to decide which categories of patent applications should receive expedited review, with nothing in the language of the AIA explicitly forbidding or discouraging the PTO

---

[51] Specifically, the AIA states the PTO: may "provide for prioritization of examination of applications for products, processes, or technologies that are important to the national economy or national competitiveness without recovering the aggregate extra cost of providing such prioritization…" America Invents Act sec. 25 (to be codified at 35 U.S.C. § 2(b)(2)(G)). While the AIA also allows for up to ten thousand patent applications a year to receive expedited processing upon the payment of a $4800 fee, the PTO's Prioritization Authority enables it to expedite the processing of patent applications without the payment of any additional fees.

from taking into account the cost of patent processing when exercising its prioritization authority. To the extent that examination-cost considerations may play a role in prioritization discussions within the Agency, our findings raise concern over whether this provision of the AIA will fulfill its legislative purpose. Though our findings suggest that these distortionary prioritization practices reversed course in 2004, it remains possible that the Agency could reinvigorate them moving forward, especially now that they have brought such practices back to a more balanced baseline and especially given the possibility that Congress could revive a strong form of fee diversion.

C. Inherent Risk in PTO Fee Structure

This analysis has largely focused on those behaviors that the PTO may partake in during instances of binding budget constraints. The PTO's fee structure bears on the choices that the Agency may make in such instances in light of certain particularities associated with that structure—i.e., the fact that some types of patents garner higher fees relative to others and the fact that all technologies pay the same examination fees despite the fact that some technologies clearly cost more to examine relative to others.

However, the Agency's fee structure is relevant for our analysis in a perhaps more fundamental way. We began our discussion with the more critical observation that the PTO's fee structure creates an inherent risk of financial instability that may lead to situations of binding budget constraints (necessitating the above responses) in the first instance. That is, we contend that the inadequacies of the examination fees coupled with the subsidization of the examination process by fees assessed on those who have already successfully navigated that process creates a financial risk that the Agency's incoming fee revenue may not be sufficient to meet its examination demands.

This basic observation suggests that, in addition to modifying those particularities of the fee-structure discussed above—e.g., setting technology-specific examination fees in proportion to examination costs for that technology—the patent-examination distortions under discussion in this paper and in Frakes and Wasserman (2013) would be less paramount under a funding structure that imposed no such inherent financial risk. One such alternative structure would be one where examination fees came closer to covering the costs of patent processing (which constitute the vast majority of the Agency's operational costs), a structure that would reduce the need for cross-subsidization and thus reduce the risk that the parties being used for cross-subsidization purposes would grow out of step with applicants.

Of course, perhaps more is at stake in setting a fee structure than just the incentives of the Agency. Arguably the biggest drawback from a structure that were to set examination fees closer to costs would be the loss of any subsidization of patent filings that comes through lower fees. Of course, studies to date have suggested that patent filings are rather price inelastic (de Rassenfosse and van Pottelsberghe de la Potterie 2012; Wilson 2008). Moreover, the actual fees paid to the PTO itself for the examination of a patent application are a fraction of the overall legal and preparation cost of securing a patent.[52] Ultimately, it is unclear the extent to which social welfare would be jeopardized through less subsidization of the patent applications.

In addition, weaker reliance on patent renewal / maintenance fees as a source of cross subsidization may come at the expense of the non-funding-related objectives of renewal fees. Maintenance fees are believed to serve an important function in effectively shortening the lifetime of certain patents and hence may decrease the

---

[52] A utility patent application of minimal complexity costs on average $10,000 to prepare and prosecute while a utility patent application of relatively high complexity costs on average $16,000 to prepare and prosecute (Law Practice Management Committee 2009).

static costs associated with patents. Of course, it need not be the case that bringing examination fees closer to costs must carry with it the extinguishment of renewal fees altogether. Alternative structures could maintain such fees and simply preclude the Agency from using them for funding purposes (e.g., depositing them to general revenues), or could use them as a direct source of subsidization of applications (as distinct from using lower fee levels as the mechanism of subsidization). A structure of this latter variety might allow the Agency to achieve both of the non-funding functions it hopes to achieve with its current structure without sacrificing the ability to absorb demand shocks through the imposition of cost-based examination fees.[53]

Ultimately, greater research is warranted on the question of the optimal fee structure of the PTO. This research may be even more valuable now than ever as a result of another feature of the recently passed AIA, which provides the PTO, for the first time, the authority to set the structure and level of its fees through the administrative rulemaking process (a task previously retained by Congress). The PTO has recently exercised this new authority, though in a manner that retains the same heavy reliance on post-allowance fees to cross-subsidize the cost of examinations, thereby maintaining the financial vulnerability of the past.[54]

---

[53] Of course, using renewal fees as an immediate source of funds to subsidize applicants would effectively transfer the risk of incurring a relative decline in renewal fees from the Agency (under the status quo) to the applicants themselves.

[54] Online Appendix C provides greater details on the PTO's first exercise of its fee-setting authority. Importantly, the time required to change fee levels through the rulemaking process may leave the Agency in a position where it may turn to the distortionary practices discussed in this paper to restore any temporary imbalances.

REFERENCES

Acemoglu, Daron and Amy Finkelstein. 2007. "Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector." *Journal of Political Economy*, 116 (5): 837-80.

Bessen, James E. and Brian J. Love. 2013. "Make the Patent "Polluters' Pay: Using Pigouvian Fees to Curb Patent Abuse." *California Law Review Circuit*, forthcoming.

Carpenter, Daniel, Evan J. Zucker, and Jerry Avorn. 2008. "Drug-Review Deadlines and Safety Problems." *New England Journal of Medicine* 358: 1354-1361.

Cockburn, Iain, Samuel Kortum, and Scott Stern. 2003. "Are All Patent Examiners Equal? Examiners, Patent Characteristics, and Litigation Outcomes." in Patents in the Knowledge-Based Economy: 46-47.

de Rassenfosse, Gaetan and Bruno van Pottelsberghe de la Potterie. 2007. "Per Un Pugno Di Dollari: A First Look at the Price Elasticity of Patents." *Oxford Review of Economic Policy* 23: 558-604.

de Rassenfosse, Gaetan and Bruno van Pottelsberghe de la Potterie. 2013. "The Role of Fees in Patent Systems: Theory and Evidence." *Journal of Economic Survey* 27: 696-716.

Federal Trade Commission. 2003. *To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy*.

Frakes, Michael D. and Melissa F. Wasserman. 2013. "Does Agency Funding Affect Decisionmaking?: An Empirical Assessment of the PTO's Granting Patterns." *Vanderbilt Law Review*, 66: 67-125.

Gillette, Clayton P. and Thomas D. Hopkins. 1987. "Federal User Fees: A Legal and Economic Analysis." *Boston University Law Review* 67: 795-874.

Gillette, Clayton P. 1988. "User Fees at the Environmental Protection Agency." Pp. 19-27 in *User Fees and Environmental Protection A Legal Perspective on User Fees*, edited by Thomas D. Hopkins. Washington, D.C.

Griffin, Charles. 1987. "User Charges for Health Care in Principle and Practice." Economic Development Institute, Seminar Paper No. 37. The World Bank, Washington D.C.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2002. "The NBER Patent-Citations Data File: Lessons, Insights, and Methodological Tools." Pp. 403-460 in *Patents, Citations, & Innovations: A Window on the Knowledge Economy*, edited by Adam B. Jaffe & Manuel Trajtenberg.

Hegde, Deepak. 2012. **"**Funding and Performance at the US Patent and Trademark Office." 30 *Nature Biotechnology*: 148-150.

Hegde, Deepak and Bhaven Sampat. 2009. "Examiner citations, applicant citations, and the private value of patents." *Economics Letters*, 105(3): 287-289.

Jaffe, Adam B. and Josh Lerner. 2004. *Innovation and Its Discontents.* Princeton, NJ.: Princeton University Press.

Law Practice Management Committee. 2009. American Intellectual Property Law Association, *Report of Economic Survey*.

Lemley, Mark A. 2001. "Rational Ignorance at the Patent Office." *Northwestern University Law Review*, 95: 1495-1532.

Lemley, Mark A. and Bhaven Sampat. 2008. "Is the Patent Office a Rubber Stamp?" *Emory Law Journal,* 58: 181-203.

Lemley, Mark A. and Bhaven Sampat. 2012. "Examiner Characteristics and Patent Office Outcomes." *Review of Economics and Statistics*, 94(3): 817-827

Lichtman, Doug and Mark A. Lemley. 2007. "Rethinking Patent Law's Presumption of Validity." *Stanford Law Review*, 60: 45-72.

London Economics. 2010. "Economic Study on Patent Backlogs and System of Mutual Recognition: Final Report to the Intellectual property Office," *available at* http://www.ipo.gov.uk/p-backlog-report.pdf.

Magliocca, Gerard N. 2007. "Blackberries and Barnyards: Patent Trolls and the Perils of Innovation." *Notre Dame Law Review*, 82: 1809-1838.

Mann, Ronald. 2013. "The Idiosyncrasy of Patent Examiners: Effects of Experience and Attrition." Mimeo.

Merges, Robert P. 1999. "As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform." *Berkeley Technology Law Journal*, 14: 577-615.

Moser, Petra. 2007. "Why Don't Inventors Patent?" Working Paper No. 13294. National Bureau of Economic Research, Cambridge, Mass.

Office of Management and Budget. 2002. Budget of the U.S. Government Fiscal Year 2002.

Olson, Mary K. 2004. "Managing Delegation in the FDA: Reducing Delay in New Drug Review." *Journal of Health Politics, Policy, and Law*, 29: 397-430.

Pakes, Ariel. 1986. "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica,* 54: 755-784.

Schankerman, Mark and Ariel Pakes. 1986. "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period." *The Economic Journal*, 96: 1052-1076.

Sepehri, Ardeshir and Robert Chernomas. 2001. "Are User Charges Efficiency- and Equity-Enhancing? A Critical Review of Economic Literature with Particular Reference to Experience from Developing Countries." *Journal of International Development*, 13: 183-209.

Thomas, John R. 2001. "Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties." *University of Illinois Law Review*, 305: 316-353.

Trajtenberg, Manuel. 1990. "A Penny for your Quotes: Patent Citations and the Value of Innovations." *RAND Journal of Economics*, 21: 172-187.

U.S. Patent and Trademark Office. 1993. Annual Report, Fiscal Year 1993.

U.S. Patent and Trademark Office. 1994. Annual Report, Fiscal Year 1994.

U.S. Patent and Trademark Office. 2002. Performance and Accountability Report Fiscal Year 2002.

U.S. Patent and Trademark Office. 2008. Performance and Accountability Report Fiscal Year 2008.

U.S. Patent and Trademark Office. 2012a. Performance and Accountability Report Fiscal Year 2012.

U.S. Patent and Trademark Office. 2012b. Proposed Rule Setting and Adjusting Patent Fees. *Federal Register* 77:55028-55078.

U.S. Patent and Trademark Office. 2013. Final Rule Setting and Adjusting Patent Fees. Federal Register 78: 4213-4284.

Watson, Traci. 1992. "Patent Office Drops Plan to Raise Fees." *Nature*, 356: 645-646.

