Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

Examining Patent Examiners: Present Bias, Procrastination and Task Performance

Ryo Nakajima、Michitaka Sasaki、Ryuichi Tamura

19 September, 2020
DP2020-015
https://ies.keio.ac.jp/en/publications/13244/

Keio University

1858

CALAMVS GLADIO FORTIOR

Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office@adst.keio.ac.jp

19 September, 2020

Examining Patent Examiners: Present Bias, Procrastination and Task Performance
Ryo Nakajima、Michitaka Sasaki、Ryuichi Tamura
Keio-IES DP2020-015
19 September, 2020
JEL Classification: D03, J01, K29, O34
Keywords: patent examination; procrastination; present bias; quasihyperbolic discounting

Abstract

This paper explores the unproductive procrastination behavior of patent examiners, probes whether such behavior is caused by present-biased preferences, and estimates the magnitude. We set out a quasihyperbolic discounting model where a patent examiner is assigned a biweekly quota of patent application reviews and determines the level of effort by the deadline. We estimate the present-bias factor of each patent examiner based on patent prosecution data in the U.S. and find that the proportion of present-biased individuals exceeds the majority. We demonstrate that the job separation rate is higher for less present-biased patent examiners, and a fragmented work quota can improve patent examination quality and timeliness.

Ryo Nakajima
Faculty of Economics, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
nakajima@econ.keio.ac.jp

Michitaka Sasaki
Organization for Research Initiative and Promotion, Tottori University
4-101 Koyama-cho Minami, Tottori 680-8550, Japan
michitakasasaki@gmail.com

Ryuichi Tamura
Faculty of International Studies and Regional Development, University of Niigata Prefecture
471 Ebigase, Higashi-ku, Niigata-City, Niigata 950-8680, Japan
ry.tamura@gmail.com

Acknowledgement : This work was supported by JSPS, Grant-in-Aid for Scientific Research (C) 18K01620 2018-2022.

Examining Patent Examiners: Present Bias, Procrastination and Task Performance

Ryo Nakajima [*] Michitaka Sasaki [†] Ryuichi Tamura [‡]

September 19, 2020

Abstract

This paper explores the unproductive procrastination behavior of patent examiners, probes whether such behavior is caused by present-biased preferences, and estimates the magnitude. We set out a quasihyperbolic discounting model where a patent examiner is assigned a biweekly quota of patent application reviews and determines the level of effort by the deadline. We estimate the present-bias factor of each patent examiner based on patent prosecution data in the U.S. and find that the proportion of present-biased individuals exceeds the majority. We demonstrate that the job separation rate is higher for less present-biased patent examiners, and a fragmented work quota can improve patent examination quality and timeliness.

JEL classification: D03, J01, K29, O34

Keywords: Procrastination, Patent Examination, Present Bias, Quasihyperbolic Discounting

1 Introduction

Procrastination permeates various aspects of our daily lives (Steel 2007; Rozental and Carlbring 2014). We often delay starting an unpleasant but important task until the deadline is close, only to find out we did not complete the task by the deadline or that the task completed in a rush was poorly done. Consequently, procrastination ends up being *unproductive*.

In economics, unproductive procrastination is understood in light of present-biased preferences (Akerlof 1991; O'Donoghue and Rabin 1999). Present-biased persons tend to overweight immediate cost over future reward and defer actions, even when it would be better to act immediately. The lack of self-control in behavior leads to procrastination (O'Donoghue and Rabin 1999). The theory of present-biased preferences is widely accepted to explain procrastination behavior, but there is only limited experimental evidence supporting it.[1] Even field evidence is rare.[2]

[*]Demartment of Economics, Keio University; nakajima@econ.keio.ac.jp (corresponding author).
[†]Organization for Research Initiative and Promotion, Tottori University.
[‡]Faculty of International Studies and Regional Development, University of Niigata Prefecture.

[1]See the review by Frederick et al. (2002) or the more recent survey articles by Cohen et al. (2020).
[2]See, for example, DellaVigna (2018).

An empirical difficulty arises because the *intention-action* gap is difficult, though not impossible, to measure. To verify whether present bias causes procrastination, researchers must compare what an individual planned to do (i.e., intention) with what he or she did (i.e., action): a more extensive gap implies a more serious self-control problem, indicating a more substantial present bias. However, in the typical environment of field studies, the data concerning intended actions is rarely available to empirical analysts.[3]

In this paper, we explore unproductive procrastination in a real work environment, probe whether such procrastination is caused by present-biased preferences, and estimate the magnitude. We overcome the identification challenges by exploiting a unique feature of the U.S. patent examination process. We believe that it provides a suitable empirical testing ground to prove present-bias-induced procrastination because detailed and voluminous administrative data are available from the U.S. Patent and Trademark Office (USPTO).[4] We can learn not only the timing but also the exact content of the patent examination. More crucially, patent examiners' planned actions are institutionally determined and comparable to patent examiners' actual actions. The USPTO imposes on each patent examiner a quota of patent application reviews every two weeks. All the patent examiners ought to steadily and adequately meet the biweekly quota, but some appear to deviate from it and do poor work.

This paper focuses on patent examiners' task *performance* rather than task *completion timing* to assess present-biased time preferences. Previous economic studies using field data take task completion on or near deadline as evidence for present-bias-induced procrastination (e.g. Martinez et al. 2017; Frakes and Wasserman 2020). However, people delay work until near the deadline if doing so has an option value (Dixit and Pindyck 1994). Therefore, delayed task completion may not be attributed to present-biased preferences.[5] By construct, we use variation in task performance to enhance the power of identification. Intuitively, if a delay is systematically accompanied by substandard performance, it will not be bolstered by option value theory but rather by present-bias theory because postponed decisions become counterproductive.

A possible concern of our performance-based approach is unobserved patent examiners' personal traits, such as inadequate ability. To the extent that these traits are correlated with poor performance, they may confound present bias. To circumvent this problem, we exploit exogenous variation in *deadlines*. Given that the biweekly quota is constant, holiday-induced downtime reduces the time that patent examiners can use for patent application reviews and thus effectively shortens the quota deadline.[6] The shortened deadline affects the task performance of *exponential discounting* patent examiners who opt for reduced work hours under the limited timeframe. On the other hand, the deadline pressure has a negligible effect on the performance of *hyperbolic discounting* patent examiners because they procrastinate and start a task later in any case. These considerations lead us to exploit the differences in the deadline responsiveness of task performance to identify present-biased preferences.

[3] A handful of experimental studies in controlled laboratory environments use this or similar strategies to identify present-biased time preferences (Augenblick et al. 2015; Augenblick and Rabin 2019). A common strategy is to systematically garner information on the actions that the subjects planned to choose before the experiment runs and contrasting it with that on the actions they chose in the experiment.

[4] It has been argued that U.S. patent examiners may steep in the vice of procrastination. For example, the Commissioner of Patents at the USPTO acknowledged that the high volume of patent decisions made just before the examination deadline could be attributed to patent examiners' procrastination behavior by saying, "It can be a bad habit, in some situations. It is procrastination. Moreover, in others, as I said, it could be misconduct if the work is incomplete" (Joint Hearing before the Committee on the Judiciary and Committee on Oversight and Government Reform, 2014). Nonetheless, much of the evidence is anecdotal rather than statistical.

[5] In recent work, Heidhues and Strack (2019) show formally, using a dynamic discrete choice framework, that time preferences cannot be identified based on the pattern of task completion alone.

[6] The stability assumption on the biweekly patent examination quota is empirically tested in Section 3.3.

Motivated by the abovementioned insight, we consider a simple quasihyperbolic discounting model, often referred to as a $\beta - \delta$ model (Laibson 1997). We use a continuous version of the model, where an agent is assigned a quota of tasks and determines their level of effort each day to complete the tasks (Fischer 1999, 2001; Herweg and Müller 2011). In the continuous effort allocation framework, a hyperbolic discounting agent is characterized by more steeply increasing effort, less total effort, and reduced performance on completed tasks compared to an exponential discounting agent.

We test the model's predictions using data on patent prosecution activities in the U.S. By measuring the degree of a patent examiner's underperformance in terms of the log-odds that his or her initial patent application review is unsuccessful, we implement regression analysis that allows for individual heterogeneity in the *level* of task performance and in the *difference* in task performance concerning the deadline change. The estimation results show that a poorly performed patent examiner task is *less* affected by deadline reduction, which agrees with our present-bias-induced procrastination model.

We then proceed to estimate the time preference parameters of the quasihyperbolic discounting model. The panel structure of the patent examination data allows us to follow the performance profile for each patent examiner and thus enables us to estimate the *individual-specific* present-bias factor. Since the present-bias factors are estimated as the number of patent examiners, the dimensionality problem occurs in the estimation process. We address this issue by means of a Bayesian inference approach with the Markov chain Monte Carlo (MCMC) method.

The estimation results provide strong evidence that present bias is widespread among patent examiners. Specifically, more than half the patent examiners have a present-bias factor less than one. The findings are shown to be robust against alternative specifications, including the prior distribution, utility curvature, and reward setup. Additionally, we validate our empirical findings using an out-of-sample prediction method.

Finally, we relate our findings to policy issues on patent system reform (Jaffe and Lerner 2004; Lemley and Shapiro 2005). We draw two policy implications from our analysis. First, given that attrition rates are significantly higher for less present-biased patent examiners than for more present-biased patent examiners, the employee retention policy should be targeted to the former group. Second, reducing the patent examination quota can improve patent examination quality and timeliness. A simulation result shows that if the currently adopted two-week quota is cut in half with a one-week deadline, initial patent examination failure may decline by approximately 30 percent, and the patent term adjustment period may be reduced by approximately one week. For a pharmaceutical patent, this reduction could yield substantial consumer benefits.

Our paper is connected to several strands of the literature. First, the studies relevant to our research are those eliciting present-biased preferences from time-inconsistent behavior associated with a real-effort task. Robust support for present bias has been provided by laboratory-based experiments, where experiment participants perform unpleasant tasks and are paid upon completion (Augenblick et al. 2015; Augenblick and Rabin 2019). A handful of field experiment studies also explore the time-inconsistent behavior of college students assigned time-consuming tasks (Ariely and Wertenbroch 2002; Bisin and Hyndman 2020). Since the information on intended actions is often not available in field experiments, the demand for a self-imposed deadline is used as smoking-gun evidence for (sophisticated) present bias.

As stated earlier, little evidence has been provided by observational field studies to support the theory of present-biased preferences on procrastination for real work effort, but some impressive results have been obtained in recent work. For example, Martinez et al. (2017) explore procrastination behavior observed in tax return filing and find substantial evidence of present bias. This finding is corroborated by an out-of-sample prediction, where a quasihyperbolic discounting model outperforms an exponential discounting model in predicting people's response to a policy change concerning tax return incentives.

Concerning the research subject, the study most closely related to ours is Frakes and Wasserman (2020). The authors investigate the U.S. patent examination process and conclude that patent examiners' uneven spike-looking work pattern can be attributed to their present-biased preferences.[7] Our paper differs in the following ways: (i) we provide quantitative evidence for present-bias-induced procrastination by estimating individual-specific present-bias factors, and (ii) we perform counterfactual simulations based on the estimated parameters to evaluate whether some policies would improve patent examination efficiency.

Second, our paper complements studies that estimate the present-bias factor using observational data.[8] In this context, a wide range of intertemporal economic decisions have been studied, including consumption and saving (Laibson et al. 2015), job search (Paserman 2008), medical examination (Fang and Wang 2015), labor supply and welfare participation (Fang and Silverman 2009; Chan 2017), and tax return filing (Martinez et al. 2017). These studies use a structural approach to estimate $\beta - \delta$ parameters in a dynamic discrete choice model. The discrete choice-based approach, which is promising in certain situations, may not suit our research needs.[9] We propose a novel approach to estimate the present-bias factors based on a continuous effort allocation model.

Third, our paper mirrors a batch of studies that attribute the variability in patent examination quality to patent examiners' traits and the patent examination environment(Cockburn et al. 2002; Lemley and Sampat 2012; Frakes and Wasserman 2017). Specifically, our finding that the patent examination deadline may impact patent examination performance is consistent with Frakes and Wasserman (2017), who find that tighter patent examiner time constraints result in low patent examination quality.

Finally, the results of our paper also have implications for the literature on the work performance of "experts," including juridical judges (Coviello et al. 2015), journal referees (Chetty et al. 2014), emergency doctors (Chan 2018), and paramedics (Brachet et al. 2012). While the contexts and perspectives are diverse, a common thread of these studies is that proper task management can enhance task productivity, which is in line with our conclusion that task efficiency may improve if a task quota is appropriately subdivided.

This paper proceeds as follows. Section 2 provides institutional background on the patent examination process in the U.S. The data and summary statistics are described in Section 3, and the regression-based evidence for present-bias-induced procrastination is included in Section 4. Section 5 presents a behavioral model of a patent examiner's worktime allocation based on a $\beta - \delta$ framework and illustrates the qualitative properties via simulation. Section 6 performs Bayesian inference for the model parameters to elicit the individual-specific present-bias factor of each patent examiner. Section 7 presents policy simulations concerning patent examination quality and patent pendency. Finally, Section 8 concludes.

[7]Frakes and Wasserman (2020) identify present-biased preferences of U.S. patent examiners by focusing on the near-deadline clustering of their patent review completion. Of critical importance is the regression finding that the clustering tendency is enhanced by the onset of the telecommuting program adopted by the USPTO, which is expected to make patent examiners' self-control ability more essential and further deter the actions of present-biased patent examiners. They also provide additional empirical evidence that patent reviews made near the deadline tend to be quick and of low quality, which can be attributed to present-bias-induced time crunch.

[8]Extensive surveys are provided by DellaVigna (2018) and Cohen et al. (2020).

[9]First, a single patent examination task involves multiple actions and is not a discrete once-for-all event. Second, patent examiners may finish review tasks earlier than the deadline but submit them all at once on the deadline date, as argued in Frakes and Wasserman (2020). Consequently, the information on task completion may be noisy in that the completion time recorded in the patent examination data does not necessarily coincide with the actual completion time.

2 Institutional Background

2.1 Patent Examiners

Patent examiners are the gatekeepers protecting intellectual property rights. They review a patent application to determine whether it satisfies the statutory requirements for patentability, including novelty and nonobviousness.[10]

The examination of a patent application demands a high level of technical knowledge and expertise. Given the highly technical nature of patent applications, the USPTO has nine Technology Centers (TCs) specializing in specific fields of technology. Each TC is further divided into smaller workgroups, called Art Units, which usually consist of 10 to 15 patent examiners reviewing patent applications in similar technological subfields.[11]

2.2 Patent Examination

The examination process is systematic and standardized.[12] The supervisory patent examiner (SPE), who heads the Art Unit, usually assigns the application to a patent examiner within the Art Unit in a random manner.[13] The applications allocated to a specific patent examiner are retained in a queue, called a "docket," from which he or she takes an application for examination on a first-in, first-out basis.[14]

A patent examiner's decision on patentability is called an office action. There are two major types of office actions: (1) a *first office action* is the initial examination of whether to allow or reject the claims, and (2) a *final office action* is the ultimate decision on the patentability of the application. If the claims are judged patentable in the first round of the review, the application is granted a patent. In most cases, however, patent examiners initially reject some or all claims and cite all the possible grounds for rejection in the first office action.[15] Upon receiving the applicant's response to the first office action's rejection, patent examiners issue the final office action, where the application is either allowed or rejected. Since patent examiners are encouraged to address all the statutory issues in the first office action and complete application review within *two* office actions, *the second office action is usually final*.[16]

2.3 Performance Appraisal

The patent office regularly tracks patent examiners' work performance by several metrics. The most important is the work output metric, referred to as the "count." Patent examiners receive a count at two different times: (i) when the first office action is issued for an application and (ii) when an application is disposed of by

[10] The U.S. Patent Act (Code 35) sets forth the general standards for patentability in Section 101.

[11] For example, TC 1600 handles patent applications in the category of Biotechnology and Organic Chemistry, whereas Art Unit 1641, a subdivision of TC 1600, is a workgroup of examiners who review patent applications relating to peptide or protein sequences.

[12] The process is documented in the Manual of Patent Examining and Procedure (MPEP).

[13] Lemley and Sampat (2012) mention this random assignment rule. Based on interviews with SPEs, they conclude that there is no evidence of the deliberate selection or assignment of patent applications.

[14] See MPEP §708 regarding the order of patent examination.

[15] Marco et al. (2017) shows that only approximately 10 percent of patent applications are granted at the first stage of examination in the U.S.