MODEL

A. Basic Framework

In this subsection, we introduce a simple, relatively informal model of PTO decisionmaking under financial constraints. We begin by setting forth the Agency's utility:

$$W_t = \sum_{t=1}^{\infty} \sum_{i=1}^{37} V\left(x_{it}, g_{it} - \delta_{it}, d1_t, d2_t\right)\beta^{-t}$$

where x_t is the number of applications processed at time t. In other words, the PTO places value on the processing of applications. Moreover, a decision to process an application received at time *t* at a later point in time (e.g., time *t+1*) will cause the ultimate value to be received for that application to be discounted by the appropriate discount factor (e.g., β^{-1} in the case of an application delayed by 1 year). The rate at which the PTO grants applications processed at time period *t* for technology group *i* is captured by g_{it}, while the rate at which proper innovation policy suggests that the PTO *should* grant such applications (given their quality) is captured by δ_{it}. A deviation from the proper grant rate in either direction generates a utility loss for the PTO. Finally, the PTO places some disutility on any differential treatment that it extends in either its examination queuing practices or its granting practices (i.e., deviating from proper grant rates more in the case of some technologies than other technologies). Measures capturing the degree to which the PTO engages in such disparities are captured by *d*1 (with respect to its examination queuing practices) and *d*2 (with respect to its granting practices). More specifically,

$$d1_t \sim f(b_{1t}, b_{2t}, b_{3t}, \dots b_{37t})$$

where b_{it} is the backlog of applications awaiting review in technology group *i* in year *t*. In other words, $d1_t$ is some function of the set of technology-specific application backlogs, which are themselves derived from the number of previous applications, *a*, that have yet to be disposed of by the agency. That is, the backlog for technology *i* at time *k* equals:

$$b1_k = \sum_{t=-\infty}^{t} (a_{it} - x_{it})$$

Moreover,

$$d2_t \sim f(g_{1t} - \delta_{1t}, g_{2t} - \delta_{2t,}, \dots g_{37t} - \delta_{37t})$$

That is, d2$_t$ is some function of the set of technology-specific deviations in grant rates from otherwise proper grant rates.

To summarize, the PTO's objective function is enhanced to the extent that it processes more applications and to the extent that it does so in a timely manner. Moreover, the PTO's objective function is decreased in the event that the PTO distorts its practices by either granting at excessive rates or inadequate rates (relative to the optimal innovation policy benchmark) or by extending differential examination treatment across technologies / patent types.

<div align="center">

B. Predictions

</div>

The PTO attempts to maximize the above utility through its determination of (1) the set of x's for all current and future time periods and for each technology—i.e., the count of the technology-specific applications processed each year and (2) the set of g's for all current and future time periods and for each technology—i.e.,, the technology-specific rate by which the PTO will grant those x_{it} applications that it processes each year. Note that, at time t, the PTO cannot effectively determine the actual number of applications to be processed each year for all future years, not knowing the future realizations of incoming applications. Nonetheless, we model the PTO's objective in this dynamic framework given that the PTO may elect to process some portion of the applications that it receives at time t, a_t, at a later time period. At any given time, of course, the number of applications that the Agency can examine is subject to the constraint that it is less than the number of applications received that year plus the number of examinations awaiting examination at the beginning of that year (i.e., its backlog):

$$\sum_{i=1}^{37} x_{it} \leq \sum_{i=1}^{37} a_{it} + \sum_{k=-\infty}^{t-1} \sum_{i=1}^{37} (a_{ik} - x_{ik})$$

Technically, this constraint also holds on a technology-specific basis:

$$x_{it} \leq a_{it} + \sum_{k=-\infty}^{t-1} (a_{ik} - x_{ik})$$

Of course, the PTO, being funded through its user-fee collections, is subject to one important, final constraint. That is, we model the Agency as maximizing the above utility subject to (in addition to the above constraints) a resource constraint that it must meet each period, t.

$$I_t \geq C_t, \quad \forall t$$

or

$$\sum_{i=1}^{37} x_{it} * f1 + S_t * f2 + \sum_{i=1}^{37} x_{it} * g_{it} * f3 \geq \sum_{i=1}^{37} c_{it} * x_{it}, \quad \forall t$$

where *f1* corresponds to the per-application fee collected by the PTO at the time of examination; *f2* corresponds to the per-patent maintenance fee paid by the existing stock, *S*, of patent grantees whose patents are still active and who have elected to renew their patents (for simplicity purposes, we treat renewal fees as being paid annually); and *f3* corresponds to the per-unit fee paid at the issuance of a patent. Consistent with the reality of the PTO's budgetary process over the course of our sample period, we do not model that PTO as having the ability to change fee

levels. The stock of previous patent grantees at time t from which the PTO derives maintenance fees is determined as follow (simply assuming a 20-year patent term):

$$S_t = \sum_{i=1}^{37} \sum_{k=t-20}^{t} (x_{ik} * g_{ik}) * m_{ik}$$

Patents will stay active unless the patent-holder elects not to renew them. Patent-holders renew their patents at rate m, which vary across technologies and over time.

PROPOSITION 1: Absent its resource constraint, the PTO maximizes its utility, W, by setting $x_{it} = a_{it}$, $\forall i, t$.

In other words, unconstrained, the PTO will process all applications when they arrive—i.e., they will not run a backlog of applications. One can establish this relatively straightforwardly by considering just a single application filed at time t. The PTO is ultimately obligated to review this application and will at some point register the utility associated with that review. However, considering that the value of any future review will be discounted appropriately, the PTO would be better off reviewing this application (and thus all incoming applications) at the time of arrival.

PROPOSITION 2: In the face of its resource constraint, the PTO may reach its unconstrained optimum—i.e., setting $x_{it} = a_{it}$, $\forall i, t$, -- assuming that $f1_i = c_i$, $f2 = 0$, and $f3 = 0$.

In other words, if fees (at least those fees accessible to the PTO for funding purposes) are only set on examination and at a level equal to the cost of the associated examination, the PTO's budget constraint will never bind and the Agency will be able to satisfy all of the examinations demanded of it each period with the incoming fees at its disposal. Establishing this is likewise relatively straightforward. Setting $f1_i = c_i$ and substituting this into the $f1$ term, while also setting $f2$ and $f3$ equal to 0, the resource constraint becomes:

$$\sum_{i=1}^{37} x_{it} * c_{it} \geq \sum_{i=1}^{37} c_{it} * x_{it}$$

This constraint now becomes satisfied by design. The outcome is intuitive. Unanticipated increases in application rates do not threaten the ability of the PTO to raise enough funds to cover the associated increase in costs given that revenues will increase at the same rate.

PROPOSITION 3: When $f2$ and $f3 > 0$ and $f1 < c$, there may be occasions when the PTO's resource constraint cannot be satisfied while attempting to set $x_{it} = a_{it}$, $\forall i, t$.

In other words, when the PTO sets examination fees less than examination costs and cross-subsidizes examination through the imposition of issuance and renewal fees, the PTO faces a risk that its budget constraint will bind and that it will be unable to undertake all examinations upon arrival (thus preventing it from achieving its unconstrained optimum). To show the conditions under which this may occur, we make the simplifying assumption (without loss of generality) that the PTO cannot distinguish across technologies. Its budget constraint thus becomes:

$$x_t * \text{f1} + S_t * \text{f2} + x_t * g_t * \text{f3} \geq x_t * c_t$$

Now, assume that at time t, this holds with equality when the PTO sets $x_t = a_t$. That is, assume as a baseline that the PTO just has enough incoming funds to satisfy all incoming applications. Over the next year, a_t will grow at rate $h(z)$ and will thus equal $a_t(1 + h(z))$ at time $t+1$. Assume that the PTO attempts to set $x_{t+1} = a_t(1 + h(z))$ and thus tries to review all applications that subsequent year as well. The PTO will be able to do this as long as:

$$(a_t(1 + h(z))) * \text{f1} + S_{t+1} * \text{f2} + (a_t(1 + h(z))) * g_{t+1} * \text{f3} \geq (a_t(1 + h(z))) * c_{t+1}$$

or

$$a_t(1 + h(z)) \leq S_{t+1} * \text{f2} / (c_{t+1} - g_{t+1} * \text{f3} - \text{f1})$$

or

$$h(z) \leq S_{t+1} * \text{f2} / (a_t * (c_{t+1} - g_{t+1} * f3 - \text{f1})) - 1$$

For realizations of $h(z)$ less than the right hand side, the PTO will still be able to process all applications in time $t+1$. Realizations of the application growth rate greater than this amount suggest that the PTO will not be able to examine the full application count at this time while satisfying its budget constraint. To state the above condition in terms of the relationship between the application growth rate and the growth rate in the stock of prior patent grantees, let us make the simplifying assumptions that the appropriate grant rates do not change over time, nor do application costs (i.e., $g_{t+1} = g_t$ and $c_{t+1} = c_t$).

As above, to isolate those conditions under which the PTO's financial balance would tip in an unfavorable direction, we assumed that the budget constraint would hold with equality at time t. As such, taking the time t budget constraint and solving for S_t, we find:

$$S_t = a_t(c_t - g_t * f3 - \text{f1}) / \text{f2}$$

Plugging this back into the above budget constraint for time $t+1$, we find:

$$h(z) \leq S_{t+1} / S_t - 1$$

or

$$h(z) \leq (S_{t+1} - S_t) / S_t$$

In other words, starting from a position in time t where the Agency collects just enough income to be able to process all incoming applications at that time, the PTO will likewise be able to review all incoming applications at time $t+1$ as long as the growth rate in applications over this year is less than the growth rate in the stock of patent grantees from which the Agency collects renewal fees While we do not show this in the present paper, it is relatively straightforward to demonstrate that the budgetary risks posed by the reliance on renewal fees are greater the more the Agency relies on such fees to cover examination costs—i.e., the higher proportion of c that is covered by $f2$.

PROPOSITION 4: When the PTO's budget constraint binds and when its granting rates, g_t, are constrained to follow proper grant rates, δ_{1t}, the PTO will attempt to satisfy its budget constraint while nonetheless maximizing the number of applications that it is able to process at time t, x_t, by decreasing average per-unit application processing costs, $\frac{\sum_{i=1}^{37} c_{it} * x_{it}}{x_t}$ through the prioritization of low-cost technologies in selecting the x_t applications to review at time t.

For now, assume that the Agency cannot alter its grant rates. Likewise, consider that we are in a situation at time t in which the Agency's budget constraint binds—i.e., $x_t^* < a_t$. As such, it is necessary for the PTO to delay the processing of some applications at this time. The question we now is explore is: which applications to put off? To begin, let us start with simple assumption that the PTO's objective is to process as many applications as it can this period—i.e., to make $\sum_{t=1}^{37} x_{it}$ as large as it can subject to the relevant constraints.

By construction, we know that the PTO cannot examine all a_t applications right now. With the first application that the Agency will decide to defer until a later time period, the Agency will obviously lose examination fees right now, along with the possibility of issuance fees this year. For the sake of simplicity, assume that the quality of applications is generally the same across technologies, in which event the decision to defer this first application will have the same impact on lost fee revenue this period across the various technologies. The decision as to which technology to defer is not inconsequential, however, with respect to examination costs, which vary across technologies. As such, even though the choice as to which examination to defer has no bearing on the left hand side of the budget constraint, it does bear on the right-hand side. The less weight that the Agency places on applications that carry low costs, the lower the Agency's aggregate costs will be at time t and thus, all else equal, the closer the Agency's budget constraint will come towards satisfaction. As such, if the Agency wants to quickly bring itself back into financial balance with the number of examinations that it would like to examine—i.e., as many as possible—it will one-by-one focus any necessary examination delays on applications from those technologies that cost the most to examine. Once it lowers its aggregate costs enough through this selection mechanism and reaches the point where its incoming fees can now satisfy the outgoing costs, it can stop this selective delay process.

Consider the choice of the marginal application that just brings the Agency back into balance. Under the above selection approach, this marginal application is the costliest application out of

the set of remaining applications facing the examine-now-versus-delay decision. To demonstrate the benefits of selection, if the PTO were to have selected the cheapest remaining application as opposed to the costliest remaining application, the PTO could then be in a position where its aggregate costs still surpass its incoming fee collections. Thus, even though the PTO could have stopped its delay process by choosing the most expensive remaining application, it will now need to delay at least 2 remaining applications in order to break even. This undermines the Agency's objective function and its desire to process as many applications as it can.[1]

Of course, as specified initially, the PTO also places a disutility on increases in $d1_t$, which represents the degree to which the PTO differentiates its examination delays across technologies. Without specifying any functional forms for this metric and for the values placed on it, we simply note that the PTO may face a tradeoff between: (1) the value that it places on the aggregate number of applications that it can process, which otherwise motivates it to preference some technologies, and (2) the harm that it associates with such heterogeneity in examination practices. We simply note the possibility of an interior solution being met in which the PTO will engage in some prioritization, though not to an extreme degree.

PROPOSITION 5: When the PTO's budget constraint binds, the PTO will further attempt to satisfy its budget constraint while nonetheless maximizing the number of applications that it is able to process at time t, x_t, by increasing g_{it} among those technologies / patent types that are associated with higher issuance fees, $f3$, and/or higher renewal rates, m_i.

We now briefly consider another tactic that the PTO may undertake to dealing with a binding budget constraint and now allow the PTO's selection of its grant rate, g_{it}, to deviate from the otherwise proper grant rate, δ_{it}. Quite simply, considering that the PTO's incoming fee stream increases linearly with its grant rate—i.e., that $I'(g) > 0$—the PTO stands to improve its financial position and increase the chances that its budget constraint will be satisfied by granting patents at a higher rate and thus increasing g_t. Of course, we model the PTO's objective function as placing a disutility on the degree to which its grant rate detracts from otherwise proper innovation policy—i.e., on the degree to which g_{it} and δ_{it} differ. For the purposes of brevity, we do not fully specify this analysis. Rather, we simply note the existence of this tradeoff and suggest that the PTO's dual desires of (1) increasing its revenue to allow it to process more applications while nonetheless breaking even and (2) limiting the disutility placed on the deviation of g_{it} from δ_{it}, will cause the Agency to target its distortionary granting practices on those types of applicants that garner higher post-allowances fees, $f2$ and $f3$. As discussed in greater detail in Frakes and Wasserman (2013), these high-fee-generating patent types include: (1) large entities (which pay double the fees relative to small entities) and (2) technologies that consistently renew their patents at high rates, m_i, thereby contributing more to S_t.

---

[1] As another way to suggest this selection desire on the part of the PTO, consider its budget constraint again. If the PTO were to reduce c_t, then all else equal, the budget constraint would be more likely to be satisfied and thus the Agency could find it easier to process all of the applications that it would like to examine now. As discussed in the paper, it is difficult for the PTO to achieve this across the board. However, it can effectively achieve this through selection, by focusing its efforts on the least costly applications.

Of course, as with the examination prioritization discussion, the PTO also places disutility on increases in *d2* –that is, on the degree to which the Agency's deviation from proper granting practices itself takes on a differential scale across technologies/patent types. Again, without specifying this tradeoff, we simply note that the PTO may limit the degree to which it focuses its excessive granting practices on the most lucrative applicant types. However, to the extent that an interior solution is met, the PTO may nonetheless engage in some level of targeted over-granting.

PROPOSITION 6: Following 2003, at which time Congress lifted caps that it had effectively placed on PTO fee collections, the PTO became relatively more inclined to respond to periods of binding budget constraints by increasing its grant rates, g_t; furthermore, in the post-2003 period, the PTO became relatively less inclined to responds to periods of binding budget constraints by decreasing its average processing costs, $\frac{\sum_{i=1}^{37} c_{it} * x_{it}}{x_t}$, through selective examination.

Finally, we introduce one additional aspect of policy relevant to this decisionmaking context. In the period of time prior to 2004, as discussed further in the text and in this Online Appendix, Congress effectively capped the PTO's fees at a level less than their anticipated fee revenues for the year, essentially imposing an additional constraint in the above model:

$$\sum_{i=1}^{37} x_{it} * f1 + S_t * f2 + \sum_{i=1}^{37} x_{it} * g_{it} * f3 \leq L_t$$

where L_t is a congressionally imposed cap on fee collections that varies annually between 1991 and 2003. This alters the above analysis insofar as the Agency's inclination to grant at higher rates to high-fee-generating patent types during times of binding budget constraints will likely operate to only a limited degree—that is, only up until the point where this cap is reached—over this time period. During such times, the PTO may be especially inclined to prioritize low-cost examinations as the mechanism to satisfy its budget constraint. At 2004, when these caps were lifted, the PTO had already amassed a long backlog of awaiting applications in its high-cost technology areas after years of low-cost preferential treatment—that is, *d1* had grown over this time period. All else equal, the PTO's utility will increase as this differential backlog, *d1*, shrinks. Following 2004, the PTO can achieve this reduction in disutility associated with a reduced *d1* level without sacrificing otherwise valuable throughput of patent examinations by taking advantage of the fee revenue it could obtain through preferential granting practices. Of course, such differential practices carry their own disutility, insofar as they lead to both disutility in deviating from proper grant rates and disutility in differentiating granting practices across technologies—i.e., an increase in *d2*. To the extent that the benefits from reducing *d1* outweigh the harms of increasing *d2* and from deviating g_t from δ_t, the PTO may indeed take advantage of this shift in policies to extend preferential examination treatment to high-cost technologies for a brief period of time in order to bring the technology-specific backlogs (i.e., the $b_{it}'s$) into greater balance.

BACKGROUND ON PTO BUDGETARY PROCESS

In the text, we note the PTO has not been granted access to its full fee collections. Rather, Congress has routinely utilized a portion of the Agency's fee income to fund other government operations, a practice referred to as fee-diversion. The incidence of fee diversion has varied over time with Congress relying on the practice more heavily prior to 2004 than it did thereafter. As a result, we characterize the time period between 1991–2003, wherein Congress routinely set the Agency's budget to a level that was essentially *below* both its estimated and actual fee collections, as a period of "strong-fee diversion." In contrast, we characterize the period of 2004 onwards, wherein the PTO's spending authority has been capped *at* its projected revenue stream, as a period of "weak-fee diversion."

More specifically, Congress's diversion of PTO funds during 1991-2003 took on one of two forms. First, during the fiscal years of 1991 to 1998, Congress limited the PTO's ability to spend a certain subset of fees known as surcharge revenue. In 1991, when Congress made the agency essentially user fee funded it concomitantly enacted a 69% surcharge on certain patent fees.[2] From the fiscal years of 1991 to 1998, non-surcharge fees collected from users were fully available to the PTO; however, surcharge revenue was not. In other words, if the PTO were to attempt to raise fee revenues—e.g., through excessive granting practices—it would be forced to share a portion of these returns with Congress (potentially blunting its incentives to take such actions. In 1999 the surcharge fees expired. Congress, however, did not allow the Agency to spend its entire fee collections at this time, instead it choose to limit the PTO's ability to spend its fee collections by some predetermined amount. More specifically, from the fiscal years of 1999 to 2003, Congress made certain dollar amount of fees unavailable for PTO use each year and then, in differing amounts over the years, allowed the agency to use some, but not all, prior year fees. The result, however, was that for the fiscal years of 1999 to 2003, the agency's budget was essentially set to below its estimated fee collections.[3] For example, in the fiscal year of 2002 the PTO estimated its fee collections would total 1,346 million dollars.[4] That year, Congress chose to set the office's budgetary resources for spending to 1,146.7 million dollars, of which 282.3 million was from fees collected in fiscal years 2000 and 2001 and 843.7 million dollars from fees collected during fiscal year 2002. In the fiscal year of 2002, 304.1 million dollars from fees collected that year were not available for spending.[5]

---

[2] Omnibus Budget Reconciliation Act of 1990 § 10101(a), 35 U.S.C. 41 (2006).

[3] U.S. PATENT & TRADEMARK OFFICE, PERFORMANCE AND ACCOUNTABILITY REPORT FISCAL YEAR 2003, at 54, *available at* http://www.uspto.gov/about/stratplan/ar/2003/2003annualreport.pdf; U.S. PATENT & TRADEMARK OFFICE, PERFORMANCE AND ACCOUNTABILITY REPORT FISCAL YEAR 2002, at 54, *available* at http://www.uspto.gov/web/offices/com/annual/2002/1-58.pdf; U.S. PATENT & TRADEMARK OFFICE, PERFORMANCE AND ACCOUNTABILITY REPORT FISCAL YEAR 2001, at 58, *available at* http://www.uspto.gov/web/offices/com/annual/2001/01performreport.pdf; U.S. PATENT & TRADEMARK OFFICE, ANNUAL REPORT, FISCAL YEAR 2000, at 49, *available at* http://www.uspto.gov/about/stratplan/ar/2000/00findisc.pdf; U.S. PATENT & TRADEMARK OFFICE, ANNUAL REPORT, FISCAL YEAR 1999, at 28, *available at* http://www.uspto.gov/about/stratplan/ar/1999/99mssgchief.pdf.

[4] OFFICE OF MGMT. AND BUDGET, BUDGET OF THE U.S. GOVERNMENT FISCAL YEAR 2002, APPENDIX, DEPARTMENT OF COMMERCE 227 (2002), *available at* http://www.gpo.gov/fdsys/pkg/BUDGET-2002-APP/pdf/BUDGET-2002-APP-1-6.pdf.

[5] U.S. PATENT & TRADEMARK OFFICE, PERFORMANCE AND ACCOUNTABILITY REPORT FISCAL YEAR 2002, at 54, *available at* http://www.uspto.gov/web/offices/com/annual/2002/1-58.pdf. While in theory these diverted fees could become available to the agency in future years—the PTO still has these fees on their books as "temporarily unavailable," s*ee, e.g.*, 2011 PERFORMANCE AND ACCOUNTABILITY REPORT, *supra* note 33, at 67, it is widely believed that the PTO is unlikely to ever receive the authority to spend any significant portion of these fees, INTELLECTUAL PROP. OWNERS ASS'N, UNDERSTANDING PATENT FEE DIVERSION AND HOW IT IS AFFECTED BY CURRENT SENATE AND HOUSE PATENT REFORM BILLS (2011), *available at* http://www.ipo.org/AM/Template.cfm?Section=Home&ContentID=30761&Template=/CM/ContentDisplay.cfm ("No one expects that these fees will ever be made available to the USPTO.").

Construction of Examination Cost Measure

For each of the 37 NBER technology sub-categories, we calculate an estimate of the examination cost of an average patent application within the relevant technology. To begin, we first determine the average cost to the PTO of having an examiner within each pay grade (e.g., GS level 5) examine a patent within the relevant technology, based on: (1) the number of examination hours slotted to an examiner of that pay level within that technology and (2) the cost per hour of time for that examiner (based on the salaries assigned to that pay level). To form the average examination cost over the entire technology, it is then necessary to form some weighted average of the examination costs for that technology across each of the GS pay levels within that technology. The weights used for this average will differ across technologies to the extent that technologies face different distributions of examination efforts across different pay scales (e.g., a technology that is more heavily populated by senior examiners will obviously place more weight on the average examination costs attributable to the higher GS payscales within that technology). To calculate the necessary weights, we determine the percentage of examination time within each technology that is allocated to each of the various GS pay levels, using information on (1) the distribution of examiners across the different pay levels within each technology and (2) the percentage of time, on average, that examiners of each pay level within the relevant technology spend examining patents (noting that more senior examiners will spend a greater percentage of their time on non-patent-examination activities).

A critical component of the above calculation is the hours allocated to an examiner for the purposes of processing an application, an allocation that is a function of both the technology and of the pay level of the examiner. Perhaps the key source of variation across technologies in the above calculation stems from the fact that different technologies, as a baseline, are allocated different lengths of examination time. These baseline allocation amounts are generally a function of the complexity of the associated examinations, with the more complex areas receiving more time for patent examination. Necessarily (considering the cost per hour of time to staff an examiner to undertake that examination), this variation in baseline hour allocations across technologies will contribute to overall average examination cost differentials across technologies. To put things in perspective, the baseline hour allocations range from 13.8 hours (Heating) to 27.4 (Electronic Business Methods and Software). The full distribution of these hour allocations across the 37 NBER technology sub-categories can be found in Table A1 below. This schedule remained fixed over the duration of our sample period. According to a fixed set of scaling factors (see below), examiners in different GS paygrades receive an adjustment to this baseline hours allotment, with lower pay-grade examiners receiving a greater amount of time to examine an application.

We note at the outset that the results of our empirical analysis are entirely robust to a cost estimation that focuses only on variation in these baseline hour allotments across technologies, as opposed to the primary approach summarized above.