[16] The USPTO's policy is called "compact prosecution." MPEP §2173.06 stipulates that "Under the principles of compact prosecution, the examiner should review each claim for compliance with every statutory requirement for patentability in the initial review of the application and identify all of the applicable grounds of rejection in the first office action to avoid unnecessary delays in the prosecution of the application."

allowance or abandonment. The patent office sets a productivity goal based on the count in the allotted time. Specifically, patent examiners are required to meet a target count in a *two-week period*. The biweekly quota is tailored individually and is determined based on a patent examiner's position and the technological complexity of applications.[17]

Patent examiners are scrutinized as to whether they meet the biweekly quota: noncompliance heightens the risk of dismissal from the patent agency. If patent examiners meet the biweekly quota for an average of 10 percent or more of the year, they are eligible for a special bonus that adds to their base salary. According to a report by the Office of the Inspector General (OIG), all tenured patent examiners had an annual productivity goal of 100 percent in the early 2000s. Approximately 90 percent of them were in the target range of 110 to 119 percent of the annual quota attainment and received special bonuses. An extra reward is set for even higher achievement levels—120 percent more on target—but fewer than 10 percent could attain this level.[18]

Another metric is used to evaluate examination quality. The SPEs or specialist reviewers conduct a quality assessment of randomly selected office actions by checking whether patent examiners made "errors" in the applications' patentability decisions.[19] The quality rating of a patent examiner affects his or her annual compensation and promotion possibility to an upper position.[20]

Ostensibly, both the quantity and quality goals of the patent examination are assessed. However, the quality aspect has been downplayed compared to the quantitative aspect.[21] The USPTO monitoring policy of prioritizing quantity over quality may provide room for patent examiners to game the system. It is widely believed among practitioners and policymakers that patent examiners often knowingly issue incomplete office actions to meet the productivity goal (Stephenson 2008). This ill-practice is known as "patent mortgaging."[22] Patent examiners, especially those in need of counts, tend to reject an application without sufficient review in the first round. Such quick and baseless rejection is referred to as a "shotgun rejection" (Frakes and Wasserman 2020).

Patent examiners' readiness to opt for premature patent application reviews motivates us to focus on the second office action to evaluate the quality of the first office action.[23] If a patent examiner rejects an

---

[17] The specific calculation of the count is described in detail in Marco et al. (2017). Moreover, the biweekly quota was unchanged from 1976 until 2010, according to a report published by the National Academy of Public Administration (NAPA) in 2005 and a report published by the Government Accountability Office (GAO) in 2016.

[18] See the report by the OIG (2004) . The detailed figures are as follows: 41 percent of examiners were between 100 percent and 109 percent of their productivity goal; 51 percent met the target in the range between 110 percent and 119 percent; and 8 percent of patent examiners attained a target higher than 120 percent.

[19] An examination "error" is defined as a clear instance where a patent examiner does not comply with the examining standards outlined in the *Patent Examiner Performance Appraisal Plan*, which is available from `http://www.popa.org/static/media/uploads/uploads/examiner-pap-guidelines-04_19_12-508.pdf` (accessed August 4, 2020). The quality assessment occurs at two different times: (i) after the first office action is issued on an application and before the final office action (called in-process review) and (ii) after an allowance is granted (called allowance review).

[20] A patent examiner's position is based on the General Schedule (GS) pay scale, which is the predominant pay scale in the U.S. federal government. New hires usually enter between GS-7 and GS-9 and are promoted on the GS scale as they accumulate experience. A patent examiner's authority to sign off on his or her decision on a patent application, referred to as signatory authority, depends on the GS level. "Junior" (secondary) examiners at GS-13 or below must have their decisions checked and authorized by a "senior" (primary) examiner above GS-14. A senior examiner with full signatory authority, on the other hand, may allow or reject a patent application without any additional check.

[21] NAPA (2005, 2015) discuss the appraisal procedures of patent examiners' performance in detail. Furthermore, the OIG (2015) provides detailed quantitative evidence that the actual error rate is systematically underreported in the quality review.

[22] See OIG (2015) and NAPA (2015) for a detailed explanation.

[23] Frakes and Wasserman (2020) propose a nonfinal second office action as a proxy measure for the failure of the first office action.

application without sufficient grounds in the first round of review, the office action is deemed to contain an error by the quality assessment. In that case, a patent examiner must redo a search of prior art to make up for his or her prior baseless decision in the next round of review. The second office action will *not be the final* action in that case.[24] Therefore, a *nonfinal second office action* indicates a failed first office action.[25]

Finally, patent examiners have been given broad discretion in choosing their working time. They must work eighty hours in two weeks but do not have to work eight hours a day.[26] Furthermore, even though patent examiners must self-report their work hours in a timekeeping system at the end of each biweekly period, there is no mechanism to verify whether the reported hours were correct. The OIG (2015) alleged that lax work monitoring might result in time and attendance abuse among patent examiners.

3 Data and Statistics

3.1 Data Construction

The primary data come from the USPTO Patent Examination Dataset based on information from the Public Patent Application Information Retrieval (PAIR). The PAIR provides detailed information on all patent applications filed with the USPTO that have been published.[27] We limit our analysis to a sample in a specific technical field because the original sample size is too large for a standard computing environment. In the following, accordingly, the analysis is performed using patent applications handled by TC 1600: Biotechnology and Organic Chemistry.

We consider only the *first cycle* of patent examination since patent applications are transacted through a convoluted route.[28] The cycle starts when an application is assigned to an examiner's docket and ends when a *first office action* is taken. We select applications with a first office action issuance between January 1, 2001, and December 31, 2009, because the USPTO management changed the patent examination rule after 2010.[29] We further restrict the sample along several other dimensions: we limit patent applications to regular utility applications, we exclude those applied through the Patent Cooperation Treaty system, we eliminate those reviewed by patent examiners with less than one year of experience, and we remove those whose prosecution records are incoherent. Eventually, a sample of approximately eighteen hundred thousand applications reviewed by 710 examiners in 55 Art Units remains.

We supplement our analysis using granted patent information available from the USPTO Patents View

[24]MPEP §707.07 (A) prescribes that "second or any subsequent actions on the merits shall be final, except where a patent examiner introduces a new ground of rejection that is neither necessitated by applicant's amendment of the claims nor based on information submitted in an information disclosure statement."

[25]Notably, a patent examiner does not receive any count when he or she issues a nonfinal second office action, for it is neither the first office action nor the final office action that leads to any disposal actions.

[26]The majority of patent examiners participate in a flextime schedule that started in the early 2000s, called the Increased Flextime Program (IFP). Under the IFP, they may vary the number of hours worked each day and the days worked each week, as long as they (i) meet the eighty-hour requirement and (2) satisfy core hour requirements (NAPA 2015) .

[27]The URL of the data source is `https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-examination-research-dataset-public-pair`. Graham et al. (2015) provide information and supporting documents about the dataset.

[28]For a detailed description of the patent examination process, see Marco et al. (2017).

[29]According to U.S. Government Accountability Office (GAO), the patent office adjusted the time allotted to examiners between fiscal years 2010 and 2012 and gave all patent examiners a total of 2.5 additional hours per application.

website.[30] The data are used to construct patent examiners' characteristics, including years of experience and positions, and to make quality-related patent application variables based on document-based information.[31] [32]

3.2 Descriptive Statistics

Table 1 presents summary statistics. Panel A shows statistics on patent examiners. The number of patent examiners in each year was between 400 and 600, and the number increased over time. The number of years of experience also increased because the early 2000s experienced active recruitment, while turnover declined over time. Panel B shows throughput statistics on the patent application reviews for a single patent examiner. As presented in the previous section, a first office failure occurs when the subsequent office action is not final. The number of first office actions issued by one examiner is relatively stable, and the failure rate of the first action per patent examiner remains between 14 and 15 percent, except in the final year of the study.

We now turn to the association between the timing and quality of the first office actions. Figure 1 presents (i) the distribution of the days that the first office actions are issued over weekdays in a biweekly quota period and (ii) the time course of the first office action failure rate for the same period.[33] Two findings emerge. First, the first action dates tend to cluster near the deadline.[34] Approximately half the first office action decisions are made on the last day of a biweekly period. Second, the quality of examination declines as the deadline approaches. At the end of the biweekly period, the rate of first office action failure is roughly 1.5 times greater than that at the beginning. In other words, the patent examiner's postponed actions do not appear to lead to good results and are counterproductive.[35]

3.3 Assumption of Quota Stability

We now verify a premise concerning the biweekly patent examination quota. In the analysis that follows, we assume that holidays do not alter the quota. Only patent office insiders know precisely how the quota is determined or fine-tuned. Therefore, we assess the validity of the assumption on quota stability by scrutinizing whether the number of first office actions in a biweekly period is affected by the existence of holidays; we let the data speak for themselves.

As a preliminary step, we categorize the two-week quota period by whether it includes a U.S. federal holiday. In the *standard* quota period, patent examiners use the full 10 weekdays to review patent applications. On the other hand, in *short* quota periods that include at least one holiday, less than 10 weekdays are available.

[30] The URL of the data source is `https://www.patentsview.org/download/`.

[31] We calculate a patent examiner's years of experience using the date of the earliest first office action issued by him or her as a career starting point. Moreover, we determine an examiner's position, either senior (primary examiner) or junior (secondary examiner), based on the signatory authority recorded in the issued patent data.

[32] We use the number of patent claims, the number of words, and word types in the patent document.

[33] According to the patent examiner's *Docket Management Manual*, all reviewed patent applications must be submitted on the first Monday after the two weeks given as the quota period. In this paper, we set that Monday, often referred to as *Count Monday*, as the deadline for each biweekly quota. The *Docket Management Manual* is available from `http://popa.org/static/media/uploads/uploads/DocketManagementManualVersion5.pdf` (accessed August 4, 2020).

[34] Figure G.1 in the Appendix shows the daily number of first office actions for TC 1600 for the period 2004-2009. A biweekly cyclic trend with a sharp spike on the deadline date, or the *Count Monday*, represented by a solid circle in the figure, is evident. The same patterns are observed for other years.

[35] Frakes and Wasserman (2020) have reported these features of patent examination.

Let S_{bt} be a binary indicator of a short timeframe, which is one if biweekly period b in year t includes a holiday and is zero otherwise.

We estimate the following regression model:

$$\log\left(\frac{Num\_FOA_{bt}}{Num\_Exam_{bt}}\right) = \alpha_b + \alpha_t + \alpha_1 S_{bt} + \alpha_2 \log(Backlog_{bt}) + \varepsilon_{bt} \tag{1}$$

where the subscript b indexes the biweekly period and subscript t indexes the year. The dependent variable is constructed as the ratio of Num_FOA_{bt}, which represents the total number of first office actions, relative to Num_Exam_{bt}, which represents the total number of patent examiners. The explanatory variable of interest is the short timeframe indicator S_{bt}. The variable $Backlog_{bt}$ represents the total number of pending patent applications as of the beginning of biweekly period b in year t, which is further divided into those for new applications and those for amended applications.[36] We include a variable to account for the impact of backlogs on patent examination capacity (Mitra-Kahn et al. 2013). Additionally, a biweekly period dummy α_b and year dummy α_t are included to control for time trends, and ε_{bt} represents an error term.

Notably, the variation in the short timeframe indicator S_{bt} comes from the discrepancy between the quota cycle and the calendar cycle. Since holidays are set on specific dates or specific weeks of the year, the same biweekly period in different years could have a different number of holidays.[37]

Table 2 presents the estimation results. In column 1, we report estimates for the baseline specification. In columns 2-5, we add an event dummy variable that takes a value of one if biweekly period b in year t is on Thanksgiving. In columns 3-5, we allow the number of first office actions per patent examiner to vary with the number of patent examiners. In columns 4 and 5, we add lead and lag variables of the short period dummy. While several specifications are used, the estimated coefficients of the short timeframe dummy, S_{bt}, are invariably *not* statistically significant.

The combined evidence indicates that patent examiners are given the same quota and issue the same number of first office actions even in biweekly periods with holidays. Nothing can be said, however, about the change in task *performance* in a short timeframe. Do patent examiners sacrifice their work quality to meet the quota in such a time-scarce environment? We will investigate this question in the next section.

4 Regression Evidence

This section performs regression analysis to understand patent examiners' task performance through the lens of their present-biased preferences.

Our empirical strategy is based on the association between present bias and task performance. Notably, the theory of present-biased preferences predicts that a present-biased person will delay the start of a quota and complete the work near the deadline (O'Donoghue and Rabin 1999). Consequently, the task performance of a present-biased person is substandard (Herweg and Müller 2011). How does the situation change when the task deadline is shortened? The answer is not much. A present-biased person will continue to delay the

[36] These two types of the patent application "backlogs" are associated with first action pendency and post first action pendency, respectively. See Mitra-Kahn et al. (2013) for a more detailed definition of patent backlogs and their impact on growing examination pendency in the U.K. and U.S.

[37] The biweekly quota cycle does not start on the same day every year, so it shifts slightly each year. For example, the second biweekly quota cycle period in 2005 (from January 11 to 24) includes Martin Luther King Jr. Day as a federal holiday, but the same biweekly period in 2010 (from January 5 to 18) does not. Therefore, the former is considered a short biweekly period, whereas the latter is not.

task even in time-scarce circumstances. Therefore, the performance will *not* be sensitive to increased deadline pressure. If we apply this intuition to the case of patent examination, we obtain the following predictions concerning a present-biased patent examiner's task performance: it will be (i) inadequate and (ii) insensitive to the deadline.

4.1 Empirical Specification

To assess whether the predictions stated above are supported by the data, we set out the following regression model:

$$\log(Odds_{aibt}) = \alpha_b + \alpha_t + (\phi_{0i} + \alpha_{0X} X_{0,abt} + \alpha_{0W} W_{0,ibt}) + (\phi_{1i} + \alpha_{1X} X_{1,abt} + \alpha_{1W} W_{1,ibt}) S_{bt} \tag{2}$$

where a indexes the application, i indexes the patent examiner, b indexes the biweekly period, and t indexes the year. The dependent variable $\log(Odds_{aibt})$ is the log-odds that the first office action of patent application a reviewed by patent examiner i in biweekly period b in year t is unsuccessful, which is referred to as the *log-failure odds* in the following.[38] The regressors include the short timeframe indicator S_{bt}, which takes a value of one if biweekly period b in year t contains a holiday. The biweekly period dummy α_b and year dummy α_t control for time trends.[39] The variables $X_{\bullet,abt}$ and $W_{\bullet,ibt}$ represent the characteristic vectors regarding application a and patent examiner i.[40] The complete list of control variables in the regression is presented in Table G.1 in the Appendix. We collect all coefficients of the time trend dummies and observed patent examiner and application characteristics into a single vector $\boldsymbol{\alpha}$.

The *patent examiner-specific* parameters, ϕ_{0i} and ϕ_{1i}, are our parameters of interest. The parameter ϕ_{0i} is interpreted as the log-failure odds of patent examiner i after the effects of the observed factors are accounted for. On the other hand, the parameter ϕ_{1i} is interpreted as the *ceteris paribus* difference in the log-failure odds of patent examiner i between the situations under a short timeframe ($S_{bt} = 1$) and standard timeframe ($S_{bt} = 0$). Since the two parameters are estimated by an empirical model (as opposed to being predicted by a theoretical model), we refer to them as the *empirical log-failure odds* in level and difference.

How are the empirical log-failure odds, ϕ_{0i}, and ϕ_{1i}, related? Recall the following predictions concerning the association between present bias and task performance:

1. A patent examiner with higher present bias works shorter hours because of procrastination and therefore has higher log-failure odds, implying that patent examiner i with higher present bias tends to have a larger positive value of ϕ_{0i}.

2. The shorter deadline will have a limited impact on a higher present-biased patent examiner's task performance, implying that patent examiner i with higher present bias tends to have a smaller positive value of ϕ_{1i}.

[38]To be more precise, a patent application reviewed by a patent examiner is unsuccessful if the first office action fails or, equivalently, as explained in Section 2.3, the second office action is nonfinal.

[39]In practice, the regression includes trend dummy variables at the quarter and month levels and the Thanksgiving dummy variable defined in the previous section.

[40]For ease of interpretation, they are both transformed as the deviation from the overall mean. We denote the original variable as $X^o_{\bullet,abt}$ and the overall mean as $\bar{X}^o_{\bullet}$ for all applications. We then define $X_{\bullet,abt} = X^o_{\bullet,abt} - \bar{X}^o_{\bullet}$. The same definition is applied to $W_{\bullet,ibt}$.

Therefore, we predict that the two empirical log-failure odds parameters, ϕ_{0i} and ϕ_{1i}, are *negatively* correlated. The prediction presented here is supported by a model analysis in the next section.

4.2 Baseline Evidence

We construe equation (2) as a multilevel logit model with a random intercept and a random coefficient. Following the standard procedure of multilevel analysis (Gelman and Hill 2006; Snijders and Bosker 2011), the random intercept-coefficient vector (ϕ_{0i}, ϕ_{1i}) is assumed to follow a bivariate normal distribution with means (μ_0, μ_1), standard deviations (σ_0, σ_1), and correlation σ_{01}. Let $\boldsymbol{\sigma} = (\sigma_0, \sigma_1, \sigma_{01})$ be the standard-deviation-correlation matrix. The random vector (ϕ_{0i}, ϕ_{1i}) is further assumed to be uncorrelated with any of the explanatory variables, $X_{\bullet,abt}$ and $W_{\bullet,ibt}$.