To provide more specifics regarding our primary approach, we calculated average technology-specific examination costs as follows:

$$Cost_c = \sum_{g=5}^{15}(Hours_{gc} * \left(\frac{Cost}{Hour}\right)_g * Percent\_Examination\_Time_{gc}) \, ,$$

where

c represents a given technology category and g represents a pay-grade level;

$$Hours_{gc} = Hours(baseline)_c / Scaling\_Factor_g;$$

and

$(\frac{Cost}{Hour})_g = \frac{Annual\_Salary_g}{1980};$ \qquad (where 1980 is an estimate of the number of work hours in a given year)

$$Percent\_Examination\_Time_{gc}$$
$$= (Number\_Examiners_{gc} * Percent\_Year\_Examining_{gc}) /$$
$$\sum_{w=5}^{15}((Number\_Examiners_{wc} * Percent\_Year\_Examining_{wc}) .$$

Data on the number of examiners within each technology-payscale cell (Number_Examiners$_{gc}$) was obtained pursuant to a Freedom of Information Act (FOIA) request filed with the PTO. , as was the percentage of time that examiners within that cell generally spend examining patents (Percent_Year_Examining$_{gc}$). These data allow us to determine the percentage of examination efforts within each technology that are attributable to the various pay levels within that technology (Percent_Examination_Time$_{gc}$). The baseline hours allocations across the various technologies are determined according to a schedule that we likewise received from the PTO according to a targeted FOIA request. Scaling factors are as follows: 0.55 (GS level 5), 0.7 (GS level 7), 0.8 (GS level 9), 0.9 (GS level 11), 1 (GS level 12), 1.15 (GS level 13) and 1.35 (GS level 14 and 15). Salary amounts are determined according to the standard payscale for federal government employees.[6]

Timing. We calculate the above examination cost estimate across technologies in a time-invariant manner, consistent with the relevant empirical precedent—i.e., consistent with those studies that have similarly explored how single national reforms or other shocks lead to

---

[6]   See http://www.fedjobs.com/pay/pay.html.

differential responses along some theorized dimension. The baseline hours allocations and scaling factors are already time-invariant in nature. While annual salary levels change over time, for simplicity purposes, we use levels as of 2011. The weights captured by the Percent_Examination_Time variable technically vary over time in so far as examiner payscale distributions within technologies change over time; however, certain data limitations forced us to calculate expected examination costs using the payscale distributions as they existed in 2011.[7]

The payscale distribiution data that we received was technically disaggregated by Group Art Unit (GAU). Using the current year patent-class—GAU mapping provided by the PTO at http://www.uspto.gov/_patents/resources/classification/art/index.jsp and using the patent-class—technology-sub-category mapping provided by the NBER, we aggregated this data to the NBER sub-category level. The baseline hours allocation that we received from the PTO pursuant to our directed FOIA request was provided to us at the PTO class level. Taking a weighted average of these hour allocations across the different classes allocated to each NBER sub-category, we derived a baseline examination hour allotment at the NBER sub-category technology level (weighting by the number of application disposals within the respective classes).

Distribution of costs across technologies. In Table A1 below, we present the distribution of the estimated cost estimates for each of the 37 NBER sub-categories. Since our concern in this paper is largely in demonstrating variation in costs across technologies as opposed to precisely estimating and reporting the absolute values of those technology-specific cost estimates, we present these cost figures after standardizing them such that they have a mean of 1 across all technologies. Nonetheless, for the sake of completeness, we also present the cost estimates in dollars as well. We stress, however, that these costs only reflect the variable salary component to average examination costs across technologies and do not reflect any overhead allocations or related cost allocations.

Additional Notes. It is possible that examiners are actually spending less time on their applications than the allocations specified in this workload schedule. Nonetheless, we have no information to suggest that any such deviations will alter the ordinal rankings of technology-specific examination costs to the Agency (in other words, this analysis is not substantially altered if examiners across the board have trended below these workload schedules in their examination times). Moreover, it is important to bear in mind that the focus here is on the Agency and not on the examiners. Examiners are effectively compensated and assigned work based on these workload schedules. If examiners process applications more quickly than these schedules provide for, it may be the case that they are simply able to end their work sooner than envisioned. In this event, considering that they are implicitly being compensated based on the assumption that it would have taken the full time, this schedule nonetheless depicts the examination cost to the Agency. Of course, there is another possibility. If examiners spend less than their allotted time, they may be able to churn through a greater number of applications. If

---

[7] More specifically, as explained below, we received these data at the Group Art Unit level. The information that we have to facilitate a mapping to the NBER sub-category level is only available as of the present.

examiners were simply compensated on a fixed salary, this would effectively entail that technology-specific examination costs may indeed be lower than that dictated by this analysis. However, in an event of this nature, examiners will effectively accrue a greater number of "counts" (generally they receive 1 count per examination processed). Depending on the policies in place at that precise time, examiners may generally receive overtime pay to the extent that they accrue counts beyond that implied by their workload schedule. With overtime adjustments of this nature, the baseline workload schedule that we are using to generate technology-specific examination cost estimates to the Agency remains a valid way to differentiate across technologies. After all, the Agency, as distinct from the examiner, may not necessarily save funds to the extent that the examiner spends less than the allotted time on the application.

This entire discussion may not be relevant however to the extent that examiners spend at least the full allotted time on their applications. According to a 2007 survey of patent examiners by the Government Accountability Office,[8] 70 percent of patent examiners over the prior year worked unpaid overtime to meet their production goals, some more than 30 extra hours in a 2-week period. Accordingly, in consideration of the hourly costs of an examiner's time implicit within their compensation, the number of hours that the PTO does allocate to a given examination may indeed be a fair reflection of the expected costs associated with that application (acknowledging that the examiner may spend more than that number of hours on the application, but may only be compensated effectively for the scheduled amount of time, as suggested by the referenced survey).

---

[8] See http://www.gao.gov/new.items/d071102.pdf.

Table A1: Examination Costs by Technology

Technology Category	(1) Examination Hours: Baseline	(2) Standardized Examination Cost Estimate	(3) Examination Cost Estimate ($)
Agriculture, Food, Textiles	19.2	1.01	786.78
Coating	20.9	1.08	841.68
Gas	21.7	1.13	878.28
Organic Compounds	18.8	1.01	782.28
Resins	19.2	1.03	796.51
Miscellaneous Chemical	18.4	0.97	756.50
Communications	19.0	1.00	780.16
Computer Hardware & Software	24.0	1.26	975.63
Computer Peripherals	21.9	1.17	905.86
Information Storage	16.2	0.87	671.70
Electronic Business Methods and Software	27.4	1.45	1122.75
Drugs	17.6	0.95	739.31
Surgical and Medical Instruments	14.4	0.73	566.49
Genetics	24.8	1.38	1067.92
Miscellaneous Drugs and Medical	18.8	0.99	768.08
Electrical Devices	17.9	0.94	728.83
Electrical Lighting	18.8	1.00	775.86
Measuring & Testing	18.0	0.95	733.61
Nuclear & X-rays	19.8	1.04	803.65
Power Systems	18.7	0.98	761.69
Semiconductor Devices	21.0	1.11	861.77
Miscellaneous Electrical	18.0	0.92	716.89
Mat. Proc & Handling	19.1	1.03	795.91
Metal Working	19.2	1.02	790.35
Motors & Engines & Parts	18.4	0.94	731.51
Optics	17.5	0.94	727.70
Transportation	16.9	0.90	695.66
Miscellaneous Mechanical	17.9	0.94	732.22
Agriculture, Husbandry and Food	18.6	0.97	752.83
Amusement Devices	15.9	0.83	644.73
Apparel & Textile	17.4	0.92	711.61
Earth Working & Wells	17.6	0.92	713.41
Furniture, House Fixtures	16.6	0.87	673.65
Heating	13.7	0.67	520.74

Pipes & Joints	16.6	0.85	658.16
Receptacles	15.3	0.77	604.66
Miscellaneous Other	18.4	0.95	741.03

Sustainability Score Calculation

In Section I, we predicted that the PTO would be more likely to trigger its sustainability / budget constraint as the ratio between its incoming post-allowance fee collections to outgoing examination expenditures fell. We predicted that this would be more likely to occur upon the following developments: an increase in the PTO's aggregate backlog (relative to the stock of patents from which the Agency derives maintenance fees), a decrease in its annual maintenance-fee collections, an increase in its average examination complexity (i.e., the average number of hours allocated to each examination disposed of in a given year), and a decrease in the percentage of patentees that are large entities. While we consider regressions that interact each of these factors separately with the technology-specific cost estimate, we also estimate regressions that aggregate each of these factors into one sustainability measure, a measure that is meant to be reflective of whether, on net, the PTO faces sustainability concerns. For instance, if over time, total backlog grows considerably while annual maintenance rates actually increase somewhat (which characterizes much of the sample period), how can we determine whether the PTO is in fact experiencing changes in its financial outlook? For these purposes, we construct the sustainability measure set forth in greater detail in the Online Appendix of Frakes and Wasserman (2013), which captures the impact of each measure on the PTO's financial balance in a manner that facilitates across-factor comparisons in such impacts.

Broadly, the sustainability score specified in Frakes and Wasserman (2013) for a given year equals the amount of incoming post-allowance fees for that year divided by the net examination costs associated with all of the patent applications awaiting examination at that time. The score is not meant to reflect the actual profits accruing to the PTO in a given year. Rather, it is meant to simulate how variations in the above-mentioned factors (keeping all other factors fixed) affect its general profitability. That is, it provides a meaningful and empirically relevant way of assessing the relative contributions to the PTO's financial position of each of these factors. Also, while an actual annual profitability measure may consider the costs associated with those applications disposed of during a given year, this measure considers the costs associated with all of those applications awaiting examination at that time—i.e., the backlog. The costs associated with examining the backlog represent a better sense of the external pressures being placed upon the PTO, as compared with the costs associated with those patents the PTO elected to dispose of during the year, which would be, in part, a reflection of the PTO's own response to its financial pressures. Our goal is then to evaluate how these external pressures to the agency's financial position induce it to take certain actions. As explained further below, we also modify this score in several key ways to further emphasize its external nature.

The sustainability score keeps fixed over time the fee amounts themselves (based on the 2011 amounts), again focusing only on variations in the above-mentioned factors. However, our preliminary extensions of this score based on our current understanding of fee amendments suggest that the regression results persist under this extension. Likewise, in calculating the average cost per examination, the only factor changing over time is the average complexity of the patents disposed of during that year. To calculate net costs, we multiply the average examination cost in 2010 by the ratio of the average examination complexity for the given year (based on the distribution of patents disposed of during the year) to the average examination complexity of 2010, the reference year.

We also consider several alternative constructions of this score. First, we estimate a variant of this score in which we replace the total backlog of pending applications as the relevant estimate of current application demand with a slightly more specific total backlog figure: the total inventory of patent applications awaiting a first office action. Second, for the reasons discussed in Section IV, we construct another variant in which we replace the total backlog of pending applications with the total backlog of pending applications among those technologies with below-average examination costs. We use this specific backlog count to nonetheless construct a sustainability score that is applied to all observations (high- and low-cost technologies alike). Finally, as again further discussed in Section IV, we construct yet another sustainability measure, one which is now unique to each technology-year cell as opposed to each year in the sample. With this measure, for reasons indicated in the text, we calculate a score for each technology-year in a manner that removes the components of the sustainability score that arise from that technology and year—e.g., the score applied to the Genetics category in 1995 is calculated so as to not include the Genetics-specific backlog amount.

To facilitate comparison between the results from these alternative specifications (and in general to facilitate the interpretation of the results in a meaningful fashion), we normalize all separate sustainability measures by dividing them by twice the standard deviation of such scores across all sample years. In this manner, the underlying specifications will explore the consequences of a 2-standard-deviation improvement (decline) in the financial health of the Agency, allowing us to effectively compare the behavior of the Agency (differentially across technologies of varying costs) during times in which the PTO is in "good" financial health with the Agency's behavior during times of "poor" financial health (in other words, a 2 standard-deviation shock in the sustainability score will facilitate a meaningful interpretation of the impact of the Agency's financial position on its differential examination priority practices).

In this Appendix, we demonstrate and discuss the robustness of the findings presented in Tables 2 and 3 and in Figures 1–3 to a range of specification checks and other robustness exercises. Generally, the results of these exercises demonstrate the flexibility of the findings to a number of alternative approaches and demonstrate the robustness of the conclusions to the consideration of various potentially confounding stories. For these purposes, we set forth the following tables:

- Table B1, summarizing a number of the robustness exercises that we undertake in connection with Table 2 of the Article;
- Table B2, summarizing a number of the robustness exercises that we undertake in connection with Table 3 of the article;
- Table B3, in which we replicate Table 2 of the Article except using an alternative examination delay measure as the dependent variable: the average length of time (in years) between the filing of an application and the issuance of the resulting patent, out of the full universe of patent issuances over 1991-2003 (note that we were not able to collect this data between 2007 and 2010, precluding our ability to use this measure in replicating Table 3 of the Article);
- Table B4, in which we consider a less parametric treatment of examination costs across technologies—that is, instead of exploring PTO responses to linear increases in costs across technologies, we explore how the PTO responds to more categorical increases in costs; and
- Table B5, in which we break out each of the individual components of the sustainability score and separately interact each component with the estimated examination cost variable (in order to test how fluctuations in each of these separate markers of an underfunded PTO correlate with fluctuations in the differential examination pendency trends between high- and low-cost technologies).

We then offer some textual discussions of some of these exercises, in addition to certain additional specification checks. Finally, we present a set of figures that demonstrate the robustness of Figures 1-3 of the Article to various alternative specifications.

		Difference-in-Difference Coefficient Estimate Under the Alterations Indicated (frame of reference = Column 1 of Table 2)
1)	Dropping the Genetics, Drugs and Organic Compounds Technology Categories—the high restriction-practices categories	-3.51*** (1.24)
2)	Including control variables for the technology-year percentage of total application filings that constitute requests for continued examinations and their predecessors (i.e., continuing prosecution applications, or CPA's), in addition to the square of this percentage	-3.41*** (1.22)
3)	Including technology-specific linear time trends	-2.77** (1.11)
4)	Using the natural log of the examination pendency as the operable dependent variable	-0.13* (0.07)
5)	Specifying the estimated technology cost only according to the baseline hours allocation across technologies (likewise standardized to evaluate a 2-standard deviation shift in such hours)	-3.62*** (1.24)
6)	Including control variables for the technology-year rate of written restrictions (normalized by initial application counts)	-3.5*** (1.1)
7)	Including control variables for the technology-year continuation rate (equal to the number of continuation and requests for continued examination filings divided by the total number of filings)	-2.4** (1.1)
8)	Full Filings Control Specification: Including separate control variables for the technology-year continuation rate (exclusive of requests for continued examinations), the request for continued examination filing rate, the rate of written restriction requirements and the rate of initial application filings (equal to the share of the aggregate initial application filings for that year attributable to the technology group of interest).	-1.95** (0.90)

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within patent categories over time. All regressions include patent-technology-category fixed effects and year fixed effects. Regressions are weighted by the number of first office actions undertaken in the respect technology-year cell.

		Main Difference-in-Difference Coefficient Estimate (SUSTAINABILITY*COST) Under the Alterations Indicated (frame of reference = Column 1, Row 1 of Table 3)	Coefficient Estimate of Interaction Variable (SUSTAINABILITY*COST* FEE_DIVERSION) Under the Alterations Indicated (frame of reference = Column 1, Row 2 of Table 3)
1)	Dropping the Genetics, Drugs and Organic Compounds Technology Categories—the high restriction-practices categories	-3.9*** (1.2)	11.3*** (4.1)
2)	Including control variables for the technology-year percentage of total application filings that constitute requests for continued examinations and their predecessors (i.e., continuing prosecution applications, or CPA's), in addition to the square of this percentage	-3.9*** (1.2)	12.2*** (4.5)
3)	Including technology-specific linear time trends	-3.0*** (1.1)	12.5*** (4.1)
4)	Using the natural log of the examination pendency as the operable dependent variable	-0.13** (0.06)	0.59** (0.21)
5)	Specifying the estimated technology cost only according to the baseline hours allocation across technologies (standardizing to evaluate a 2-standard-deviation shift in such hours)	-4.0*** (1.1)	10.8** (4.1)
6)	Including control variables for the technology-year rate of written restrictions (normalized by initial application counts)	-3.9*** (1.2)	11.2*** (3.9)

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within patent categories over time. All regressions include patent-technology-category fixed effects and year fixed effects. Regressions are weighted by the number of first office actions undertaken in the respect technology-year cell.

TABLE 2. THE IMPACT OF PTO FINANCIAL HEALTH ON DIFFERENTIAL EXAM DURATIONS (MONTHS, TIME TO ISSUANCE) BETWEEN HIGH- AND LOW-COST TECHNOLOGIES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SUSTAINABILITY * COST	-2.5**	-1.8*	-2.5**	-2.3**	-2.2**	-2.2***	-2.0***
	(1.0)	(0.9)	(1.1)	(1.0)	(0.9)	(0.6)	(0.6)
N	480	480	480	480	480	480	480
Sustainability Score Construction	Version 1	Version 2	Version 3	Version 4	Version 5	Version 1	Version 1
Technology-Year Filing Rates?	NO	NO	NO	NO	NO	YES	YES
Other Technology-Year Covariates?	NO	NO	NO	NO	NO	NO	YES

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within technologies over time. All regressions include technology fixed effects and year fixed effects. The sustainability score and technology-specific cost measures are each normalized such that an increase in "1" of each normalized variable represents an increase of 2 standard deviations of the respective measure. Reported coefficients are for the difference-in-difference variable of interest and reflect the degree to which a 2-standard-deviation improvement in the sustainability score is associated with a change in the differential exam pendency between a technology with examination cost X and a technology with an examination cost equal to 2 standard deviations below X. Regressions are weighted by the number of first office actions associated with each technology-year cell. Version 1 of the sustainability score bases the examination demand facing the Agency on the total backlog of applications awaiting review. Version 2 calculates the score applying to a given technology-year cell by excluding contributions to the score from that technology-year cell. Version 3 bases the examination demand facing the Agency on the total backlog of applications awaiting review among the low-cost technologies only. Version 4 of the sustainability score excludes the average examination cost from the underlying calculation. Finally, Version 5 bases the examination demand facing the Agency on the total backlog of applications awaiting a first office action. Patent processing data is from the PTO.

Note that the above results are likewise robust to the use of the natural log of the examination duration (time to issuance) as the dependent variable. For instance, the log specification suggests that a 2 standard deviation drop in the sustainability score is associated with a 6 percent longer examination duration for high-cost relative to low-cost technologies, as captured by a 2 standard deviation shift in examination costs across technologies (significant at the 5% level).

TABLE B4. THE IMPACT OF PTO FINANCIAL HEALTH ON DIFFERENTIAL EXAM DURATIONS BETWEEN HIGH- AND LOW-COST TECHNOLOGIES: NONPARAMETRIC TREATMENT OF ESTIMATED COST LEVELS (1991-2003)

	(1)
Dependent Variable: Time to Issuance (Months). Unit of Observation = Technology / Year	
(Reference group: 0-33rd Percentile)	
SUSTAINABILITY * (33rd – 66th Percentile)	-0.03 (0.94)
SUSTAINABILITY * (66th-100th Percentile)	-3.49** (1.68)

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within patent technology categories over time. All regressions include patent-category fixed effects and year fixed effects. Regressions are weighted by the number of first office actions completed in each technology-year cell. The sustainability score measure is normalized such than an increase in "1" represents an increase of 2 standard deviations of the actual sustainability score measure.

TABLE B5. SEPARATE EFFECTS OF TWO-STANDARD-DEVIATION SHIFTS IN EACH OF THE KEY COMPONENTS OF THE SUSTAINABILITY SCORE

	(1)	(2)
BACKLOG / PATENT STOCK RATIO * COST	6.0** (2.3)	2.11 (1.89)
AVERAGE RENEWAL RATE * COST	1.3** (0.6)	-1.28** (0.58)
AVERAGE EXAM COMPLEXITY * COST	4.2*** (1.5)	4.29*** (1.51)
LARGE ENTITY PERCENTAGE * COST	1.2** (0.5)	0.62 (0.76)

* significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are reported in parentheses and are clustered to correct for autocorrelation within technology categories over time. All regressions include technology-category fixed effects and year fixed effects. Regressions are weighted by the number of first office actions completed in each technology year cell. Column 1 actually represents results from 4 separate regressions, where each factor is considered individually in a given regression. Column 2 represents results from a single specification, where each factor is considered together.

Restriction Practices

There is a potential concern that the PTO's restriction practices may be responsible for the observed differential trends in the Agency's examination priorities. The PTO has the authority to issue a restriction requirement when two or more independent or distinct inventions are claimed in a patent application.[9] Thus, the Agency's restriction authority enables it to divide a single patent application into two or more patent applications. Restriction practices, however, are not uniform across all technologies areas but instead are more heavily relied upon in the

---

[9]   U.S. Patent & Trademark Office, Manual of Patent Examination Procedure § 802.02 (8th ed., rev. 9, 2012).

biotechnology and organic chemistry fields.[10] Moreover, the PTO appears to have dramatically increased their issuance of restriction requirements in the 2000s.[11] Because there is an average three month delay between when a restriction requirement is issued and when substantive examination of a restricted application begins,[12] there may be concern that certain patterns of restriction practices could be driving our results. To date, empirical studies suggest that restriction requirements are not solely concentrated in high cost technologies.[13] Moreover, even if the PTO's restriction requests are disproportionately concentrated in complex technologies, the frequency of utilization of such practices must be correlated with both changes in the Agency's financial health as well as the 2004 reform in order to explain the above results.[14] To further alleviate any residual concerns, we estimate the pattern of results presented in Tables 2 and 3 after dropping drugs, genetics, and organic compounds technologies together—the fields wherein the PTO relies upon restriction practice the most.[15] Our results remain virtually unchanged (as demonstrated by Tables B1 and B2), providing strong support that the PTO's restriction practice does not explain our findings. Out of an abundance of caution, we also filed a Freedom of Information Act Request to the PTO to obtain data on the Agency's restriction practices on a yearly, patent-class level. In the last row of Tables B1 and B2 we include controls for the technology-year utilization rate of restriction practices (normalized by initial filing counts). The results remain virtually identical to those estimated in Tables 2 and 3 of the text.

Technology-Specific Linear Time Trends

Since the unit of observation in our empirical analysis is a given technology-year combination, we obviously cannot account for all unobservable characteristics specific to given technology-year cells. However, we can at least implicitly account for that portion of those unobservable characteristics that moves slowly and linearly over time through the imposition of these additional controls. The validity of this approach of course rests on the perhaps unknowable question of the existence of any such slowly-moving linear factors. In any event, the fact that our results are unchanged with the inclusion of such controls (as demonstrated by Tables B1 and

---

[10]   Dennis   Crouch,   *Likelihood   of   Office   Action   Rejections*,   PATENTLYO,   (June   15,   2010), http://www.patentlyo.com/patent/2010/06/likelihood-of-office-action-rejections.html (noting that based on 20,000 file histories of published patent applications the restriction requirement rates vary across technology centers where Biotechnology and Organic Chemistry has the highest rate at 54%, followed by Chemical and Material Engineering at 21%, followed by Transportation, Construction, Electronic Commerce, and Agriculture at 19%, followed by Semiconductor Electrical and Optical Systems and Components and Mechanical Engineering, Manufacturing Products both at 18%, followed by Communications at 7% and finally Computer Networks, Multiplex Video Distribution, and Security and Computer Architecture, Software and Information Security both at 6%).

[11]   Shine S. Tu, et al., *Squeezing More Patent Protection from a Smaller Budget Without Compromising Quality*, 2 LANDSLIDE 35, 37 (2009) (finding the PTO's use of restriction practices in the biotechnology practice group has increased from approximately 3% of patent applications receiving restriction requirements in 1993 to over 50% of patent applications receiving restriction requirements by 2008; while the authors note in 1993 that approximately 1.5% of the patent applications in the biotechnology practice group were issued a restriction requirement, this number appears to be inconsistent with their reported findings that 1,000 restriction requirements were filed out of approximately 32,000 first office actions—1,000/32,000 ≈ 3%).