We estimate the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ by means of a maximum-likelihood method. Table 3 reports the main results.[41] Panel A presents the marginal effect of the short timeframe indicator S_{bt} on the first office action failure, while panel B shows the estimates of the standard-deviation-correlation matrix $\boldsymbol{\sigma}$. As measures of goodness-of-fit, the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the area under the receiver operating characteristic curve (AU.ROC) are presented in the last panel.[42]

For robustness considerations, we apply several specifications. We begin with a simple specification in column 1, where neither application-specific nor patent examiner-specific *observed* heterogeneities are allowed to interact with the short timeframe dummy.[43] We then proceed to flexible models with more interaction terms. In column 2, the interaction terms include the examiner characteristic variable vector, $W_{1,ibt}$. In column 3, the interaction terms are expanded to include both the application and examiner characteristic vectors, $X_{1,abt}$ and $W_{1,ibt}$.

The estimation results reveal several findings. First, the results are qualitatively unchanged under the different specifications. Nonetheless, the specification in column 3 appears to be overfit because it contains 264 regressors, 1.5 times as many as the specifications in column 1 and 2, but it does not exhibit substantial improvement in the goodness-of-fit measures. Therefore, we consider the model in column 2 as the preferred specification in what follows. Second, the marginal effects of the short timeframe on examination failure are positive and statistically significant (p-value of 0.008 in the preferred specification). As an overall average, the failure odds increase approximately one percent if an application is reviewed in a short biweekly period rather than in a standard biweekly period. Third, the estimates of σ_0 and σ_1 are positive and statistically significant (both p-values are 0.000 in the preferred specification), which indicates that substantial unobserved heterogeneity exists in the log-failure odds in both levels and differences. Last, and most importantly, the estimate of correlation σ_{01} is negative and statistically significant (p-value of 0.009 in the preferred specification), which strongly corroborates our prediction of the empirical log-failure odds.

[41] Estimates of the coefficients for variables other than those shown in Table 3 are reported in Table G.2 in the Appendix. However, the estimates for the time trend dummies (i.e., year, quarter, month, Thanksgiving), biweekly period dummies, art unit dummies, and technological class dummies are not provided. The estimates on the interaction term of S_{bt} and $(X_{1,abt}, W_{1,ibt})$ are not also reported in Table 3. The estimation results are available from the authors upon request.

[42] AU.ROC, which ranges from zero to one, is a measure of a model's ability to discriminate the binary outcomes of the observation (Hosmer and Lemeshow 2000). A higher AU.ROC value indicates better discriminatory power of the model. In particular, a value of 0.5 indicates no more discriminative power than a coin flip, while a value of 1.0 indicates perfect discrimination. Since the calculated AU.ROC value exceeds 0.7 for our estimates, the model's fit is considered satisfactory in terms of discriminatory power.

[43] This situation corresponds to the case where the set of interaction variables $\{X_{1,abt}, W_{1,ibt}\}$ is empty.

4.3 Additional Evidence

There are several concerns in interpreting the baseline regression results as evidence of present-biased preferences.

The first concern is that the regression model may be misspecified. The assumption of the random parameters' joint normality or the assumption of no correlation between the random parameters and the covariates may not hold. To address this issue, we estimate the logit regression, where the intercept-coefficient parameters, (ϕ_{0i}, ϕ_{1i}), are treated as *fixed effects*. While the fixed-effect estimates become less precise than the random-effect estimates in general, the distribution assumption is much weaker than that for random effects. The results are reported in column 4 of Table 3. No evidence is provided to refute the interpretation of the results stated above.

The second concern regards the endogeneity of the regressors. In the baseline specification, it is assumed that the short timeframe dummy, S_{bt}, is *exogenous* in the sense that it is uncorrelated with unobserved factors that could influence the log-failure odds. However, this assumption may be violated if the examiners *intentionally* manipulate the order of examination or conduct application sorting.[44] For instance, patent examiners may prefer reviewing an "easy" application in a time-crunch situation.[45] This scenario is considered a typical omitted variable problem because the regression model fails to control for application characteristics by which patent examiners are guided to arrange the order of patent review but are unobserved by researchers.

One general way to mitigate the omitted-variable problem is to include as many attributes as believed to be appropriate. We exploit the dataset of *granted patents*, which includes detailed information on the patent contents.[46] The estimation results are presented in column 5 of Table 3. Reassuringly, the estimated association between ϕ_{0i} and ϕ_{1i} is negative and significant at the 10 percent level (p-value of 0.075).

As a further check on application sorting, we run a regression that controls for the *lead* and *lag* effects of the shortened deadline. The lead and lag variables of the short timeframe dummy are denoted by S_{b-1t} and S_{b+1t}, respectively. The model is then restated as

$$\begin{aligned} \log(Odds_{aibt}) = \quad & \alpha_b + \alpha_t + (\phi_{0i} + \alpha_{0X} X_{0,abt} + \alpha_{0W} W_{0,ibt}) + \\ & (\phi_{1i} + \alpha_{1X} X_{1,abt} + \alpha_{1W} W_{1,ibt}) S_{bt} + \phi_{2i} S_{b-1,t} + \phi_{3i} S_{b+1,t} \end{aligned} \tag{3}$$

This specification is motivated by a scenario that we believe may occur in the patent examination process.[47] Suppose that a patent examiner reviews easy applications in short biweekly periods and difficult applications in the rest of the biweekly periods. If application sorting is allowed, he or she may review difficult patent applications in the biweekly period before or after a short biweekly period. Consequently, the task performance before or after a short biweekly period decreases because the task load is transmitted from one period to another. Such "burden-shift effects" are captured by the parameters ϕ_{2i} and ϕ_{3i} in the augmented regression

[44] Review order manipulation is not permitted, in principle, in the patent examination process. MPEP §708 stipulates that, as a general rule, "Each examiner will give priority to that application in his or her docket, whether amended or new, which has the oldest effective U.S. filing date." However, it is not guaranteed that the examiners conform to the rule in actual operations.

[45] We consider an application to be "easy" if it is technically simple and easily understandable so that patent examiners can decide its patentability without much prior search. On the other hand, if an application is "difficult," it is technically complex and takes a long time for patent examiners to review.

[46] While less than 60 percent of the total number of applications were granted patents, we believe that the limited data coverage is compensated by in-depth knowledge of the patent characteristics.

[47] In this scenario, patent examiners are assumed to know in advance what applications are in their task list (or "docket" using patent lingo) and can judge the level of their technical difficulty before the actual examination begins.

model.

The estimation result is presented in column 6 of Table 3, which shows that the marginal effects of S_{b-1t} and S_{b+1t} are not statistically significant (p-values of 0.399 and 0.613, respectively). Furthermore, even if the burden-shift effects are considered, the parameter σ_{01}, which represents the correlation between ϕ_{0i} and ϕ_{0i}, remains negative and statistically significant (p-value is 0.000).[48] Since no affirmative support for application sorting is observed, in the following analysis, we assume that the endogeneity of the short timeframe indicator is not severe.

As a final concern, one may worry that the empirical log-failure odds, $\hat{\phi}_{0i}$ and $\hat{\phi}_{1i}$, are confounded by unobserved patent examiner characteristics, such as inability or incompetency for tasks. However, the view is difficult to explain coherently with the data. One would expect a patent examiner with less task competency to issue more examination failures in a standard timeframe (i.e., a high value of ϕ_{0i}). However, at the same time, one would expect that the individual, because of inferior capability, would be even more likely to fail in a short timeframe (i.e., a high value of ϕ_{1i}). This predicted *positive* association between ϕ_{0i} and ϕ_{0i} is not compatible with the evidence of a *negative* correlation consistently found in the estimated results in Table 3.

4.4 Graphical Summary

It is worthwhile to summarize the regression results graphically. Figure 2 presents a scatterplot of the estimated empirical log-failure odds $(\hat{\phi}_{0i}, \hat{\phi}_{1i})$, which are calculated via empirical Bayes inference.[49] Visual inspection of the figure confirms that the association of the pair of empirical log-failure odds is negative. Figure 2 shows a fitting curve where the estimated parameter, $\hat{\phi}_{0i}$, is regressed on the other estimated parameter, $\hat{\phi}_{1i}$, with the first- and second-order terms. We find that the relationship is negative and convex.[50]

However, it remains unclear whether a behavioral theory can rationalize the observed pattern in the empirical log-failure odds. In the next section, we address this issue in detail.

5 A Model of Procrastination

We present a behavioral model of procrastination based on the theory of present-biased preferences. To elucidate the mechanism behind the reduced-form findings in the previous section, we focus on the effort allocation of a representative patent examiner who reviews patent applications in two weeks. To formalize time inconsistency in the decision making, we adopt a quasihyperbolic discounting model to describe continuous work time allocation (Fischer 1999, 2001; Herweg and Müller 2011).

[48] We estimate the standard deviations and correlation matrix of the parameters ϕ_{0i}, ϕ_{1i}, ϕ_{2i} and ϕ_{3i}. We denote by σ_k and $\sigma_{kk'}$ the standard deviation of parameter ϕ_{ki} and the correlation of parameterϕ_{ki} and $\phi_{k'i}$, respectively, for $k, k' = 0, 1, 2, 3$. The estimation results are presented in Table G.3 in the Appendix.

[49] There is a close link between random coefficient models and Bayesian statistical models. The random coefficients used in the multilevel regression model are analogous to random parameters modeled in Bayesian frameworks. See Snijders and Bosker (2011) for a detailed discussion and the formula used to obtain the empirical Bayesian estimates of the random intercept and coefficients.

[50] The estimated relationship is given by $\hat{\phi}_{0i} = -1.665 - 3.88\hat{\phi}_{1i} + 2.528\hat{\phi}_{1i}^2$.

5.1 Basic Model

Recall that every patent examiner must review a certain number of patent applications in two weeks.[51] Since patent application review demands laborious effort, the tasks are rarely completed in one day; therefore, the labor is allocated over a sequence of days.

The specific timeline is as follows. On the first day of a biweekly period, a patent examiner is assigned a quota of patent application reviews. Then, he or she dynamically allocates work time to perform the task. Note that under a flexible work arrangement policy, the patent examiner can work any day and any number of hours a day within the two weeks.[52] At the end of the biweekly period, the patent examiner must submit all the assigned work to the patent office.[53] If the patent office assesses the submitted task to be of satisfactory quality, the patent examiner receives a reward. The biweekly schedule is then repeated.

Let us assume that every patent examiner must complete the assigned task by the two-week deadline and is therefore given no chance to work on the task before or after the designated two-week period. Since the rules forbid patent examiners from allocating the quota across biweekly periods, the terminal time of the dynamic optimal allocation problem is in the finite time domain of two-week period.[54]

5.1.1 Work Time Allocation Decision

Suppose that a patent examiner has an assigned task to complete before deadline day D and determines work hours w_d for each day $d = 1, \cdots, D$. Assuming no work on weekends and holidays, there are $D = 10$ days available under the standard timeframe and $D = 9$ days under the short timeframe. In the following paragraphs, we set $D = 10$ unless stated otherwise. Let an increasing and concave function $u(l)$ denote the instantaneous utility function of a patent examiner who uses l hours for leisure. Working w_d hours on day d brings utility for l_d hours of leisure, which is given by $u(l_d) = u(24 - w_d)$. Let $S_D = \sum_{d=1}^{D} w_d$ denote the accumulated or total work hours on the final day D, and let $R(S_D)$ denote a reward for S_D hours of accumulated work. We assume that the reward function $R(S_D)$ is increasing and concave in S_D. An intertemporal tradeoff—a smaller-sooner reward vs. a larger-later reward—exists. Since the task deprives a patent examiner of leisure time, today's work depletes a small amount of utility. However, the hard work is compensated by a large chance of receiving a future reward.

In line with previous studies on procrastination (O'Donoghue and Rabin 1999), we use the β-δ quasi-hyperbolic discount function proposed by Laibson (1997), where the discount factor on day d is given by $\beta\delta^d$. The parameter β represents the present-bias factor—that is, the degree to which the patent examiner

[51] To be precise, every patent examiner must review patent applications and take office actions on whether they are patentable in a biweekly period. A patent examiner's quota is set by the number of counts that he or she receives when issuing first office actions or disposing of patent applications, as explained in Section 2.3. For simplicity, we assume that the number of applications that a patent examiner reviews is the same as the number of office actions he or she issues.

[52] As explained in footnote 26, a flextime schedule, the IFP, is available to patent examiners.

[53] Figure 1 shows that the examiners submit approximately half their decisions before the deadline. However, for expositional brevity, we assume that the agent submits all completed reviews on the last day of a biweekly period. This simplification does not change the central insight of the model since our focus is not on the timing of submission but on the quality of the submitted task.

[54] We acknowledge that the finite-horizon model is restrictive and that it would be more realistic to formulate the time allocation problem for a longer horizon than two weeks. It is nonetheless true that the assumption of a constant quota on which the model is based does not appear to contradict the data, as was shown by regression analysis in Section 3.3. Therefore, in favor of tractability, we proceed with the empirical analysis by adopting the finite-horizon framework.

favors the present over the future—whereas δ represents the standard discount factor. Given the form of time preferences, the total utility function of the agent on day d is given by

$$U_d = u(24 - w_d) + \beta \left\{ \sum_{k=1}^{D-d-1} \delta^k u(24 - w_{k+d}) + \delta^{D-d} R\left(\sum_{k=1}^{D} w_k \right) \right\} \tag{4}$$

We model a single patent examiner as a composite of intertemporal selves. Each self on day d decides how many hours he or she will work beyond that day until the final day D to maximize the total utility function U_d, which is given by equation (4). In the literature, two extreme kinds of different agents, in terms of future beliefs, have appeared: a naive agent who fails to predict future behavior and a sophisticated agent who recognizes his or her present bias. The behavior of naive agents is far more numerically tractable than that of sophisticated agents.[55] Hence, while it might oversimplify reality, the assumption of a naive agent is widely adopted by empirical studies analyzing time-inconsistent behavior (DellaVigna 2018). We thus follow previous studies and assume naivety in the rest of this analysis.

The decision making of a naive patent examiner proceeds in the following way. The self on day d determines a particular path of working time $\{w_d^*, w_{d+1}, \cdots, w_{D-1,D}\}$ that maximizes the total utility function U_d. However, when the next day $d+1$ arrives, the new self does not comply with the working time allocation chosen by the old self but updates it to the new time allocation path $\{w_{d+1}^*, w_{d+1}, \cdots, w_D\}$, which is optimally based on the current total utility U_{d+1} at that period. Notably, the planned work time w_{d+1} and the actual work time w_{d+1}^* do not coincide in general. We denote by $S_D^* = \sum_{d=1}^{D} w_d^*$ the corresponding total work time on the final day D.

5.1.2 Functional Form Specification

To solve the behavioral model presented above in practice, we must determine the shape of the instantaneous utility function $u(l)$ and the reward function $R(S)$. For the utility function, we assume the constant relative risk aversion (CRRA) form with γ being the relative risk aversion coefficient. To simplify the argument, we set the value at $\gamma = 1$ and consider a log utility form. Later, we discuss the robustness of the results against other values of the relative risk aversion coefficient.

For the reward function, we specify the function form considering the institutional features of the patent examination process. Let r be a fixed reward for a patent examination quota.[56] As stated in Section 2.3, a patent examiner receives a reward if the completed reviews are judged to be satisfactory. Suppose that the patent office assesses the patent examiner's task as unsuccessful if the work hours that he or she spent on the task fall short of a particular stochastic threshold, denoted by $\bar{S}$. Assume for simplicity that random variable $\bar{S}$ follows an exponential distribution with shape parameter τ. We denote the cumulative distribution function of $\bar{S}$ by $\text{Prob}(\bar{S} < S|\tau) = G_{\bar{S}}(S|\tau) = 1 - \exp\left(\frac{\log\tau}{80} S\right)$. Given the examination failure likelihood, the expected reward that a patent examiner who works S hours receives is $R(S) = r\,\text{Prob}(S > \bar{S}) = r - r\exp\left(\frac{\log\tau}{80} S\right)$.

The shape parameter τ has a straightforward interpretation. Even if a patent examiner works the standard eighty hours for two weeks, he or she may commit an examination error by chance and not be rewarded. The

[55] This characteristic has been noted by previous studies (Laibson et al. 2015; DellaVigna 2018). Indeed, under the assumption of a sophisticated patent examiner, we encounter computational fragility in the dynamic decision system, but we observe no such problem under the assumption of naivety.

[56] More specifically, the reward r is the utility value that a patent examiner perceives for the "counts" that he or she receives when issuing the first office actions for two weeks.

probability that this situation occurs $\text{Prob}(\bar{S} < 80) = \tau$. Therefore, the parameter τ represents the admissible patent examination failure rate—i.e., the *admissible failure rate*. This value captures some allowable degree of patent examination error because a certain amount of human error is inevitable.