[12]   Julie Burke, U.S. PTO Quality Assurance Specialist Technology Center 1600, FY08 Restriction Petition Update and Burden, *available at* www.cabic.com/bcp/120308/JBurke_RPD.ppt (noting that average turnaround time for restriction petitions was 91 days in the fiscal year of 2004 through 2007 and 103 days in the fiscal year of 2008).

[13]   *See* Crouch, *supra* note 10 (noting that restriction rates and the complexity of the art are not positively correlated; i.e., restriction rates in complex areas such as computer networks and software are low (6%) whereas the restriction rates of less complex areas such as mechanical arts and transportation are higher (18-19%)). Thus, empirical evidence suggests that there is no positive correlation between restriction rates and the examination costs of patent applications.

[14]   Even if such spurious correlation does occur, it could not possibly explain the increase in the PTO's grant rate towards patents the Agency stands to profit the most from issuing in 2004.

[15]   *See* Crouch, *supra* note 10.

B2) lends additional support to the proposition that some unobservable factor other than the Agency's financial health and funding structure is responsible for the observed patterns.

Randomization Inference

Standard errors may be inaccurately estimated in difference-in-difference specifications when there are a limited number of overall analytical or treatment groups.[16] This concern may not be terribly pronounced in the present context considering that (with thirty-seven technology categories each with different cost levels) there are effectively a larger number of treatment groups. In any event, we also perform hypothesis tests on the estimated coefficient of the *SUSTAINABILITY*COST* variable in the primary empirical specification (Column 1 of Table 2) using a randomization inference approach,[17] which allows for an estimation of the distribution of the treatment effect that is valid under any number of groups. For these purposes, we run five thousand simulations, where, with each simulation, we randomly assign each technology category a different estimated cost level, based on the distribution of cost levels actually observed.[18] We find that the estimated difference-in-difference coefficient reported in Column 1 Table 2 easily falls within the first percentile of the empirical distribution of the five thousand estimated coefficients from the simulations, consistent with a p-value of less than 0.01—i.e., statistical significance at the 1% level.

Nonparametric treatment of estimated examination costs.

The primary results explore the interaction between the sustainability score and examination costs using a linear treatment of technology-specific costs. In Table B5, we allocate technology categories into one of three groups, based on their examination cost percentile: (1) bottom 33rd percent, (2) 33rd-66th percent, and (3) top 33rd percent. We assign each technology category three dummy variables indicating whether or not the respective category falls into the relevant tercile. We then interact each such dummy variable with the sustainability score. We include each interaction in a single regression, leaving out the dummy representing the bottom tercile, which will serve as the reference group. The results suggest that the distortionary examination-timing practices of the Agency are concentrated in the highest-cost technologies. In response to a deterioration in the Agency's financial state, it does not appear that the Agency differentially treats its examination priorities between the bottom and middle terciles of examination costs.

Components of Sustainability Score.

Rather than simply exploring how the differential examination pendencies across high- and low-cost technologies changes in connection with fluctuations in the composite sustainability score, we also estimate regressions in Table B5 that break out the key components to that score and

---

16.   *See, e.g.*, Timothy G. Conley & Christopher R. Taber, *Inference with "Difference-in-Differences" with a Small Number of Policy Changes*, 93 REV. ECON. & STAT. 113 (2011).

17.   *See, e.g.*, Esther Duflo, Rachel Glennerster & Michael Kremer, *Using Randomization in Development Economics: A Toolkit, in* 4 HANDBOOK OF DEVELOPMENT ECONOMICS 3895, 3895–3962 (T. Paul Schultz & John A. Strauss eds., 2007).

18.   *See* Jonathan Gruber & Daniel M. Hungerman, *Church Versus the Mall: What Happens When Religion Faces Increased Secular Competition*, 123 Q. J. ECON. 831 (2008).

estimate how the differential examination delays change in connection with fluctuations within each of these components, independently. Column 1 actually represents results from 4 separate regressions, where each factor is considered individually in a given regression. Column 2 represents results from a single specification, where each factor is considered together, allowing us to partial out any correlations. To facilitate the interpretation of these results, we interpret the regressions with respect to a 2-standard-deviation increase in the magnitude of each of the respective factors. For the purposes of this illustration, we break out the following factors:

(1) The PTO's aggregate backlog of pending examinations for the given year normalized by the stock of patents available to generate post-allowance fees that year—i.e., the sum of the patents issued that year, four years previously, eight years previously, and twelve years previously. An increase in this backlog ratio would suggest a weakening in the PTO's financial balance, in connection with which one would presume to observe (in the 1991-2003 period) an increase in differential examination delays between high- and low-cost technologies—i.e., a positive coefficient.

(2) the average maintenance rate for patents eligible for renewal that year, averaging the four-year, eight-year, and twelve-year rates for ease of presentation. A decrease in this rate would suggest a weakening in the PTO's financial balance, in connection with which one would presume to observe an increase in the differential examination delays between high- and low-cost technologies—i.e., a negative coefficient.

(3) the average examination complexity (i.e., average examination hours) of the patents disposed of that year. An increase in this average complexity (e.g., due to a shift towards applications in more complex areas) would suggest a weakening in the PTO's financial balance, in connection with which one would presume to observe an increase in the differential examination delay between high- and low-cost technologies—i.e., a positive coefficient.

(4) the percentage of the patent stock available to generate post-allowance fees that year (e.g., the patents issued that year and each of four-, eight,- and twelve-years previously) that are large entities. A decrease in this rate would suggest a weakening in the PTO's financial balance, in connection with which one would presume to observe an increase in the differential examination delays between high- and low-cost technologies—i.e., a negative coefficient.

Again, fluctuations in these aggregate factors over time can disrupt the financial balance reached between incoming fee collections and the costs associated with the examinations expected of the Agency, leaving the Agency potentially in need of funds. Instead of aggregating all of these factors into one sustainability score, this specification breaks that score out into these separate parts. In the collective specification from Column 2, the sign of the estimated coefficients for the aggregate backlog, maintenance-rate, and average examination complexity interaction terms are consistent with the predictions set forth in the above list.

Supplemental Figures

To demonstrate the robustness of the graphical analyses presented in Figures 1-3 of the Article, we provide the following supplemental figures:

- Figure B1: replication of Figure 1 with the inclusion of controls for initial application filings (and their squares);
- Figure B2: replication of Figure 1 with the inclusion of controls for initial application filings (and their squares) and for various additional technology-year covariates (e.g., average number of claims among issued patents);

- Figure B3: replication of Figure 2 with the inclusion of controls for initial application filings (and their squares) and for various additional technology-year covariates (e.g., average number of claims among issued patents);
- Figure B4: replication of Figure 1, but using as the relevant dependent variable the average length of time between the filing of an application and the ultimate issuance of a patent (among the full universe of patent issuances over the sample period);
- Figure B5: replication of Figure 1, but using as the relevant dependent variable the average length of time between the filing of an application and the ultimate issuance of a patent (among the full universe of patent issuances over the sample period), including controls for initial application filings (and their squares) and for various additional technology-year covariates (e.g., average number of claims among issued patents);
- Figure B6: replication of Figure 2, but using as the relevant dependent variable the average length of time between the filing of an application and the ultimate issuance of a patent (among the full universe of patent issuances over the sample period), including controls for initial application filings (and their squares) and for various additional technology-year covariates (e.g., average number of claims among issued patents);
- Figure B7: replication of Figure 1, but over the 1986-2010 period, as opposed to the 1991-2010 period—that is, including the 5 year period prior to the adoption of the PTO's present fee structure (and near fully user-fee funded model);
- Figure B8: replication of Figure 1, but using the growth rate (relative to the base year of 1990) in the examination pendency as the dependent variable in the dynamic difference-in-difference regression, as opposed to the level of the examination pendency; and
- Figure B9: replication of Figure 1, but using the natural log of the examination pendency as the dependent variable.

1.

Figure B1: Dynamic Difference-in-Difference Results
Overlaid with Sustainability Score
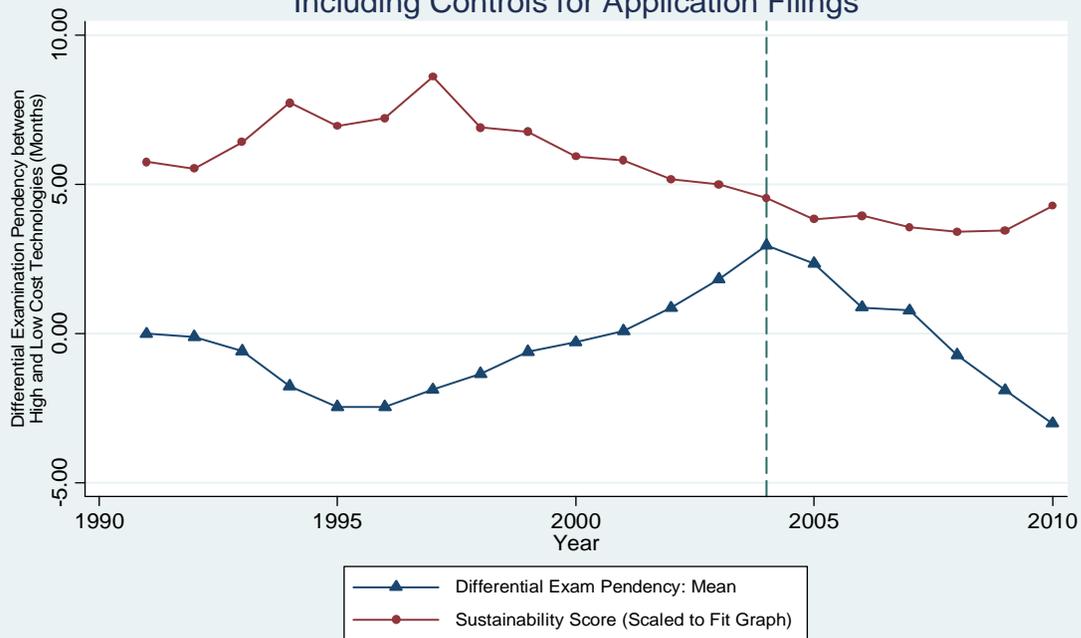Including Controls for Application Filings

Figure B2: Dynamic Difference-in-Difference Results
Overlaid with Sustainability Score
Including Controls for Application Filings and
Technology-Year Covariates
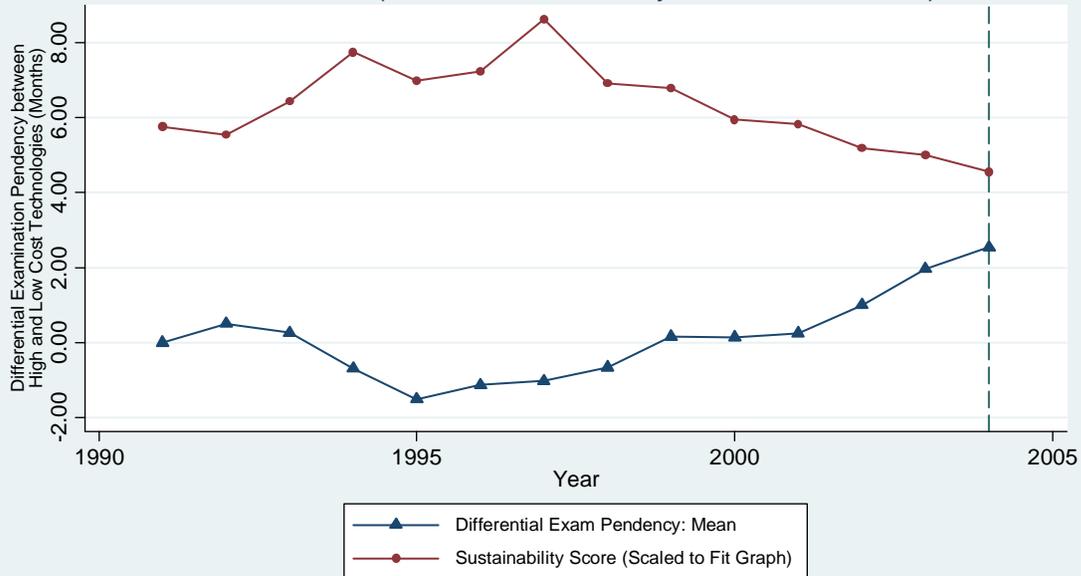1991-2004 (Period of Availability of Covariate Data)

Figure B3: Dynamic Difference-in-Difference Results Overlaid with Alternative Sustainability Score Including Controls for Application Filings and Technology-Year Covariates (1991-2004)

Differential Exam Pendency: Mean
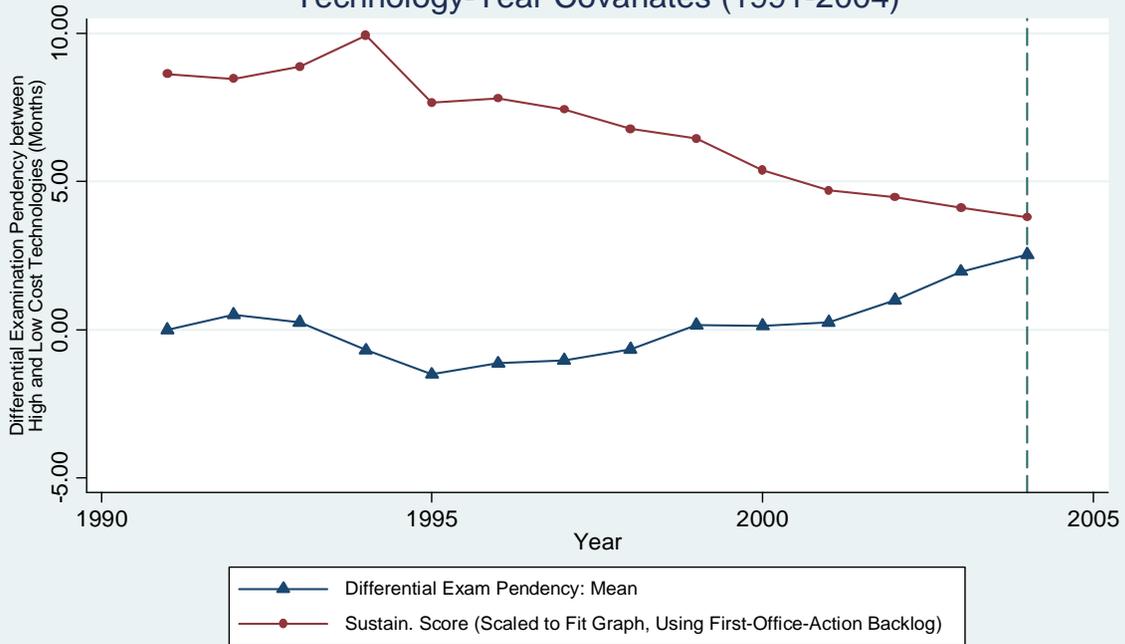Sustain. Score (Scaled to Fit Graph, Using First-Office-Action Backlog)

Figure B4: Dynamic Difference-in-Difference Results Overlaid with Sustainability Score Dependent Variable = Total Examination Duration 1991-2004 (Period of Availability of Total Duration Data)
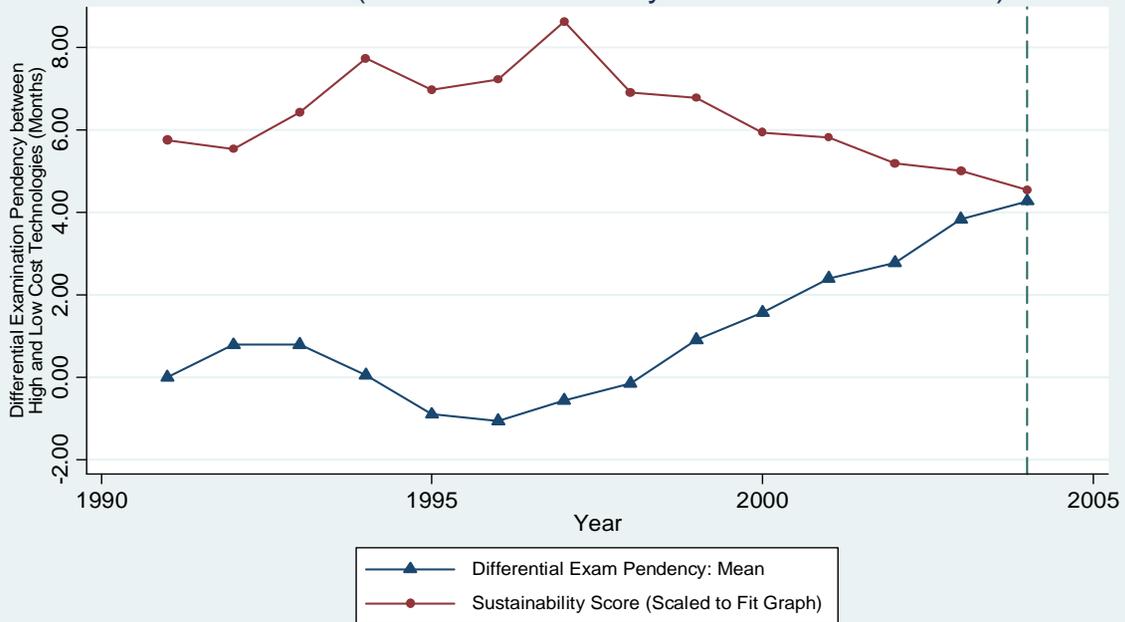
Differential Exam Pendency: Mean
Sustainability Score (Scaled to Fit Graph)

Figure B5: Dynamic Difference-in-Difference Results
Overlaid with Sustainability Score
Including Controls for Application Filings and
Technology-Year Covariates
Dependent Variable = Total Examination Duration
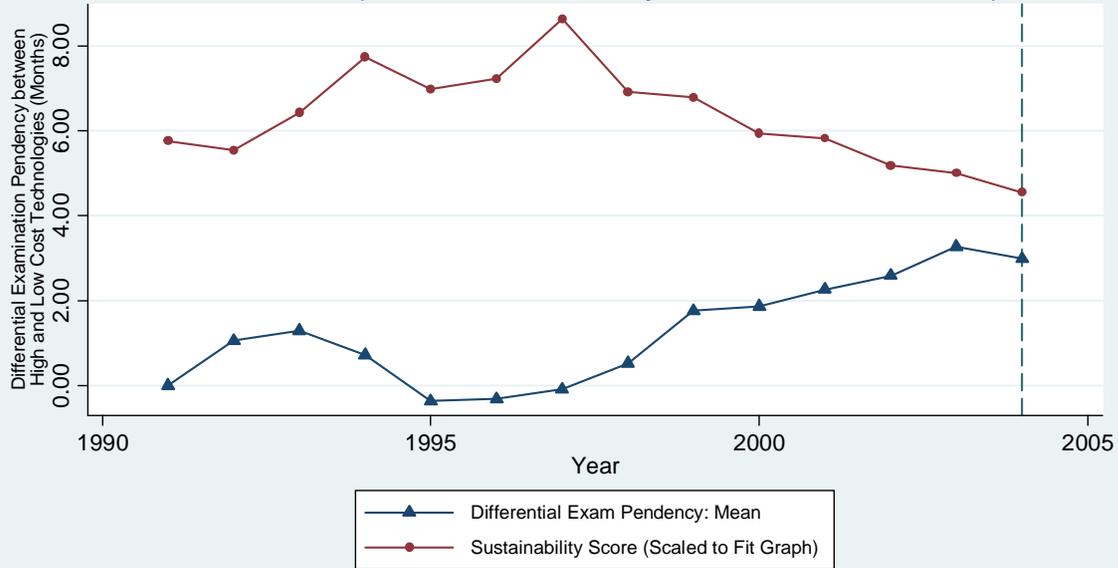1991-2004 (Period of Availability of Total Duration Data)

Differential Examination Pendency between High and Low Cost Technologies (Months)

Year

— Differential Exam Pendency: Mean
— Sustainability Score (Scaled to Fit Graph)

Figure B6: Dynamic Difference-in-Difference Results
Overlaid with Alternative Sustainability Score
Including Controls for Application Filings and
Technology-Year Covariates
Dependent Variable = Total Examination Duration
1991-2004 (Period of Availability of Total Duration Data)

Differential Examination Pendency between High and Low Cost Technologies (Months)

Year

— Differential Exam Pendency: Mean
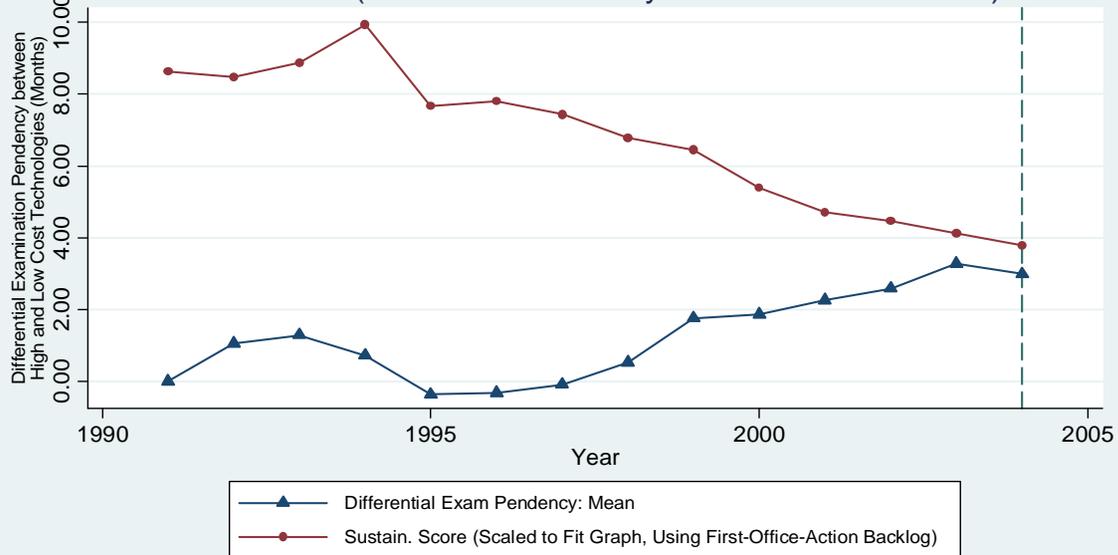— Sustain. Score (Scaled to Fit Graph, Using First-Office-Action Backlog)