We are now ready to characterize the work time decision of the patent examiner as a function of parameters (β, δ, τ, r). Let us denote the total utility function by $U_d(\beta, \delta, \tau, r)$. Similarly, we denote the single-day work time and the total work time that the patent examiner chooses by $w_d^*(\beta, \delta, \tau, r)$ and $S_D^*(\beta, \delta, \tau, r)$, respectively.

5.1.3 Reward Specification

We are interested in specifying the parameters (β, δ, τ, r) of the behavioral model. However, insufficient information is available to identify them all. In particular, little information about the reward r that the patent office pays each patent examiner is disclosed. Therefore, we impose an additional restriction on possible values of the reward r and assume that it is determined at the level under which a *time-consistent* patent examiner works the standard eighty hours in two weeks.[57] In Appendix A, we present a simple principal-agent model to interpret the reward in light of an incentive scheme.

Under this scenario, the patent office sets an appropriate level of reward to achieve an examination failure rate as low as the admissible failure rate τ, assuming the patent examiner behaves in a time-consistent manner. The reward chosen by the patent office, denoted by r^*, must satisfy the equation $S_D^*(1, \delta, \tau, r^*) = 80$, where the left-hand side represents the total work hours for an exponential discounting patent examiner with $\beta = 1$. Since the total work hours $S_D^*(\beta, \delta, \tau, r)$ are increasing in the reward r, the target reward r^*, if it exists, is uniquely determined.[58] To consider its dependency on the other parameters, we denote the reward by $r^*(\delta, \tau)$.

Using the notations defined above, the total work time that a patent examiner chooses is given by $S_D^*(\beta, \delta, \tau, r^*(\delta, \tau))$. For notational convenience, we collect the set of parameters (β, δ, τ) into a single vector of the structural parameters $\boldsymbol{\theta}$ and reparametrize the total work time as $S_D^*(\boldsymbol{\theta})$. In general, the total work time $S_D^*(\boldsymbol{\theta})$ is different from eighty hours, except for the case where a patent examiner is an exponential discounter with $\beta = 1$. The disagreement stems from a misbelief by the patent office about the patent examiner's present-biased preferences.

5.1.4 Task Performance

To conclude the model's description, we formulate a patent examiner's task performance as a function of the structural parameters $\boldsymbol{\theta}$. The model predicts that the rate of examination failure is $\tau \exp\left(\frac{S_D^*(\boldsymbol{\theta})}{80}\right)$. The *theoretical log-failure odds*, denoted by $LO_D(\boldsymbol{\theta})$, are therefore given by

$$LO_D(\boldsymbol{\theta}) = \log\left\{\frac{\tau \exp\left(\frac{S_D^*(\boldsymbol{\theta})}{80}\right)}{1 - \tau \exp\left(\frac{S_D^*(\boldsymbol{\theta})}{80}\right)}\right\}. \tag{5}$$

We assume that the parameter $\boldsymbol{\theta}$ of the model is exogenous and out of the patent office's and patent examiner's control. In the following, we determine the value of $\boldsymbol{\theta}$ at which the model best fits the data.

[57] We adopt this method of determining the reward following Fischer (1999).

[58] The proof is presented in Appendix A.

5.2 Calibration Exercise

In this subsection, we perform a numerical calibration. We find that the *empirical* log-failure odds in the level and difference exhibit negative correlation across examiners. Therefore, we expect the corresponding *theoretical* log-failure odds to have the same association pattern.

The calibration procedure is straightforward. We consider a representative patent examiner and calculate the optimal work load under the standard timeframe for various values of the parameters $\boldsymbol{\theta}$. Since equation (5) does not have a closed-form solution, we solve it numerically for the deadline period of $D = 10$ (which is the workdays under the standard timeframe) and calculate the log-failure odds of $LO_{10}(\boldsymbol{\theta})$. We repeat the same simulation procedure for short timeframes with a deadline period of $D = 9$ (which is the workdays under the short timeframe) and obtain the related log-failure odds $LO_9(\boldsymbol{\theta})$. We denote the difference in the log-failure odds by $DLO(\boldsymbol{\theta}) = LO_{10}(\boldsymbol{\theta}) - LO_9(\boldsymbol{\theta})$. In what follows, when mentioning the log-failure odds under the standard timeframe, $LO_{10}(\boldsymbol{\theta})$, we ignore the subscript and write it as $LO(\boldsymbol{\theta})$ to avoid notational clutter. Accordingly, we denote by $(LO(\boldsymbol{\theta}), DLO(\boldsymbol{\theta}))$ the theoretical log-failure odds in level and difference.

We explore the trajectories of the log-failure odds of a pool of hypothetical examiners who differ in the present-bias factors.[59] Denote by β_i the present-bias factor of patent examiner i. We let the individual value of β_i take values from 0.1 to 1.1 with a step size of 0.1. On the other hand, we assume that the discount factor is the same for all examiners at $\delta = 0.95$ and let the admissible failure rate take a value of $\tau = 0.05$, 0.10 or 0.15.[60] We define a vector $\boldsymbol{\theta}_i = (\beta_i, \delta, \tau)$ and denote by $LO(\boldsymbol{\theta}_i)$ and $DLO(\boldsymbol{\theta}_i)$ the theoretical log-failure odds of patent examiner i in the level and the difference, respectively.

Figure 3 presents the simulation results. The leftmost panel shows, for reference purposes, a scatterplot of the *empirical* log-failure odds, which is the same graph as Figure 2, while the panels in the second to the rightmost columns show scatterplots of the *theoretical* log-failure odds for different values of the admissible failure rate. A darker color indicates a higher degree of present-biased preference (i.e., a lower value of the present-bias factor). While the log-failure odds trajectory is perfectly inelastic for highly present-biased patent examiners, the association between the level and difference becomes negative as a patent examiner's present-biased preferences become mild. Thus, we can conclude that the theoretical log-failure odds can accommodate the observed negative correlation of the empirical log-failure odds when individual present-bias factors are heterogeneous and distributed in the middle range.

6 Structural Model Estimation

We proceed to estimate the patent examiner-specific present-bias factor. This section consists of several parts. The first part is devoted to derivation of the likelihood function for the key parameters. The second part briefly explains a Bayesian estimation method. In the third part, the baseline results are presented, and a sensitivity analysis is performed. The final part shows the out-of-sample model validation.

[59] Appendix B provides simulated work patterns of patent examiners with distinct values of present-bias factor β. We find that the outcomes are generally consistent with the findings on time allocation in previous theoretical papers (Herweg and Müller 2011).

[60] We choose the values for the admissible failure rate of τ to cover the actual failure rate of 0.15. See the prior distribution settings of the parameter described in the next section.

6.1 Likelihood Function

We start by deriving the likelihood function. We substitute the *empirical* log-failure odds, (ϕ_{0i}, ϕ_{1i}) for the *theoretical* log-failure odds $(LO(\boldsymbol{\theta}_i), DLO(\boldsymbol{\theta}_i))$ in the log-failure odds given by equation (2).

Accordingly, the log-likelihood that examination failure occurs takes the following form:

$$\begin{aligned}\ell(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \quad & \sum_{i \in N} \sum_{t \in T_i} \sum_{b \in B_t} \sum_{a \in A_{ibt}} \log \Lambda \Big\{ \alpha_b + \alpha_t + LO(\boldsymbol{\theta}_i) + \alpha_{0X} X_{0,abt} + \alpha_{0W} W_{0,ibt} + \\ & DLO(\boldsymbol{\theta}_i) S_{bt} + (\alpha_{1X} X_{1,abt} + \alpha_{1W} W_{1,ibt}) S_{bt} \Big\}\end{aligned}$$

where Λ represents the cumulative distribution function of the logistic distribution, N represents the set of patent examiners, T_i represents the set of years in which patent examiner i is in the patent office, B_t represents the set of biweekly periods in year t, and A_{ibt} represents the set of applications reviewed by patent examiner i in biweekly period b in year t. We abuse the notation and denote by N the number of patent examiners.

For ease of exposition, we consolidate the main parameters into a vector $\boldsymbol{\theta} = (\boldsymbol{\beta}, \delta, \tau)$ and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_N)$. We also include the remaining "incidental" parameters in a vector $\boldsymbol{\alpha} = (\alpha_b, \alpha_t, \alpha_{0X}, \alpha_{0W}, \alpha_{1X}, \alpha_{1W})$.

6.2 Estimation Method

There are two computational aspects to consider. First, no analytical form is available for the theoretical log-failure odds $LO(\boldsymbol{\theta}_i)$ and $DLO(\boldsymbol{\theta}_i)$ or the log-likelihood $\ell(\boldsymbol{\theta}, \boldsymbol{\alpha})$. Therefore, a simulation-and-estimation procedure is required to obtain the maximum-likelihood estimates of the parameter $\boldsymbol{\theta}$. Since the simulated likelihood is not smooth in the parameter, derivative-based methods, such as quasi-Newton methods, are not suitable for maximizing the likelihood function. Second, the log-likelihood function is of high dimensionality. Indeed, since there are approximately seven hundred patent examiners, we must specify as many individual-specific present-bias factors $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_N)$. Furthermore, since the model incorporates hundreds of explanatory variables, including time trend dummies and technological group dummies, the total number of incidental parameters is substantial in the full specification. Such high dimensionality of the parameter space further increases the computational cost because non-derivative-based optimization algorithms are inefficient and likely to become stuck in local optima.

This paper adopts a Bayesian approach to circumvent the computational challenges. In the structural model context, particularly when simulation-estimation exercise is called for, previous studies have noted several advantages of Bayesian methods over classical estimation methods.[61]

According to Bayes' rule, the marginal joint posterior for the key parameter $\boldsymbol{\theta}$ is given by

$$\pi(\boldsymbol{\theta}) = \frac{\int \exp \ell(\boldsymbol{\theta}, \boldsymbol{\alpha}) \pi_0(\boldsymbol{\theta}, \boldsymbol{\alpha}) d\boldsymbol{\alpha}}{\int \int \exp \ell(\boldsymbol{\theta}, \boldsymbol{\alpha}) \pi_0(\boldsymbol{\theta}, \boldsymbol{\alpha}) d\boldsymbol{\theta} d\boldsymbol{\alpha}} \tag{6}$$

where $\pi_0(\boldsymbol{\theta}, \boldsymbol{\alpha})$ represents the prior distribution. Since the multiple integrals in the equation above cannot be

[61]The following three points are considered advantages of Bayesian methods over classical methods. First, since Bayesian methods rely on integration instead of optimization, inference works well even when the likelihood function is neither smooth nor unimodal. Second, the posterior integration is less computationally expensive than maximizing the likelihood function. Third, although the Bayesian estimator has the same asymptotic behavior as the classical estimator, it can have desirable estimation properties, such as consistency and efficiency, under less restrictive conditions. See Train (2009) for more details on this issue. Jiang et al. (2009) argue the advantage of Bayesian methods with MCMC over standard GMM estimation methods for random-coefficient logit models.

evaluated analytically, we use MCMC algorithms for the integrals. However, if the parameter space is of high dimensions, it is still unwieldy for MCMC algorithms to compute the posterior distribution with sufficient accuracy. For high-dimensional models, Markov chains do not mix sufficiently fast over the parameter space (Jackman 2009). If we let an MCMC algorithm run for a sufficiently long time, it will eventually reach an equilibrium distribution; however, the computational cost will be prohibitively high.

To lessen the computational burden and make the Bayesian inference practically tractable, we adopt a pragmatic approach to reduce the dimensions of the parameter space. We concentrate the posterior distribution for the key parameter $\boldsymbol{\theta}$ by conditioning it on the incidental parameter $\hat{\boldsymbol{\alpha}}$:

$$\pi(\boldsymbol{\theta}|\hat{\boldsymbol{\alpha}}) = \frac{\exp\left\{\ell(\boldsymbol{\theta},\hat{\boldsymbol{\alpha}})\right\}\pi_0(\boldsymbol{\theta})}{\int \exp\left\{\ell(\boldsymbol{\theta},\hat{\boldsymbol{\alpha}})\right\}\pi_0(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{7}$$

where $\hat{\boldsymbol{\alpha}}$ represents the estimates of the incidental parameters obtained from the reduced-form estimation in Section 4, and $\pi_0(\boldsymbol{\theta})$ represents the prior distribution.

We acknowledge that the concentrated posterior distribution (7) may misspecify the true marginal posterior distribution (6). However, the latter is also considered an approximation of the complex reality concerning the patent examination.[62] Therefore, our approach is considered a constructive way to reach reasonable conclusions about the parameters of interest.

We assign uniform distributions for the prior $\pi_0(\boldsymbol{\theta})$ with sufficiently wide supports to cover most of the estimates in previous studies. The specific choices are as follows. First, the present-bias parameter $\boldsymbol{\beta}$ is assumed to span the interval $[0.1, 1.1]$. We allow the value to exceed 1.0 to consider the possibility of future bias (Takeuchi 2011). Second, the annual discount factor parameter δ is in the interval $[0.9, 1.0]$. Note that since the model assumes daily decision making, we convert the annual discount factor to a daily discount factor. Third, the parameter of the admissible failure rate, τ, is assumed to lie in the range $[0.01, 0.20]$. Since it is interpreted as an unavoidable chance of examination failure for patent examiners, we choose a sufficiently broad interval to cover the actual failure rate of 0.15. We impose a prior independence assumption such that the priors' distributions are assumed independent for all the parameters.

Table 4 reports the Bayesian statistical results of the concentrated posterior distribution $\pi(\boldsymbol{\theta}|\hat{\boldsymbol{\alpha}})$ given by equation (7). Computations are performed using an MCMC method based on the Metropolis-Hastings algorithm. We present the detailed procedure in Appendix C. Panel A of Table 4 summarizes the estimates for the key parameters. The point estimates are the means of the marginal posterior distributions. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. The quantile points of the distribution of the mean individual present-bias factors are reported. Panel B of Table 4 provides the goodness-of-fit measures. The widely applicable information criterion (WAIC) is an information criterion similar to the AIC and BIC but is based on Bayesian statistics (Watanabe 2013; Gelman et al. 2014). As previously stated, AUC.ROC provides a measure of the models' discriminatory accuracy. Panel C reports the percentage of present-biased patent examiners. We label a patent examiner as present biased if the upper limit of the 95 percent HDI of his or her present-bias factor does not reach the value of one.

Several features emerge from the baseline estimates shown in column 1. First, and most important, is that the majority of patent examiners have present-biased preferences: the median patent examiner has a

[62] For example, the regression model assumes that the patent examiner and application-specific characteristics affect the examination outcome in a linear and additive way. Furthermore, the distribution of the error term is assumed logistic. These assumptions are standard but have never been tested.

present-bias factor of 0.60. Furthermore, the percentage of present-biased patent examiners is approximately 70 percent. The second feature of the estimates is that the discount factor δ is imprecisely estimated with an extensive 95 percent HDI. The posterior distribution is as flat as the prior distribution, implying that the data are not sufficiently informative to be conclusive about the discount factor.[63] Last, the posterior uncertainty of the admissible failure rate of τ is reasonably small. The probability that the true value lies in an interval with a width of 0.005 centered at the estimated mean of 0.044 is higher than 95 percent. Since the parameter is construed as the "target" rate of examination failure, it seems reasonable that the estimated value is below the actual examination failure rate of 0.15.

Figure 4 presents the marginal posterior distributions of the present-bias factor for patent examiners. The vertical line segment corresponds to the 95 percent HDI of individual patent examiners, where patent examiners are sorted in ascending order according to the mean represented by a black line. We use a darker color for present-biased patent examiners and a lighter color for other patent examiners. While present-bias factors are broadly distributed across patent examiners, a substantial fraction of patent examiners are present biased.

6.3 Robustness Checks

We begin by examining the sensitivity of the results to prior assumptions. Although it is possible to formulate numerous priors, we focus on changes in the prior of the discount factor δ since it is imprecisely estimated under the baseline setup. Column 2 of Table 4 presents the estimation results when we use an interval $[0.1, 1.0]$ as the support of the discount factor prior. Although the point estimate of the discount factor δ drops significantly, the credible interval remains wide, indicating that the estimated value is not accurate. More importantly, the estimated posterior distributions of the present-bias factors and the admissible failure rate are similar to those in the baseline. Similar results are obtained when the supports of the prior distributions of the present-bias factor and the admissible failure rate are widened. Table G.4 in the Appendix reports the estimation results with the support of the uniform prior of the parameter $\boldsymbol{\beta}$ set to $[0.1, 1.2]$ or that of the parameter τ set to $[0.01, 0.30]$.

Next, we perform a robustness check by changing the CRRA utility function, where the default value of the relative risk aversion coefficient is set to $\gamma = 1.0$, (i.e., the log utility function). Columns 3-4 of Table 4 report the results for changing values of γ. Encouragingly, the median values of the estimated mean of the posterior distributions do not change substantially. Table G.5 in the Appendix reports the estimation results for a wide range of γ values from 0.0 to 6.0. The percentage of present-biased patent examiners tends to decline as the curvature degree increases; however, the goodness-of-fit measure also decreases. Therefore, in terms of model fit, the default choice of $\gamma = 1.0$ appears reasonable.