Figure B7: Dynamic Difference-in-Difference Results
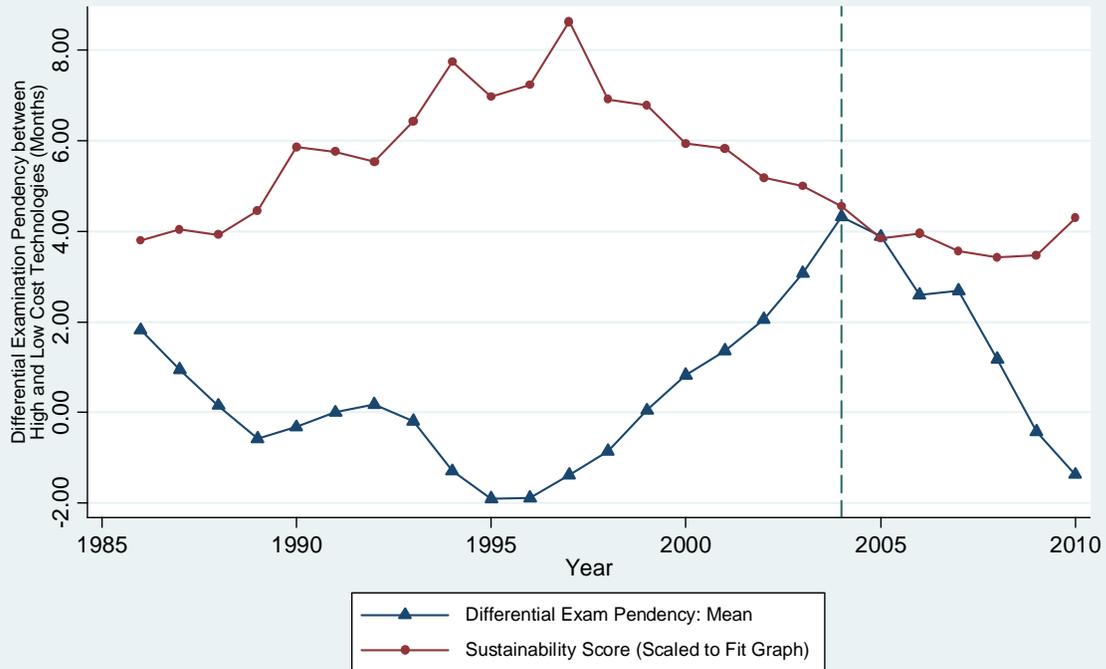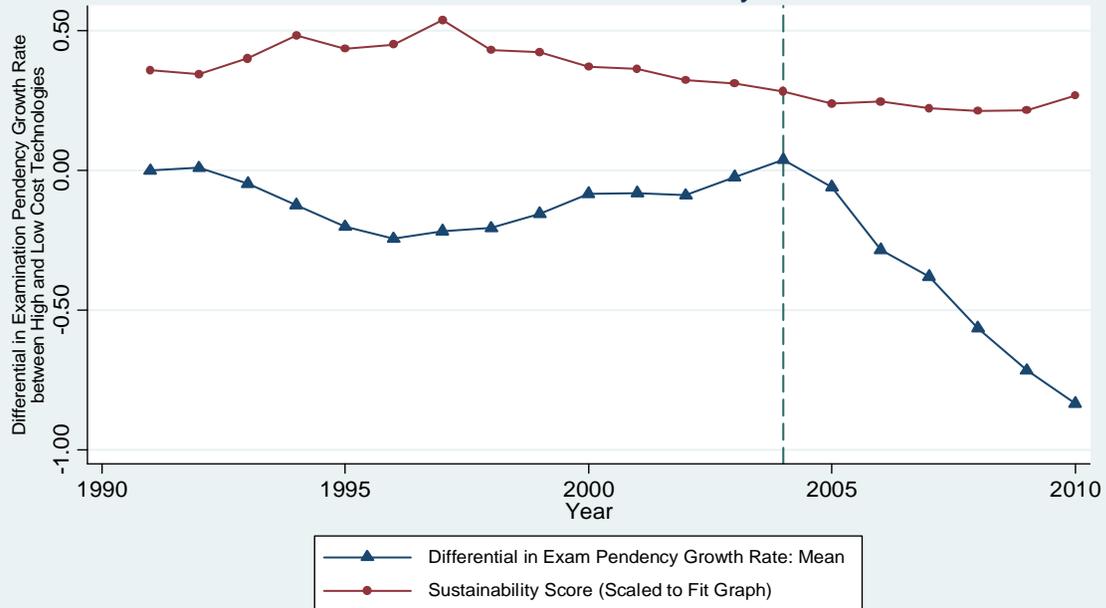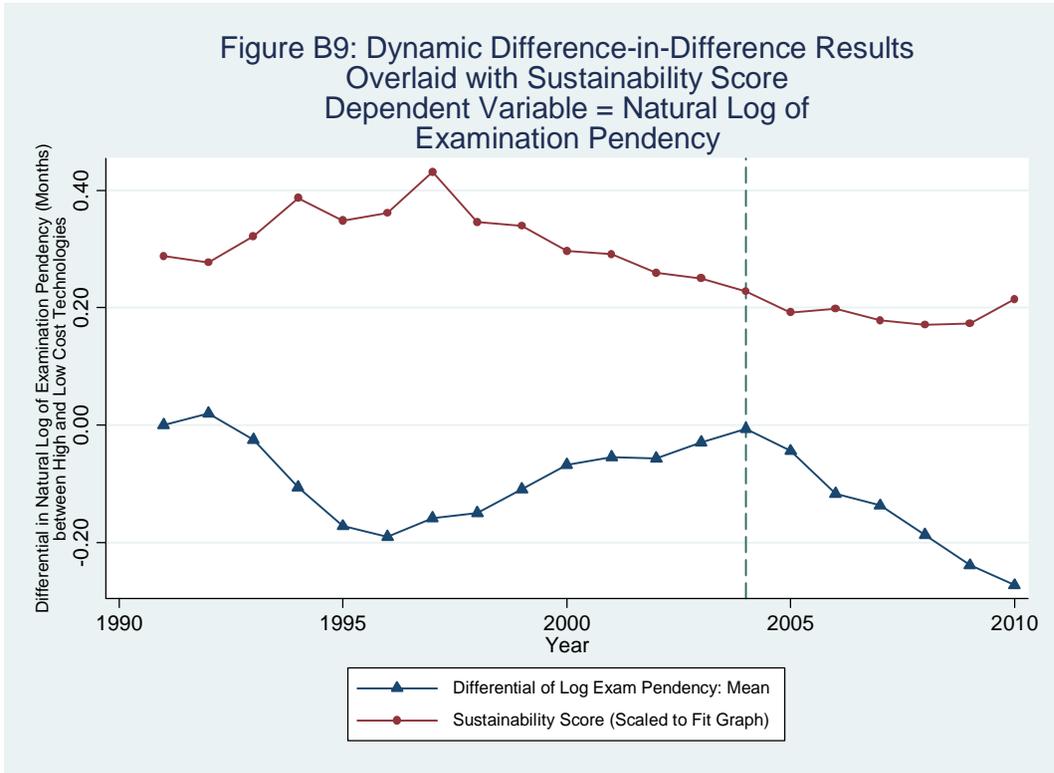Overlaid with Sustainability Score
1986-2010

Figure B8: Dynamic Difference-in-Difference Results
Overlaid with Sustainability Score
Dependent Variable = Growth Rate in
Examination Pendency

Figure B9: Dynamic Difference-in-Difference Results Overlaid with Sustainability Score Dependent Variable = Natural Log of Examination Pendency

1986-2010 Period

In Figure B7, we plot the relevant trends over all periods in which we have available data on examination timing practices—that is, from 1986-2010. For the reasons set forth in the Article, we primarily focus on the 1991-2010 period, as that represents a period in which the Agency universally followed a fully user-fee funded model, allowing us to focus on the fluctuations in the financial health of the Agency and in the fee-diversion reform. In the period prior to 1991, the Agency was only partially funded through user fees with the remainder coming from general tax revenues (nearly a 50/50 split). Even during periods of partially fixed funding of this nature, however, it may still be reasonable to think that the Agency found itself subject to cost considerations. After all, fixed budgets, by their nature may incentivize parties to curb costs to stay within their budgets. As such, at times, the Agency may have also found it worthwhile to considering delaying its costliest examinations.

Similarly, the Agency's desire to avoid excessive expenditures during this pre-1991 period of partial tax funding may have been alleviated during times in which the Agency's incoming fee revenue kept a favorable balance with the costs of the examinations expected of it at that time, even if such fees weren't designed to cover all costs. As such, it is arguably valid to nonetheless hypothesize that the relationship between the Agency's sustainability score and its differential examination pendency patterns would be of the same nature in the 1986-1990 period as it was

during the 1991-2003 period. And, in fact, as demonstrated by Figure B7, this is precisely what we find, only lending further support to the theoretical predictions of this Article.

The sustainability score was rising during the 1986-1990 period. We should note that the sustainability score is constructed so as to simulate the impact of variations in the factors incorporated in Table B5. As indicated above, it does not build in variations over time in the fee levels themselves, a very minor point when considering the 1991-2010 period, as fee fluctuations over that time period are nearly irrelevant for the score's construction. However, the fee levels increased considerably between 1990 and 1991. Nonetheless, the sustainability score over 1986-1990 does validly aggregate the influence of variations in the factors embedded within the score (e.g., annual maintenance rates), in which case it continues to serve as a marker indicative of the financial state of the Agency over that time period. On a final note, we should observe that the increase in the sustainability score over this time period may not be especially surprising considering that the Agency began to collect its first influx of 8-year maintenance fees over this period.

Growth-Rate Specifications

Upon a negative financial shock to the Agency, for the reasons set forth in the Article, the PTO may be inclined to begin delaying examinations in those technologies that cost the most to examine. By delaying some set of examinations that year, the Agency will set in play a longer queue of examinations awaiting review the following year. Thus, if the Agency feels no additional financial pressure in the following year and nonetheless sets out to review the number of applications it may otherwise have the resources to review in a given year, it will continue to maintain the queue that it inherited from the prior year (think of this as your standard traffic effect). Of course, if during this subsequent year, the Agency continues to face the same level of poor financial health that it experienced the prior year, it may be inclined to delay yet another set of applications, only adding to the queue and leading to even longer wait times. In other words, if we see a decline in the sustainability score in one year only to see it plateau at that lower level in subsequent years, it could still be the case that differential examination delays between high- and low-cost technologies could continue to grow.

Now, consider the case where the sustainability score does not simply remain flat in the following year, but instead falls further—that is, the financial health of the Agency gets worse and worse over time. This would perhaps characterize the state of the PTO following 1997, as demonstrated by Figures 1—3. In this case, one might predict that the Agency's inclination to delay some set of applications within high-cost technologies would be an inclination that itself grows over time. In which case, not only would we predict that the differential examination delay between high- and low-cost technologies grows from year to year, but that it grows at an *increasing rate*. Accordingly, to test this prediction, we modify our basic empirical specification which uses examination pendency length levels as the dependent variable and instead estimate a specification that uses the growth rate of the examination pendency length as the dependent

variable. More specifically, we estimate this specification by setting the dependent variable for each technology-year cell equal to the ratio between the examination pendency for that technology and year and the examination pendency for that technology in 1990, the year before the sample period begins. In Figure B8, we then replicate Figure 1 using this modified dependent variable. Though with different slopes obviously, we document a nearly identical pattern of swings in the differential examination pendency measures to that found in Figures 1—3, consistent with these additional theoretical predictions.

Note that those specifications that use the natural log of the examination pendency (see Figure B9) may also be seen as demonstrating the impact of sustainability fluctuations on the rate of change of the differential examination pendency.

ONLINE APPENDIX C

ANALYSIS OF PTO'S NEW FEE STRUCTURE

The PTO has recently utilized its new-fee setting authority to promulgate a new fee schedule. Unfortunately, the Agency's new fee structure does not eliminate the financial risk in examining patent applications, nor is it likely to diminish the Agency's tendencies to delay reviewing its costliest obligations. More specifically, the new fee schedule modestly increases examination fees, substantially decreases issuance fees, and substantially increases renewal fees.[19] The ratio of post-allowance to examination fees remains virtually unchanged.[20] As a result, the PTO will remain heavily dependent on issuance and renewal fees to fund its examination process. As long as the PTO is subsidizing the examination process through fees generated by other activity, the financial health of the Agency will continued to be threatened, as the payment of these post-allowance fees may grow out of step with the examination demands on the PTO. Thus, under the new proposed fee structure, the Agency's financial risk associated with processing applications remains, and the PTO retains a financial incentive to decrease its operational costs by prioritizing the examination of its least costly patent applications.

---

[19] U.S. PTO, Final Rule Setting and Adjusting Patent Fees, 78 FED. REG. 4,212 (Jan. 18, 2013) (to be codified at 37 C.F.R. 1, 41, and 42) (increasing examination fees by 27%, which equates to a $340 increase, decreasing issuance fees by 46%, which equates to an $810 decrease, and increasing the first stage maintenance fee due at 3.5 years by 39%, which equates to a $450 increase, increasing the second stage maintenance fees due at 7.5 years by 24%, which equates to a $700 increase, increasing the third stage maintenance fees due at 11.5 years by 54%, which equates to a $2,590 increase).

[20] Under the proposed fee schedule the ratio of post-allowance to examination fees is approximately 8.5 (960+1600+3600+7400)/1600 = 8.5) whereas under the current fee schedule the ratio of post-allowance to examination fees is approximately 8.6 ((2040+1130+2850+4730)/1250 = 8.6). *Id.* at 4,224–4,225.