To further confirm the robustness of the baseline results, we explore different assumptions concerning the reward setup. In Section 5.1, we assume that the patent office sets the reward at a level where the average examination failure rate is τ for a *time-consistent* patent examiner. What if the reward is targeted to a *time-inconsistent* patent examiner? Columns 5-6 of Table 4 report the results under the assumption that the reward is determined for patent examiners whose present-bias factor is $\beta = 0.4$ or 0.8. Table G.6 in the Appendix shows more detailed results, where the parameter range is extended to between 0.1 and 0.9. In all cases, the estimated posterior distribution of the present-bias factors does not change significantly from the

[63]This is evident from the marginal posterior distribution of δ presented in Figure C.3 in Appendix C.

baseline. Furthermore, the goodness-of-fit measures of the alternative specifications are no better than the benchmark specification.

Finally, we examine the heterogeneity assumption on time discounting. Throughout the analysis, patent examiners are assumed to have the same discount factor, while the present-bias factors are individual specific. What if patent examiners are exponential discounters with individual-specific discount factors? Should the task performance variation across patent examiners be attributed to discount factor differences rather than present-bias factor differences? To assess this possibility, we estimate an *exponential discounting* model with $\beta = 1$, allowing the discount factor to be individual specific, δ_i, for patent examiner i.[64] We then modify the assumptions concerning the reward scheme. Specifically, let us assume that the patent office has a *common* belief about the patent examiner's discounting factor, which we denote by ζ. Since the patent office's belief is not necessarily correct, the belief parameter $\zeta \neq \delta_i$ for some patent examiner i.

We estimate the posterior distributions of the key parameters $(\boldsymbol{\delta}, \zeta, \tau)$, where $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$ is a vector of individual discount factors. Table 5 reports the results. Column 1 provides a summary of the estimates when assuming uniform priors on $[0.1, 1.0]$ for the discount-factor-related parameters $\boldsymbol{\delta}$ and ζ and a uniform prior on $[0.01, 0.20]$ for the admissible failure rate τ. Since the discount factors are individual specific, the quantile points of the distribution of the mean individual discount factors are reported.

Overall, the results direct us not to use the exponential discounting model. There are at least three reasons for this. First, the posterior of the individual-specific discount factor is neither precisely estimated nor sufficiently discriminative to separate patent examiners with higher discounting factors from those with lower discounting factors. [65] Second, the median of the mean marginal posterior distribution of the discount factor is 0.55 in the annual term and is thus implausibly low. It should be noted, however, that the result depends on the prior assumption. Column 2 shows the estimation results where the priors of the discount-factor-related parameters, $\boldsymbol{\delta}$ and ζ, are uniform on $[0.9, 1.0]$. While the estimated range of the discount factors appears plausible in this case, it is considered an artifact of the restrictive priors. Finally, notwithstanding the priors, the goodness-of-fit values, measured in terms of the WAIC and AUC.ROC, are substantially lower than those for the hyperbolic discounting model.

6.4 Model Verification

In this subsection, we attempt to validate our behavioral model based on *out-of-sample* prediction. We adopt an out-of-sample validation approach in line with previous studies (e.g., Fehr and Goette 2007).[66] The idea is straightforward. We examine whether our estimate of present-bias factors will accurately forecast patent examination delays.

The model prediction is based on a sample not used for estimation. While the present-bias factors are estimated from the data on patent examination *performance* in the initial review process, the predictability of the model on which we focus is assessed by the information on patent examination *delay* in the entire review

[64] To deepen our understanding of the estimated results, we perform a calibration exercise under the assumption that all patent examiners are exponential discounters with heterogeneous discount factors. The results are presented in Appendix D. As demonstrated, the model prediction fits the empirical data very poorly. The statistical inference performed in this section corroborates the qualitative findings of the calibration results.

[65] The inaccuracy of the posterior estimates is also evident from Figure G.2 in the Appendix, where the 95 percent HDI of the marginal posterior distribution of the individual discount factor, $\bar{\hat{\delta}}_i$, is presented for each patent examiner i.

[66] See DellaVigna (2018) for other out-of-sample validation exercises in behavioral economics.

process. Since the "double use" of data for estimation and prediction is avoided, our validation approach can be considered out-of-sample rather than within-sample.

To garner information on whether the examination of an issued patent is delayed, we focus on the lengthened patent term, referred to as *patent term adjustment (PTA).* Under U.S. patent law, a patent is eligible for day-for-day adjustment of the standard twenty-year patent term, and the length of the extension is based on delays caused by the fault of the patent office. Among several reasons for which a patent term may be extended, one is that the patent examiners fail to meet specified timeframes.[67] Therefore, we employ the PTA period as a reasonable proxy for delayed patent examination.

We regress the PTA period of an issued patent on the present-bias measure of the patent examiner who reviewed the patent.[68] We consider a negative value of the standardized estimated present-bias factor of patent examiner i and refer to it as $PresentBiasMeasure_i$.[69] A higher value indicates that patent examiner i has a lower present-bias factor β_i.

The regression requires special considerations because the distributional pattern of the PTA period is limited in a way that the support of the distribution is a set of nonnegative integers characterized by a massive spike at zero.[70] While we could consider several regression models that accommodate such distributional features, we adopt tobit regression because it exhibits a better fit to the data than alternative specifications.[71]

The tobit regression model can be stated as

$$\begin{aligned} &PTA^*_{ait} = \alpha_t + \alpha_X X_{at} + \alpha_W W_{it} + \rho PresentBiasMasure_i + \varepsilon_{ait} \\ &PTA_{ait} = PTA^*_{ait} \text{ if } PTA^*_{ait} \geq 0, \text{and } PTA_{ait} = 0 \text{if } PTA^*_{ait} < 0, \end{aligned} \tag{8}$$

where a indexes the patent application, i indexes the patent examiner, and t indexes the year. The left-hand side variable, PTA_{ait}, is the actual days of the extended patent term associated with the "latent" variable PTA^*_{ait}. The primary explanatory variable is the present-bias measure $PresentBias_i$. As additional explanatory variables, we include the year dummies α_t and the application and patent examiner characteristics variables, (X_{at}, W_{it}), which are the same as those used in the regression analysis in Section 4.2. The unobserved error term ε_{ait} is assumed to be normally distributed.

Table 6 reports the main results from the tobit regression.[72] We focus on the estimated marginal effect of the present-bias measure on the PTA period. The baseline results in column 1 show that the estimated marginal effect is statistically significant (p-value of 0.000), which implies that an application reviewed by a patent examiner classified as more heavily present-focused tends to be subject to longer days of PTA.

The specification in column 2 includes a dummy variable, $Failure_{ait}$, that represents whether the first office action of patent application a reviewed by patent examiner i is judged unsuccessful. The estimated coefficient is positive and statistically significant (p-value of 0.000). This result is not surprising because a first office action failure calls for additional examination rounds (i.e., the second office action is nonfinal) and

[67]PTA due to patent office delay is provided under 35 U.S.C. §154 (b) 1.

[68]The sample consists of applications that were granted patents because unapproved patent applications do not have PTA records.

[69]To be more precise, we define$PresentBias_i \equiv -\left\{\bar{\hat{\beta}}_i - \text{Avg}(\bar{\hat{\beta}}_i)\right\}/\text{SD}(\bar{\hat{\beta}}_i)$, where the mean of the marginal posterior distribution of the present-bias factor $\bar{\hat{\beta}}_i$ is taken from the baseline estimates reported in column 1 of Table 4.

[70]The histogram of the PTA period is presented in Figure G.3 in the Appendix.

[71]We use count regression models, including Poisson, negative binomial, zero-inflated Poison, and zero-inflated negative binomial models. The specifications and estimation results are presented in Appendix E.

[72]Table G.7 in the Appendix reports the marginal effects of the variables that are not presented in Table 6.

is thus likely to extend the patent adjustment term. More importantly, the estimated marginal effect of the present-bias measure on the PTA period remains positive and statistically significant (p-value of 0.000).

To isolate the effect of the present-bias factor from potential confounding factors, we estimate a random-effect tobit model that incorporates a normally distributed patent examiner-specific error term. Column 3 indicates that even after patent examiner-specific unobserved heterogeneity is controlled for, the present-bias measure correctly predicts delayed examination.

Finally, in column 4, we re-estimate the tobit regression for patent applications examined from 2010 to 2018. Note that the present-bias factors were estimated using a sample of patent applications reviewed between 2001 and 2009. The use of samples from different periods for estimation and prediction allows for a more robust out-of-sample validation. We find that the patterns of estimates are quite similar to those depicted above.[73]

All the regression results presented above demonstrate that our present-bias measure exhibits good out-of-sample predictive performance for patent examination delay. Therefore, we conclude that the results support the validity of the model and the assumptions embedded in the model.

7 Policy Implications

This section brings behavioral insights into policy issues concerning the patent examination process. We explore the topic from two angles: (i) the attrition of patent examiners and (ii) the structure of the patent examination quota.

7.1 Patent Examiner Attrition

Public administration experts have expressed broad concern regarding the high attrition rate of patent examiners in the U.S. (NAPA 2005). For example, in the early 2000s, approximately 70 percent of patent examiners left the patent office within five years of joining, causing substantial loss of training investment and human capital. While the patent agency had made several efforts to hire and retain qualified patent examiners, they were not considered sufficient (GAO 2007).

In this context, we investigate whether behavioral factors could affect the job separation of patent examiners. We focus on the subgroup of patent examiners who entered the patent office as junior examiners between 2001 and 2009[74]. We further categorize those who left the position within five years after the first date they issued their first office action as the *leaver* group and the remaining as the *stayer* group.[75]

Figure 5 compares histograms of the present-bias measure defined in the previous section between the *leaver* and *stayer* groups. The two distributions are noticeably different. The distributions differ significantly according to the Mann-Whitney test (p-value of 0.039), and the distribution of the *stayer* group first-order stochastically dominates the distribution of the *leaver* group according to Somers' D statistic (p-value of 0.031).

[73]We test the predictive performance of the present-bias measures for patent examiners included in the pre-2010 sample and the post-2010 sample, which is approximately 60 percent of the original sample.

[74]The target group of patent examiners consists of 38 percent of all patent examiners in the original sample

[75]We define a patent examiner as having left the patent office when his or her patent examination record has been disrupted for more than a year. Therefore, it cannot be ruled out that a patent examiner is mistakenly judged as a *leaver* when promoted from a primary examiner position (GS-14) to a more senior and managerial position, such as supervisory patent examiner or technology center director (GS-15 or above). While such possibilities are likely rare, it should be considered that our attrition data contain some measurement errors.

Therefore, patent examiners with *less* present-biased preferences are *more* likely to leave the position.[76]

This finding enables us to better understand the implications on the efficiency of the patent examination process. The intense turnover at the U.S. patent agency was once described as "churning of so many new patent examiners" (GAO 2007) . A more suitable phrase suggested by the estimation results above would be *cream skimming* of new patent examiners: the less present-biased patent examiners who perform high-quality reviews are more likely to leave, while more present-biased patent examiners who tend to delay reviews are more likely to remain.

Based on these analyses, we suggest that the patent office adopt a strategy to mitigate the higher job separation of the less present-biased patent examiners. The Patents Hoteling Program (PHP) introduced in the mid-2000s was one effort taken by the USPTO to increase the patent examiner retention rate by providing greater work flexibility (NAPA 2005, 2015). Nevertheless, the telework environment, which limits the patent office's ability to monitor the patent examiners, might encourage the heavily present-biased patent examiners to procrastinate and therefore would result in a counterproductive effect. Therefore, it is recommended that the telework program be made available only to patent examiners with less present-biased preferences. Our elicited measure of present-biased preference may help separate the wheat from the chaff.

7.2 Examination Quota

What work quota would best enhance patent examiner task performance? Since present-biased patent examiners are prone to squander the allotted time, shortening the deadline would improve their time management and consequently increase their task performance. We verify this intuition by means of a counterfactual simulation.

Suppose that the patent office cuts the patent examination quota in half from the original amount and shortens the deadline from 10 days (the number of workdays in a biweekly period) to 5 days (the number of workdays in a weekly period). To measure the impact of this potential quota reform, we simulate a model in which the time discounting function of examiner i is based on $(\bar{\bar{\beta}}_i, \bar{\bar{\delta}})$, which are the mean values of the marginal posterior distributions estimated by the baseline model reported in Table 4. Furthermore, the reward under the hypothetical quota rule is set to the level under which the failure rate of a time-consistent patent examiner who works forty hours a week is equal to $\bar{\bar{\tau}}$, which is the mean value of the estimated posterior distribution of the admissible failure rate.

Table 7 reports the differences in patent examination accuracy and patent pendency between the current and hypothetical quota rules for the whole sample (row 1) and for the bottom and top quintile groups based on the present-bias measure (rows 2-3). Column 1 shows the predicted average reduction in the first office action failure rate. In column 2, the predicted decrease in the PTA period is presented.[77] The values in the brackets are the percentage changes relative to the original magnitudes. The fragmented task quota substantially improves patent examination accuracy and timeliness, with a more significant impact for patent examiners with a higher degree of present bias.

[76] In Appendix F, we perform a survival analysis of patent examiners' job duration using a Cox proportional hazard regression. As expected, we find less present-biased patent examiners have a higher hazard of leaving the patent office.

[77] The quota structure alters the PTA length through a change in the first office action failure probability. Specifically, if first office action failure becomes less likely, the number of office actions required before a patent is issued decreases. The effect of the quota change on the first office action failure is reported in column 1 of Table 7, and the effect of the first office action failure on the PTA period is reported in column 2 of Table 6 in the previous section. The synthesis of the two effects yields the predicted decrease in PTA, as shown in column 2 of Table 7.

How large is the monetary benefit gained from the quota reform? A back-of-the-envelope calculation can approximate the amount. Considering that we have analyzed the patent prosecution process in the fields of Biotechnology and Organic Chemistry, let us take the case of a hypothetical blockbuster drug with $1 billion in annual sales in the U.S.[78] If a pharmaceutical patent protects the brand-name drug, no one except the patentee can manufacture and sell the drug, except in licensed production. However, once the patent expires, low-price bioequivalent generic drugs can be produced and will flood the market. Therefore, the drug's original manufacturer could earn more profit for the extended patent term. The situation would be changed, however, if it were given a shorter PTA period.

Based on the results presented in Table 7, the shorter deadline reduces PTA by 6.62 days, on average, allowing the entry of a generic drug as early as that number of days. If 90 percent of consumers switch from the brand-name drug to a generic drug, the patent holder's sales loss caused by the entry of a generic substitute would amount to approximately $ 16.3 million.[79][80] If the generic drug is sold for 75 percent less than the brand-name drug, the potential consumer gains are estimated at $ 12.2 million.[81]

We can obtain more detailed estimates of consumer benefits for a specific brand drug. Let us take SPRYCEL, a blockbuster cancer drug with approximately $1.2 billion in annual sales in the U.S. market.[82] The drug is protected by patents held by Bristol-Myers Squibb, one of which is due to expire in 2020. Among the patents, the one with the longest PTA is U.S. Patent 7,491,725, whose patent term has been extended by 417 days.[83]

Table 8 presents the results of counterfactual experiments for the patent under the following three scenarios: (i) The original patent examiner reviews under the hypothetical one-week quota rule. (ii) A hypothetical exponential discounting patent examiner reviews under the original two-week quota rule. (iii) A hypothetical discounting patent examiner reviews under the hypothetical one-week quota rule. The quota reforms result in a reduction in PTAs and can create enormous consumer benefits.[84]

8 Conclusion

The majority of people procrastinate: our analysis shows that patent examiners are no exception. We fit a quasihyperbolic discounting model with patent examiners' task performance data and find that present-biased preferences are widespread among patent examiners. The result implies that present-focused patent examiners may have a consequential negative impact on patent examination quality because they may review patent applications incompletely in situations wherein there are time constraints. We demonstrate that overall patent examination quality may deteriorate due to the attrition of less present-biased patent examiners. We show by simulation that patent quality and pendency may be improved by reducing the quota and shortening the deadline.

[78] The definition of a "blockbuster" drug follows from Cutler (2007). Munos (2009) showed that approximately 20 percent of new drugs approved between 1950 and 2008 in the U.S. achieved blockbuster status expressed in year-2000 equivalent dollars.

[79] According to GAO (2012), the rate at which a generic drug is substituted for a brand-name drug is 93 percent.

[80] The estimated amount is based on the following formula: 1000 million dollars/(365days) × (6.62days) × 0.90.

[81] The information on generic drug prices comes from a report published by the U.S. Congressional Budget Office (CBO) .

[82] The sales figure is for 2019, as presented in the annual (10-K) report of Bristol-Myers Squibb.

[83] Lajeunesse, DiMarco, Galella and Chidambaram (2009) "Process for Preparing 2-Aminothiazole-5-Aromatic Carboxamides as Kinase Inhibitors," U.S. Patent No. 7,491,725.

[84] The calculation is based on the generic-to-brand substitution and price ratios shown above.

Our paper has several limitations. First, some model assumptions are not directly verified. They may be overly restrictive or even false to describe complex decision making in the real-work environment of patent examiners. For example, one may consider the dynamic setting of our behavioral model unjustified. We formulate a patent examiner's intertemporal worktime allocation over two weeks. However, the actual patent examiners may be more far-sighted and allocate their work effort considering future events that would occur beyond the biweekly timeline. To relax the restriction regarding the time horizon, one may need a more flexible but more complicated dynamic programming model. However, this is not the goal we pursue here. We leave the development of a more situationally relevant model to future research.

Second, our research focus is on present-biased preferences in isolation to understand procrastination behavior. However, a recent experimental study suggests that procrastination can be associated with limited memory or forgetfulness (Taubinsky and Rees-Jones 2018). Moreover, if such factors play roles in intertemporal decision making, the present-biased factors may be estimated with a bias (Ericson 2017). Therefore, the sheer size of our estimates on the patent examiner-specific present-biased factors should be interpreted with caution.

Third, we find a somewhat puzzling behavioral pattern based on our procrastination model. Notably, the reduced-form regression results in Section 4.2 indicate that some, though not many, patent examiners *improve* their task performance when the quota period is shortened. As the second quadrant of Figure 2 illustrates, the log-odds differences are negative for those patent examiners, suggesting that the odds that the initial examination fails may decline in a tight-time situation. The estimates are not consistent with our model, where constant or increasing failure odds are predicted. Additionally, it remains an open question why the attrition rate is significantly higher for less present-biased patent examiners than more present-biased patent examiners. Further research should be pursued to better understand these behavioral patterns.

Finally, our conclusions on unproductive procrastination are drawn from the analysis of a subset of patent examiners. In this regard, they may not be generalizable to the broad population of patent examiners. We opt to focus on patent examiners in TC 1600 because the requirement of computational resources is prohibitively high if the scope of the analysis is expanded to the whole set of U.S. patent examiners. While supplemental analysis using the data for patent examiners in other technological fields yields qualitatively similar results, more extensive research should be conducted to understand the whole picture of unproductive procrastination of U.S. patent examiners.[85]

Notwithstanding these limitations, this paper provides firsthand measurable evidence of patent examiners' present-bias-induced procrastination, which helps us understand whether or to what extent present-biased preferences hinder efficient patent prosecution and make policy suggestions on how to redesign the system.

[85]We reproduce the analysis using data of patent examiners in TC 2100: Computer Architecture and Software. The main estimation results of the present-biased factors are presented in Table G.8 in the Appendix. Although the results vary with the value of the utility curvature parameter, the conclusion that a majority of patent examiners have present-biased preferences remains unchanged.

References

Akerlof, George A, "Procrastination and Obedience," *The American Economic Review*, 1991, *81* (2), 1–19.

Ariely, Dan and Klaus Wertenbroch, "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological science*, 2002, *13* (3), 219–224.

Augenblick, Ned and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," *Review of Economic Studies*, 2019, *86* (3), 941–975.

_ , Muriel Niederle, and Charles Sprenger, "Working Over Time: Dynamic Inconsistency in Real Effort Tasks," *The Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Bisin, Alberto and Kyle Hyndman, "Present-bias, Procrastination and Deadlines in a Field Experiment," *Games and economic behavior*, 2020, *119*, 339–357.

Brachet, Tanguy, Guy David, and Andrea M Drechsler, "The Effect of Shift Structure on Performance," *American Economic Journal: Applied Economics*, 2012, *4* (2), 219–46.

Chan, David C, "The Efficiency of Slacking Off: Evidence from the Emergency Department," *Econometrica*, 2018, *86* (3), 997–1030.

Chan, Marc K, "Welfare Dependence and Self-Control: An Empirical Analysis," *The Review of Economic Studies*, 2017, *84* (4), 1379–1423.

Chetty, Raj, Emmanuel Saez, and László Sándor, "What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics," *Journal of Economic Perspectives*, 2014, *28* (3), 169–88.

Cockburn, Iain M, Samuel Kortum, and Scott Stern, "Are All Patent Examiners Equal? The Impact of Examiner Characteristics," Technical Report, NBER Working Paper No. 8980 2002.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White, "Measuring Time Preferences," *Journal of Economic Literature*, 2020, *58* (2), 299–347.

Coviello, Decio, Andrea Ichino, and Nicola Persico, "The Inefficiency of Worker Time Use," *Journal of the European Economic Association*, 2015, *13* (5), 906–947.

Cutler, David M, "The Demise of the Blockbuster?," *New England Journal of Medicine*, 2007, *356* (13), 1292–1293.

DellaVigna, Stefano, "Structural Behavioral Economics," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 1, Elsevier, 2018, pp. 613–723.

Dixit, Avinash K and Robert S Pindyck, *Investment under Uncertainty*, Princeton university press, 1994.

Ericson, Keith Marzilli, "On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation," *Journal of the European Economic Association*, 2017, *15* (3), 692–719.

Fang, Hanming and Dan Silverman, "Time-Inconsistency and Welfare Program Participation: Evidence from the NLSY," *International Economic Review*, 2009, *50* (4), 1043–1077.

— **and Yang Wang**, "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions," *International Economic Review*, 2015, *56* (2), 565–596.

Fehr, Ernst and Lorenz Goette, "Do Workers Work More If Wages Are High? Evidence from a Randomized Field Experiment," *The American Economic Review*, 2007, *97* (1), 298–317.

Fischer, Carolyn, "Read This Paper Even Later: Procrastination with Time-Inconsistent Preferences," 1999. mimeo.

— , "Read This Paper Later: Procrastination with Time-Consistent Preferences," *Journal of Economic Behavior & Organization*, 2001, *46* (3), 249–269.

Frakes, Michael D and Melissa F Wasserman, "Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data," *Review of Economics and Statistics*, 2017, *99* (3), 550–563.

— **and** — , "Procrastination at the Patent Office?," *Journal of Public Economics*, 2020, *183*, 104140.

Frederick, Shane, George Loewenstein, and Ted O'donoghue, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 2002, *40* (2), 351–401.

Gelman, Andrew and Jennifer Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, 2006.

— **, Jessica Hwang, and Aki Vehtari**, "Understanding Predictive Information Criteria for Bayesian Models," *Statistics and Computing*, 2014, *24* (6), 997–1016.

Graham, Stuart JH, Alan C Marco, and Richard Miller, "The USPTO Patent Examination Research Dataset: A Window on the Process of Patent Examination," 2015. mimeo, USPTO Economic Working Paper No. 2015-4.

Heidhues, Paul and Philipp Strack, "Identifying Present-Bias from the Timing of Choices," 2019. mimeo.

Herweg, Fabian and Daniel Müller, "Performance of Procrastinators: on the Value of Deadlines," *Theory and Decision*, 2011, *70* (3), 329–366.

Hosmer, David W and Stanley Lemeshow, *Applied Logistic Regression*, 2 ed., Wiley-Interscience, 2000.

Jackman, Simon, *Bayesian Analysis for the Social Sciences*, John Wiley & Sons, 2009.

Jaffe, Adam B and Josh Lerner, *Innovation and Its Discontents: How Our Broken Patent System Is Endangering Innovation and Progress, and What to Do About It*, Princeton, NJ: Princeton University Press, 2004.

Jiang, Renna, Puneet Manchanda, and Peter E Rossi, "Bayesian Analysis of Random Coefficient Logit Models Using Aggregate Data," *Journal of Econometrics*, 2009, *149* (2), 136–148.

Joint Hearing before the Committee on the Judiciary and Committee on Oversight and Government Reform, "Abuse of USPTO's Telework Program: Ensuring Oversight, Accountability and Quality," Technical Report, House of Representatives (113-117 and 113-150), 113th Cong. 2014.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–478.

_ , Peter Maxted, Andrea Repetto, and Jeremy Tobacman, "Estimating Discount Functions with Consumption Choices over the Lifecycle," 2015. mimeo.

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

_ and Carl Shapiro, "Probabilistic Patents," *Journal of Economic Perspectives*, 2005, *19* (2), 75–98.

Marco, Alan C, Andrew A Toole, Richard Miller, and Jesse Frumkin, "USPTO Patent Prosecution and Examiner Performance Appraisal," 2017. mimeo, USPTO Economic Working Paper 2017-08.

Martinez, Seung-Keun, Stephan Meier, and Charles Sprenger, "Procrastination in the Field: Evidence from Tax Filing," 2017. mimeo.

Mitra-Kahn, B, A Marco, M Carley, P D'Agostino, P Evans, C Frey, and N Sultan, "Patent Backlogs, Inventories and Pendency: An International Framework," 2013. mimeo, A joint UK Intellectual Property Office and US Patent and Trademark Office Report.

Munos, Bernard, "Lessons from 60 Years of Pharmaceutical Innovation," *Nature Reviews Drug Discovery*, 2009, *8* (12), 959–968.

National Academy of Public Administration (NAPA), "US Patent and Trademark Office: Transforming to Meet the Challenges of the 21st Century," Technical Report 2005.

_ , "The United States Patent and Trademark Office: A Telework Internal Control and Program Review," Technical Report 2015.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *The American Economic Review*, 1999, *89* (1), 103–124.

Paserman, M Daniele, "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," *The Economic Journal*, 2008, *118* (531), 1418–1452.

Rozental, Alexander and Per Carlbring, "Understanding and Treating Procrastination: A Review of a Common Self-Regulatory Failure," *Psychology*, 2014, *5* (13), 1488.

Snijders, Tom AB and Roel J Bosker, *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage, 2011.

Steel, Piers, "The Nature of Procrastination: A Meta-Analytic and Theoretical Review of Quintessential Self-Regulatory Failure.," *Psychological bulletin*, 2007, *133* (1), 65.

Takeuchi, Kan, "Non-parametric Test of Time Consistency: Present Bias and Future Bias," *Games and Economic Behavior*, 2011, *71* (2), 456–478.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *The Review of Economic Studies*, 2018, *85* (4), 2462–2496.

Train, Kenneth E, *Discrete Choice Methods with Simulation*, Cambridge University Press, 2009.

U.S. Congressional Budget Office (CBO), "Effects of Using Generic Drugs on Medicare's Prescription Drug Spending," Technical Report Pub. No. 4043 2012.

U.S. DEPARTMENT OF COMMERCE Office of Inspector General (OIG), "United States Patent and Trademark Office: USPTO Should Reassess How Examiner Goals, Performance Appraisal Plans, and The Award System Stimulate and Reward Examiner Production," Technical Report IPE-15722 2004.

_ , "United States Patent and Trademark Office: USPTO Needs to Strengthen Patent Quality Assurance Practices," Technical Report OIG-15-026-A 2015.

U.S. Government Accountability Office (GAO), "The United States Patent and Trademark Office: Hiring Efforts Are Not Sufficient to Reduce the Patent Application Backlog," Technical Report GAO-07-1102 2007.

_ , "DRUG PRICING: Research on Savings from Generic Drug Use," Technical Report GAO-12-371R 2012.

_ , "Intellectual Property: Patent Office Should Define Quality, Reassess Incentives, and Improve Clarity," Technical Report GAO-16-490 2016.

Watanabe, Sumio, "A Widely Applicable Bayesian Information Criterion," *Journal of Machine Learning Research*, 2013, *14* (Mar), 867–897.

Tables and Figures

Table 1: Summary statistics

	year									average
	2001	2002	2003	2004	2005	2006	2007	2008	2009	
(A) Patent Examiner Statistics										
number of patent examiners	432	447	473	514	563	573	595	599	636	536.89
experience year	5.64	6.13	6.58	6.73	6.84	7.22	7.64	8.24	8.60	7.07
entry rate	12.92	9.98	9.39	11.17	11.71	7.36	7.40	5.74	4.03	8.86
exit rate	6.36	5.70	3.52	4.86	3.18	5.92	4.25	4.03	2.55	4.49
(B) Throughput Statistics										
number of first office actions	1.80	1.81	1.68	1.47	1.62	1.87	1.96	1.84	1.77	1.76
failure rate of first office action	0.16	0.14	0.15	0.16	0.16	0.15	0.15	0.14	0.10	0.15

Notes: This table reports summary statistics on patent examiners and patent examination throughput. For the statistics on patent examiners, all values are annual averages. For the statistics on patent examination throughput, the average values per patent examiners per biweekly period are presented.

Table 2: The impact of a short timeframe on the number of first office actions per two weeks

Variable names	(1)	(2)	(3)	(4)	(5)
(A) Parameter estimates					
Short timeframe					
S_b for reference period	-0.068	-0.051	-0.054	-0.052	0.002
	[0.113]	[0.218]	[0.192]	[0.184]	[0.975]
S_{b+1} for lag period				0.054	
				[0.219]	
S_{b-1} for lead period					-0.054
					[0.219]
log($Backlog$) for new applications	0.046	0.047	-0.085	-0.116	-0.116
	[0.681]	[0.676]	[0.557]	[0.433]	[0.433]
log($Backlog$) for amend applications	-0.385	-0.373	-0.376	-0.353	-0.353
	[0.010]	[0.008]	[0.011]	[0.014]	[0.014]
$Num\_Examiners$			1.115	1.335	1.335
			[0.285]	[0.199]	[0.199]
$Thanksgiving$ Dummy		-0.203	-0.201	-0.22	-0.22
		[0.006]	[0.005]	[0.001]	[0.001]
(B) Measures of goodness-of-fit					
Adj. R-squared	0.66	0.672	0.674	0.68	0.68
No. of observations	234	234	234	233	233

Notes: This table reports the main estimation results for the regression model (1). The estimated coefficients of the variables except for the year and biweekly period dummies are presented. The p-values in brackets are calculated based on robust standard errors clustered at the art unit level. The adjusted R-squared (Adj. R-squared) values are presented.

Table 3: Reduced-form regression evidence for present-bias-induced procrastination.

	(1)	(2)	(3)	(4)	(5)	(6)
(A) Marginal effect of the short timeframe						
S_b for reference period	0.010	0.010	0.009	0.011	0.007	0.009
	[0.007]	[0.008]	[0.011]	[0.001]	[0.152]	[0.026]
S_{b+1} for lag period						-0.002
						[0.613]
S_{b-1} for lead period						-0.003
						[0.349]
(B) Standard deviation-correlation matrix						
σ_0	0.674	0.674	0.674	0.957	0.679	0.688
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
σ_1	0.205	0.206	0.171	0.658	0.220	0.190
	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
σ_{01}	-0.267	-0.271	-0.292	-0.286	-0.234	-0.372
	[0.008]	[0.009]	[0.023]	[0.000]	[0.075]	[0.000]
(C) Measures of goodness-of-fit						
AIC	0.763	0.772	0.764	0.751	0.779	0.763
BIC	0.772	0.773	0.779	0.762	0.796	0.773
AUC.ROC	0.700	0.700	0.701	0.704	0.726	0.704
No. of observations	176797	176797	176797	175163	98357	176,797

Notes: This table reports the main estimation results for the regression model (2). We construe the model as a multi-level logit model with a random intercept and random coefficients and estimate the parameters, $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ by a maximum likelihood method, except for the results in column 5, where a fixed-effects approach is used. Panel A presents the marginal effect of the short timeframe indicator S_{bt} on the first office action failure, while panel B shows the estimates of the standard deviation-correlation matrix $\boldsymbol{\sigma}$. The p-values in brackets are calculated based on robust standard errors clustered at the art unit level. As measures of goodness-of-fit, the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the area under the ROC curve (AU.ROC) from the receiver operating characteristics (ROC) analysis are presented in panel C. The AIC and BIC are normalized by the number of observations.

Table 4: The estimation results for the key parameters of the present-bias-induced procrastination model.

	baseline	wide prior	utility curvature		reward target	
	(1)	(2)	(3)	(4)	(5)	(6)
		$\delta \in [0.1, 1.0]$	$\gamma = 0$	$\gamma = 2.0$	$\beta = 0.4$	$\beta = 0.8$
(A) Marginal posterior means						
Present bias factors $\bar{\hat{\beta}}_i$						
1st Qt.	0.430	0.432	0.353	0.407	0.465	0.549
	(0.282, 0.619)	(0.312, 0.570)	(0.286, 0.435)	(0.316, 0.517)	(0.377, 0.557)	(0.442, 0.662)
2nd Qt.	0.600	0.599	0.556	0.585	0.603	0.733
	(0.495, 0.716)	(0.495, 0.714)	(0.423, 0.715)	(0.463, 0.748)	(0.317, 0.930)	(0.449, 1.030)
3rd Qt.	0.792	0.792	0.790	0.802	0.732	0.902
	(0.642, 0.966)	(0.574, 1.045)	(0.395, 1.085)	(0.584, 1.053)	(0.567, 0.903)	(0.653, 1.088)
Discount factor $\bar{\hat{\delta}}$	0.949	0.321	0.950	0.949	0.949	0.949
	(0.902, 0.997)	(0.104, 0.904)	(0.903, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.999)
Admissible failure rate $\bar{\hat{\tau}}$	0.044	0.044	0.055	0.050	0.033	0.063
	(0.040, 0.047)	(0.041, 0.046)	(0.053, 0.057)	(0.047, 0.052)	(0.030, 0.037)	(0.062, 0.065)
(B) Measures of goodness-of-fit						
WAIC	0.756	0.756	0.756	0.756	0.767	0.757
AUC. ROC	0.692	0.692	0.692	0.692	0.692	0.690
(C) Present-biased patent examiners (%)						
	69.014	70.000	62.113	63.380	85.493	54.789

Notes: This table reports the Bayesian estimation results based on the posterior distribution given by equation (7). The means of the marginal posterior distributions are presented for the key parameters. For the individual-specific present bias factors, the first to third quartiles of the marginal posterior means' distribution are presented. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. As goodness-fit-measures, we report the widely applicable information criterion (WAIC), which is normalized by the number of observations, and the area under the ROC curve (AU. ROC) from the receiver operating characteristics (ROC) analysis. A patent examiner is defined as present biased if the upper limit of 95 percent HDI of his or her present-bias factor does not reach the value of 1.0.

Table 5: The estimation results for the key parameters of the time-consistent procrastination model.

	(1) $\delta_i, \zeta \in [0.1, 1.0]$	(2) $\delta_i, \zeta \in [0.9, 1.0]$
(A) Marginal posterior means		
Discount factor $\bar{\hat{\delta}}_i$		
1st Qt.	0.537	0.950
	(0.129, 0.963)	(0.904, 0.996)
2nd Qt.	0.554	0.950
	(0.140, 0.965)	(0.904, 0.996)
3rd Qt.	0.570	0.950
	(0.141, 0.967)	(0.904, 0.996)
Patent office's belief $\bar{\hat{\zeta}}$	0.552	0.950
	(0.123, 0.977)	(0.903, 0.997)
Admissible failure rate $\bar{\hat{\tau}}$	0.131	0.131
	(0.129, 0.133)	(0.129, 0.132)
(B) Measures of goodness-of-fit		
WAIC	0.799	0.800
AUC. ROC	0.592	0.592

Notes: This table reports the Bayesian estimation results under the assumption that patent examiners are exponential discounters having individual-specific discount factors. The means of the marginal posterior distributions are presented for the key parameters. For the individual-specific discount factors, the first to third quartiles of the marginal posterior means' distribution are presented. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. We assume different priors for the discount-factor-related parameters: uniform priors on $[0.1, 1.0]$ for the results in column 1 and uniform priors on $[0.9, 1.0]$ for the results in column 2.

Table 6: The effect of the present bias measure on the PTA period using a tobit regression model

	(1)	(2)	(3)	(4)
(A) Marginal effect estimates				
$PresentBiasMeasure$	31.815	18.777	30.222	32.266
	[0.000]	[0.000]	[0.007]	[0.000]
$Failure$		150.293	145.673	110.855
		[0.000]	[0.000]	[0.000]
(B) Measures of goodness-of-fit				
AIC	9.477	9.466	9.410	8.192
BIC	9.488	9.476	9.421	8.203
No. of observations	97951	97951	97951	101740

Notes: This table reports the marginal effects on the PTA period of a patent application based on the regression model (8). The p-values in brackets are calculated based on robust standard errors clustered at the art unit level. Akaike information criteria (AIC) and Bayesian information criteria (BIC) are given for each specification. The AIC and BIC are normalized by the number of observations. The sample consists of applications that were granted patents because unapproved patent applications do not have PTA records. The sample period is 2001-2009 for the estimates in columns 1-3, and the sample period is 2010-2017 for the estimates in column 4. We include the random effects for the regression specification presented in column 3.

Table 7: The results of counterfactual simulation

	Failure rate reduction (%)	PTA reduction (days)
the whole sample	0.07	6.62
	[30.44]	[2.66]
the bottom quintile group (the least present-biased)	0.02	1.07
	[17.02]	[0.58]
the top quintile group (the most present-biased)	0.12	0.12
	[30.28]	[4.41]

Notes: This table demonstrates the differences in patent examination accuracy (column 1) and patent pendency (column 2) between the current and hypothetical quota rules. The values in the brackets are the percentage changes relative to the original magnitudes. We assume the patent office cuts the patent examination quota in half from the original amount and shortens the deadline from 10 days to 5 days. The percentage reductions of the first office action failure, presented in column 1, are predicted by the behavioral model of present-bias-induced procrastination, whose key parameter values are given by the mean values of the posterior distributions reported in Table 4. To obtain the predicted values of the PTA reduction, which are reported in column 2, we combine the values of the predicted change of the first office action failure probability reported in column 1, with the marginal effects of the first office action failure on the PTA reported in Table 6.

Table 8: The impacts of patent examiner change and the quota structure change on a pharmaceutical patent

Counterfactual scenarios	PTA reduction (days)	Consumer benefits ($1M)
(i) The original patent examiner reviews under the hypothetical one-week quota rule.	12.66	28.10
(ii) A hypothetical exponential discounting patent examiner review under the original two-week quota rule.	43.24	95.95
(iii) A hypothetical discounting patent examiner reviews under the hypothetical one-week quota rule.	54.28	120.46

Notes: This table shows the change in the PTA period for a pharmaceutical patent and the associated increase in consumer benefits under hypothetical scenarios. We take U.S. patent 7,491,725, related to a blockbuster cancer drug, called SPRYCEL, whose annual sales are approximately $1.2 billion. To predict the values of the PTA reduction, we base our model on the behavioral model of present-bias-induced procrastination, whose key parameter values are given by the mean values of the posterior distributions reported in Table 4. The consumer benefits are calculated based on the following assumptions: (i) the longer patent term delays the introduction of a generic drug of SPRYCEL to the market, (ii) 90 percent of consumers switch when a generic drug appears on the market, and (iii) the generic drug is sold for 75% less than the brand-name drug.

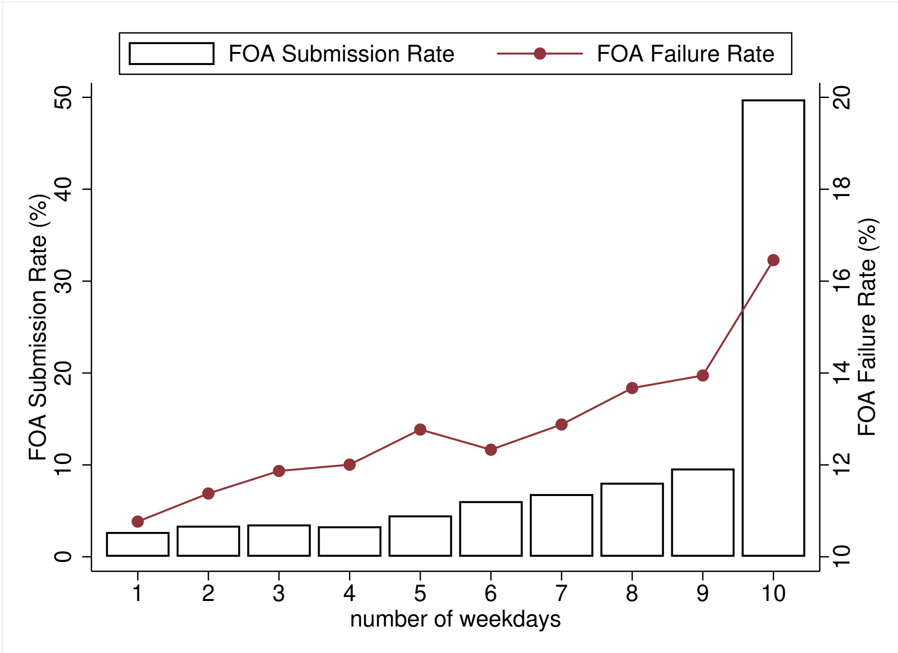

Figure 1: A distribution of the first office action (FOA) submission rates (bar graph) and the time course of the FOA failure rate (line graph). The horizontal axis represents the number of weekdays for the biweekly quota period.

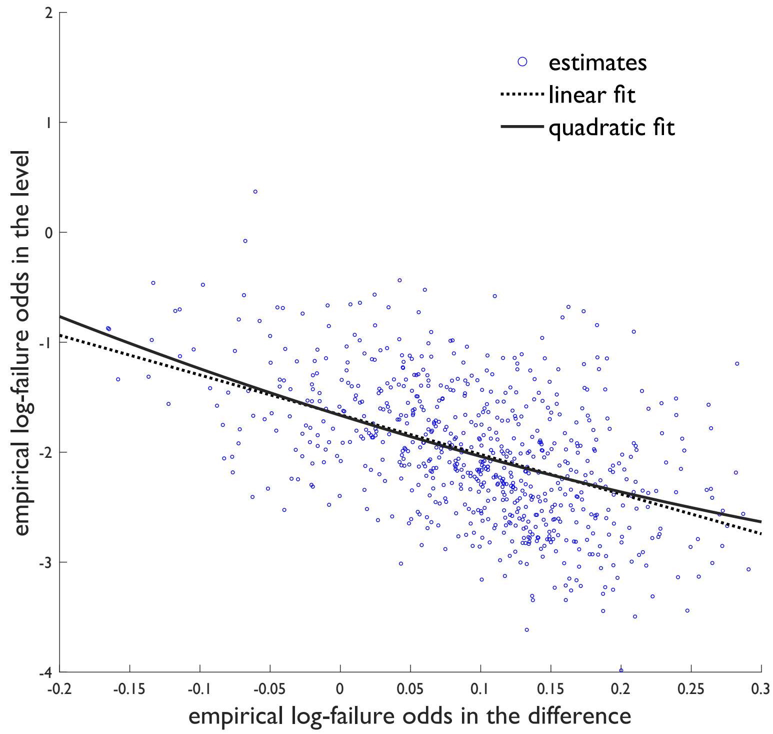

Figure 2: A scatterplot of the estimated empirical log-failure odds. The vertical axis represents the empirical log-failure odds in the level, $\hat{\phi}_0$, and the horizontal axis represents the empirical log-failure odds in the difference, $\hat{\phi}_1$. Both first- and second-order polynomial fits are shown.

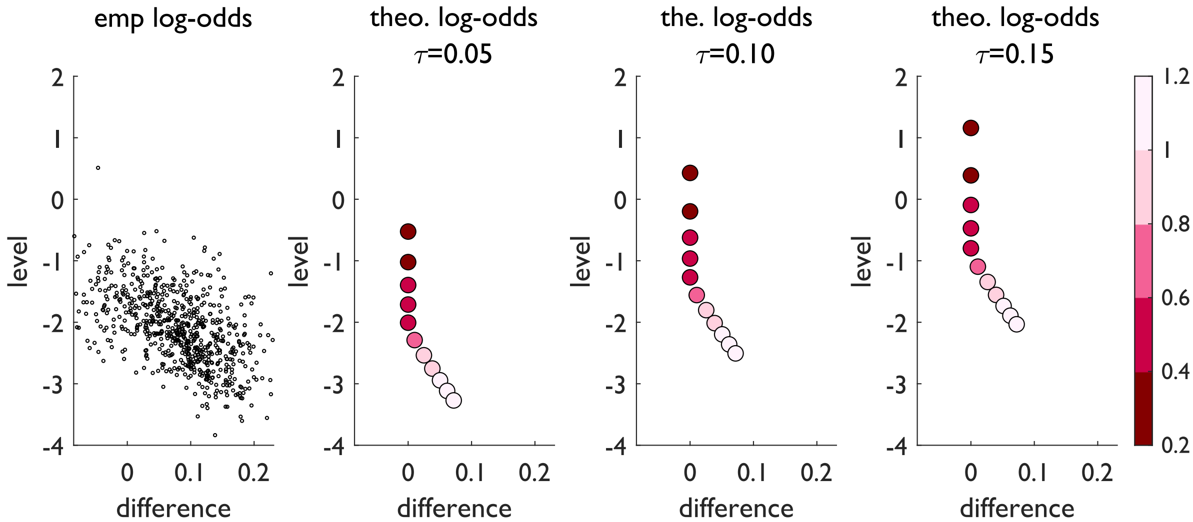

Figure 3: The empirical log-failure odds and the simulated theoretical log-failure odds. The vertical axis represents the log-failure odds in the level, and the horizontal axis represents the log-failure odds in the difference. A darker color indicates a higher degree of present-biased preference.

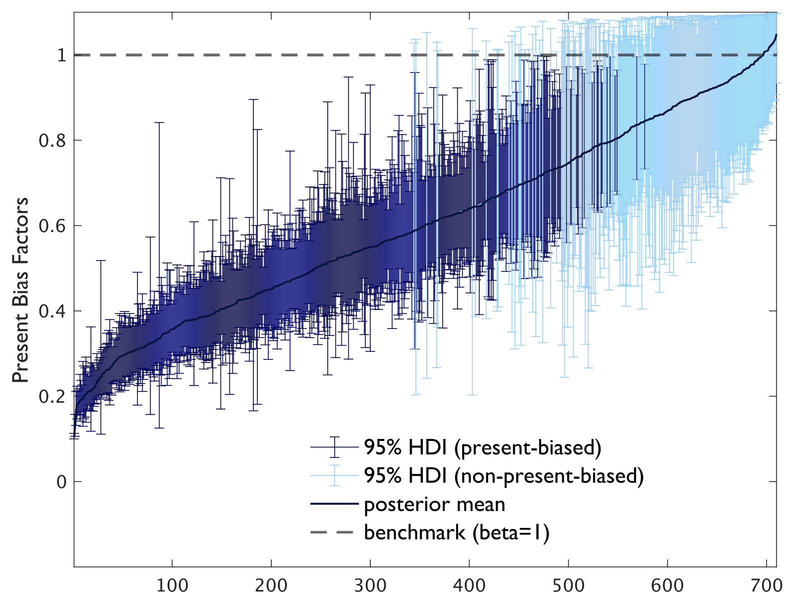

Figure 4: Estimated marginal posterior distributions of the present bias factor for patent examiners. The bar represents the 95% HDI of the marginal posterior distribution of the present bias factor of a patent examiner. A darker color is for the present-biased patent examiners, while a lighter color is for the other patent examiners.

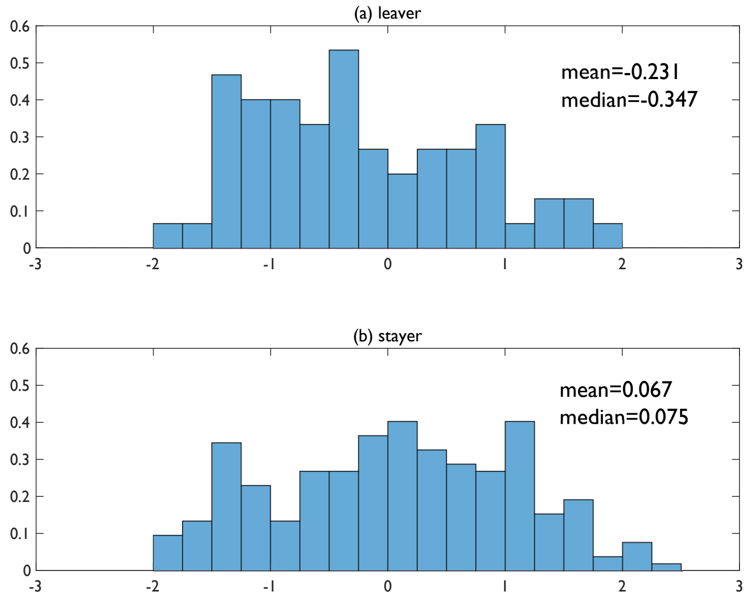

Figure 5: Histograms of the present bias measure between the leaver group (top) and the stayer group (bottom). The present bias measure is defined by a negative value of the standardized present bias factor. The mean and median of the distribution are inset in the figure.

Appendix A Reward Scheme

Suppose that a contract is signed between the patent office (the principal) and a patent examiner (agent) specifying a reward that is paid at the time a task quota is completed. As in the standard principal-agent model, the principal cannot observe the agent's work time and thus controls the reward to make the agent work to the level considered desirable.

The decision processes of the patent office and patent examiner are modeled as a sequential game. The timing of events is as follows.

1. The patent office offers a reward scheme to the patent examiner.
2. If the patent examiner accepts the reward scheme, he or she dynamically allocates his or her work time to complete the quota by the deadline.
3. The patent office evaluates the submitted work and pays the patent examiner according to the reward scheme.

The game is solved in a backward manner. The analysis in Section 5 provides the solution for the process shown in 2 and 3 above. To be more concrete, under a scheme in which the probability of receiving a reward r for working the standard eighty hours in two weeks is τ, a patent examiner with time preference parameters (β, δ) chooses to work $S_D^*(\beta, \delta, \tau, r)$ hours in two weeks.

We now turn to the patent office's decision. How does it determine the reward r^*? To answer this question, we must know the patent office's goals. However, the agency's publicized policy goals are too general and do not provide sufficient information to determine the reward for the individual patent examination process.[86] Therefore, we presume that the patent office's goal is to keep the failure rate below the admissible failure rate τ. Under this policy target, the patent office lets the patent examiner work the standard eighty hours for two weeks, while the patent office cannot directly observe if the patent examiner is working that many hours.

We further assume that the admissible failure rate τ is common knowledge. If the patent office correctly knew the patent examiner's time preferences (β, δ), it could achieve the intended goal concerning the patent examination failure. However, in the real work environment, it is challenging for the patent agency to obtain private information on the time preferences for every patent examiner. Therefore, we assume that the patent office has *partial* knowledge of the patent examiner's time preferences. Concretely, we assume that the patent office perfectly knows the value of the discount factor δ but believes that the patent examiner is an exponential discounter with present-bias factor $\beta = 1$.

Under the reward determination scheme presented above, the reward r^* that the patent office selects satisfies the following equation: $S_D^*(1, \delta, \tau, r^*) = 80$. The next proposition ensures that the number of rewards r^* that solves the equation is no greater than one.

Proposition 1. *The total work time $S_D^*(1, \delta, \tau, r)$ of a time-consistent patent examiner is increasing in the reward r for any values of δ and τ. In other words, when two schemes with different rewards $r_1 < r_2$ are given, the exponential discounting patent examiner works more under the higher reward scheme so that $S_D^*(1, \delta, \tau, r_1) < S_D^*(1, \delta, \tau, r_2)$.*

[86] According to NAPA (2005), the USPTO has announced strategic goals, including to "optimize the quality and timeliness of the patents and trademarks, and to provide leadership to improve intellectual property right in the United States and the World."

Proof. We prove the result for the case where the instantaneous utility takes the CRRA form.

The time allocation problem for an exponential discounting patent examiner is formulated as

$$\max_{w_1,\cdots,w_D} \left\{ \sum_{d=1}^{D} \delta^{d-1} u(24 - w_d) + \delta^D R\left(\sum_{d=1}^{D} w_d\right) \right\}$$

The first-order condition is given by

$$u'(24 - w_d^*) = \delta^{D+1-d} R'\left(\sum_{k=1}^{D} w_k^*\right)$$

for $d = 1, \cdots, D$

We can immediately see from the condition that $u'(24-w_d^*) = \delta u'(24-w_{d+1}^*)$. Knowing that $u'(24-w) = -(24-w)^{-\gamma}$ for the CRRA utility function u, we have the following recursive equation:

$$(24 - w_{d+1}^*) = \delta^{1/\gamma}(24 - w_d^*)$$

Consequently, the optimal total work time can be written as

$$S_D^* = \sum_{k=1}^{D} w_k^* = \left\{ \frac{1-\delta^{\frac{D+1}{\gamma}}}{1-\delta^{\frac{1}{\gamma}}} \right\} (w_1^* - 24) + 24D$$

To avoid notational clutter, we use the following expression:

$$S_D^* = A(\delta) w_1^* + B(\delta)$$

where $A(\delta) = (1-\delta^{\frac{D+1}{\gamma}})/(1-\delta^{\frac{1}{\gamma}})$ and $B(\delta) = 24(D - A(\delta))$.

By substituting the equation above into the first-order condition for the first day $d = 1$, we obtain the following equation:

$$\frac{\gamma}{(24-w_1^*)^{1+\gamma}} + r\delta^D \left(\frac{\log \tau}{80}\right) \exp\left\{\frac{\log \tau}{80}\big(A(\delta)w_1^* + B(\delta)\big)\right\} = 0$$

The last term on the left-hand side of the equation above follows because $R'(S_D) = -\frac{r \log \tau}{80} \exp\left(\frac{\log \tau}{80} S_D\right)$

Let $F(w_1, \delta, \tau, r)$ denote the function that appears in the left-hand side of the equation above. The optimal first day work time w_1^* then satisfies $F(w_1^*, r, \delta, \tau, r) = 0$. The partial derivatives of the function have the following signs:

$$F_1(w_1, r) = \frac{\gamma(1+\gamma)}{(24-w_1)^{2+\gamma}} + r\delta^D A(\delta) \left(\frac{\log \tau}{80}\right)^2 \exp\left\{\frac{\log \tau}{80}\big(A(\delta)w_1^* + B(\delta)\big)\right\} > 0$$
$$F_2(w_1, r) = \delta^D \left(\frac{\log \tau}{80}\right) \exp\left\{\frac{\log \tau}{80}\big(A(\delta)w_1^* + B(\delta)\big)\right\} < 0$$

since $\gamma > 0$, $\delta \in [0,1]$, $\tau \in [0,1]$ and $A(\delta) > 0$. To establish the result of the proposition, we first turn to the work time that the patent examiner chooses at the initial day and show that it is increasing in the reward to be received.

Consider two rewards $r_1 < r_2$, and denote the corresponding optimal work time on day one by $w_1^*(r_1)$ and

$w_1^*(r_2)$. By construction, these values satisfy $F\big(w_1^*(r_1), \delta, \tau, r_1\big) = F\big(w_1^*(r_2), \delta, \tau, r_2\big) = 0$. Since the function $F\big(w_1, \delta, \tau, r\big)$ is decreasing in r, we know that $0 = F\big(w_1^*(r_1), \delta, \tau, r_1\big) > F\big(w_1^*(r_1), \delta, \tau, r_2\big)$. This result implies that $0 = F\big(w_1^*(r_2), \delta, \tau, r_2\big) > F\big(w_1^*(r_1), \delta, \tau, r_2\big)$. Given that the function $F\big(w_1, \delta, \tau, r\big)$ is increasing in w_1, we can say that $w_1^*(r_1) < w_1^*(r_2)$.

The same relationship between the work time and reward holds for any day in a biweekly period. To see this, it should be noted that

$$w_d^*(r) = 24 - \delta^{\frac{d-1}{\gamma}} \{24 - w_1^*(r)\}$$

where $w_d^*(r)$ represents the optimal work time on day d as a function of reward r. Therefore, $w_d^*(r_1) < w_d^*(r_2)$ implies $w_d^*(r_1) > w_d^*(r_2)$ for day $d \geq 2$. The patent examiner increases the work time overall for all days in a biweekly period if offered a higher reward scheme.

Now, it is straightforward to show that the total work time, $S_D^*(1, \delta, \tau, r)$, is increasing in the success-contingent reward r. Given that it is the sum of the single-day work time that the patent examiner chooses, $w_d^*(r_1) < w_d^*(r_2)$ for any $d = 1, \cdots, D$ implies $S_D^*(1, \delta, \tau, r_1) < S_D^*(1, \delta, \tau, r_2)$.

$\square$

Figure A.1 demonstrates that the target reward $r^*(\delta, \tau)$ is decreasing in the admissible failure rate τ for given values of discount factor δ. The result has a clear intuition. If the patent office wishes to achieve a lower error rate, it must pay a higher reward to the patent examiner, if all other conditions are kept constant.

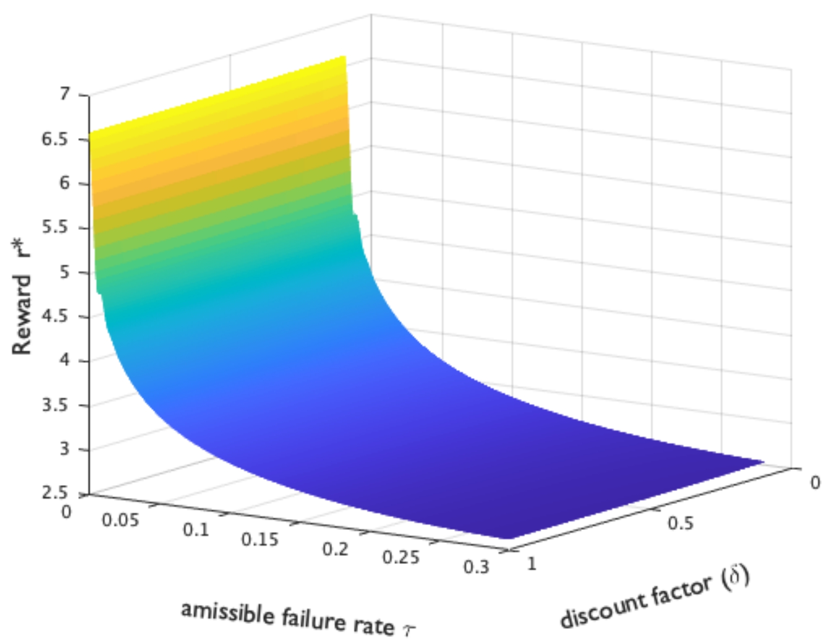

Figure A.1: Simulated values of the reward for various combinations of discount factors and admissible failure rates. The value of the reward takes the natural logarithm.

Appendix B Simulated Work Pattern

Tables B.1 and B.2 provide the calibration exercise results concerning a patent examiner's work pattern. The start day of work, total work hours, and log-failure odds are simulated for five distinct values of present-bias factor β ranging from 0.2 to 1.0 with a step size of 0.2. We set the annual discount factor $\delta = 0.95$ in both tables but change the values of the admissible failure rate in each table. We set $\tau = 0.15$ for the results given in Table B.1 and $\tau = 0.10$ for those given in Table B.2. The results are reported under the standard timeframe case and the short timeframe case. In the rightmost column, the differences in the log-failure odds between the timeframes are reported.

The outcomes are generally consistent with the findings on time allocation that the previous theoretical papers have proved (Herweg and Müller 2011). Namely, patent examiners with lower present-bias factor β tend to start the assigned tasks later and work less than those with high values of β. Furthermore, the higher the present-bias factor (i.e., the less present-biased the preferences), the larger the log-failure odds difference.

Table B.1: Simulated work pattern of a patent examiner ($\tau = 0.15$)

	standard (10 business days)			short (9 business days)			log-failure odds
β	start day	total hours	log-failure odds	start day	total hours	log-failure odds	diference
0.20	8	11.507	1.159	7	11.507	1.159	0.000
0.40	5	31.232	-0.093	4	31.232	-0.093	0.000
0.60	3	49.304	-0.797	2	49.304	-0.797	0.000
0.80	1	66.494	-1.345	1	65.614	-1.319	0.026
1.00	1	80.000	-1.735	1	78.161	-1.684	0.052
1.20	1	90.914	-2.033	1	88.236	-1.961	0.072

Notes: This table describes the relationship between the present-bias factor β and the work pattern of a patent examiner. We set the discount factor $\delta = 0.95$ and the admissible failure rate $\tau = 0.15$.

Table B.2: Simulated work pattern of a patent examiner ($\tau = 0.10$)

	standard (10 business days)			short (9 business days)			log-failure odds
β	start day	total hours	log-failure odds	start day	total hours	log-failure odds	diference
0.20	8	17.434	0.428	7	17.434	0.428	0.000
0.40	6	36.572	-0.623	5	36.572	-0.623	0.000
0.60	2	52.731	-1.270	1	52.731	-1.270	0.000
0.80	1	67.958	-1.803	1	67.199	-1.778	0.025
1.00	1	80.000	-2.197	1	78.419	-2.146	0.051
1.20	1	89.790	-2.506	1	87.491	-2.434	0.072

Notes: This table describes the relationship between the present-bias factor β and the work pattern of a patent examiner. We set the discount factor $\delta = 0.95$ and the admissible failure rate $\tau = 0.10$.

Appendix C Bayesian Inference Procedure

We compute the posterior distribution, given by equation (7), in two stages. In the first stage, we perform a grid computation of the posterior distribution. Although the computation method explores the probability distribution at only a limited point in the parameter space, it provides information on the approximate location and shape of the distribution. In the second stage, we compute the posterior density at high resolution using an MCMC method based on the Metropolis-Hastings (MH) algorithm. The MCMC approach outperforms the grid computation because it is designed to search more intensively for higher-probability regions than lower ones in the parameter space.

We start with a grid approximation of the posterior distribution via the following brute force method: (i) define a grid in the parameter space, (ii) for each grid point, calculate the values of the likelihood function and the prior density and multiply them, and (iii) normalize the multiplied values in the previous step. Finally, divide the results at each point by the sum of those of all points.

One should be careful about the grid choice because a massive number of grid points makes the computation virtually intractable. For example, suppose that we use 100 grid points for each of the structural parameters $\boldsymbol{\theta}$ when the number of examiners is $N = 719$. Although the number of required evaluation points is reduced due to the multiplicative nature of the likelihood function for independent data, there are $100 \times 100 \times 100 \times 719 \simeq 7.19 \times 10^8$ grid points to be evaluated, which requires formidable computational resources.

To avoid the computational burden, we focus on grid points for the parameters of the individual present-bias factors $\boldsymbol{\beta}$ and the admissible failure rate τ, while we fix the value of the discount factor at $\delta = 0.95$. Furthermore, we keep up to 80 grid points for each parameter. Consequently, the total number of grid points is $80 \times 80 \times 719 \simeq 4.60 \times 10^6$, a manageable size for our workstation computer to process.[87] To achieve a dense evaluation of the parameter space, we interpolate the likelihood for a nongrid point based on the already-computed values of the adjacent grid points.

Figure C.1 presents graphs of the marginal posterior densities of the present-bias factors for the first seven examiners and the admissible failure rate estimated via grid computation. [88]

We then estimate the posterior distribution using the Metropolis-Hastings sampler by means of the following steps:

1. Choose an initial value of the key parameter $\boldsymbol{\theta}^{(0)}$.
2. Choose a proposal parameter value $\boldsymbol{\theta}^{(j+1)}$ sampled from a probability distribution dependent on the previous value, $\boldsymbol{\theta}^{(j)}$.
3. Update the parameter value $\boldsymbol{\theta}^{(j)}$ by $\boldsymbol{\theta}^{(j+1)}$ with the following probability

$$r_{MH} = \min\left[1, \exp\left\{\ell(\boldsymbol{\theta}^{(j+1)}, \hat{\boldsymbol{\phi}})\right\} / \exp\left\{\ell(\boldsymbol{\theta}^{(j)}, \hat{\boldsymbol{\phi}})\right\}\right]$$

In practice, in iteration step j, we adopt the proposed value $\boldsymbol{\theta}^{(j+1)}$ if the value, r_{MH}, which is often called the Metropolis-Hastings criterion, is larger than the value taken from a uniform distribution on the interval

[87] The calculations were performed on a workstation with an 18-core Intel Xeon processor and 128 GB of memory using the computer language `Julia` 1.0.

[88] The results presented are for the baseline case, and the results from alternative specifications are available upon request.

of [0 1]; otherwise, we keep the old parameter value $\boldsymbol{\theta}^{(j)}$.

Several remarks are in order.

- We run $R = 10^6$ iterations of the two independent chains from the MH sampling algorithm, with ten percent of the total iterations used for the burn-in period.
- The sampler's initial values are determined based on the location information obtained from the grid approximation method. They are set to be the means of the grid-approximated posterior distributions of the corresponding parameters.
- A "random walk" method is used to propose a new parameter value in the sampler. Specifically, let θ be a single parameter. The proposed value $\theta^{(j+1)}$ is then generated from the current value $\theta^{(j)}$ according to the following normal distribution: $\theta^{(j+1)} \sim N\big(\theta^{(j)}, \kappa_\theta \hat{\sigma}_\theta^2\big)$ where $\hat{\sigma}_\theta^2$ represents the variance of the grid-approximated posterior distribution, and κ_θ represents a "tuning" parameter, which is calibrated such that the acceptance rate of the sampler is approximately 25 percent.
- Block-wise sampling is implemented. A detailed description is as follows: the set of the vector parameters $\boldsymbol{\theta}$ is partitioned into three blocks consisting of $\boldsymbol{\beta}$, δ and τ. The values are updated simultaneously within the first block, followed by the next block.
- The samples are "thinned" to save memory and disk space. All but the $th = 100$th observations are discarded to construct the inference.

Figures C.2 and C.3 present trace plots (plots of the value generated from the sample against the number of simulations) and histograms of the marginal posterior densities. For the individual-specific present-bias factors, those of the first seven examiners are reported. Note that the results indicated by different colors in each figure correspond to independent MH chains.

To assess the sampler's convergence, we perform two popular tests of nonconvergence: the Geweke test and the Heidelberger-Welch test.[89] The results are graphically summarized in Figure C.4, where the greater of the p-values from the tests performed on two independent samplers are plotted against the parameter set. To increase the visibility, we let solid circles represent the cases where the p-value is less than five percent, indicating that the null hypothesis is rejected. As demonstrated, the diagnostic test results vary widely by parameter, yet there is little evidence against stationarity. Indeed, the tests detect nonstationarity for less than 3 percent of all the parameters. Therefore, the MH sampling algorithm successfully generates an accurate characterization of the parameters' posterior distributions.

[89] The Geweke test compares values in the early part of the sampler to those in the latter part of the sampler to detect convergence failure. The Heidelberger-Welch test uses the Cramer-von-Mises statistics to assess evidence of stationarity. For a detailed explanation of the tests, see Jackman (2009). For implementation, we use the functions `gewekediag` and `heideldiag` in the `Mamba` package in the computer language `Julia` 1.0.

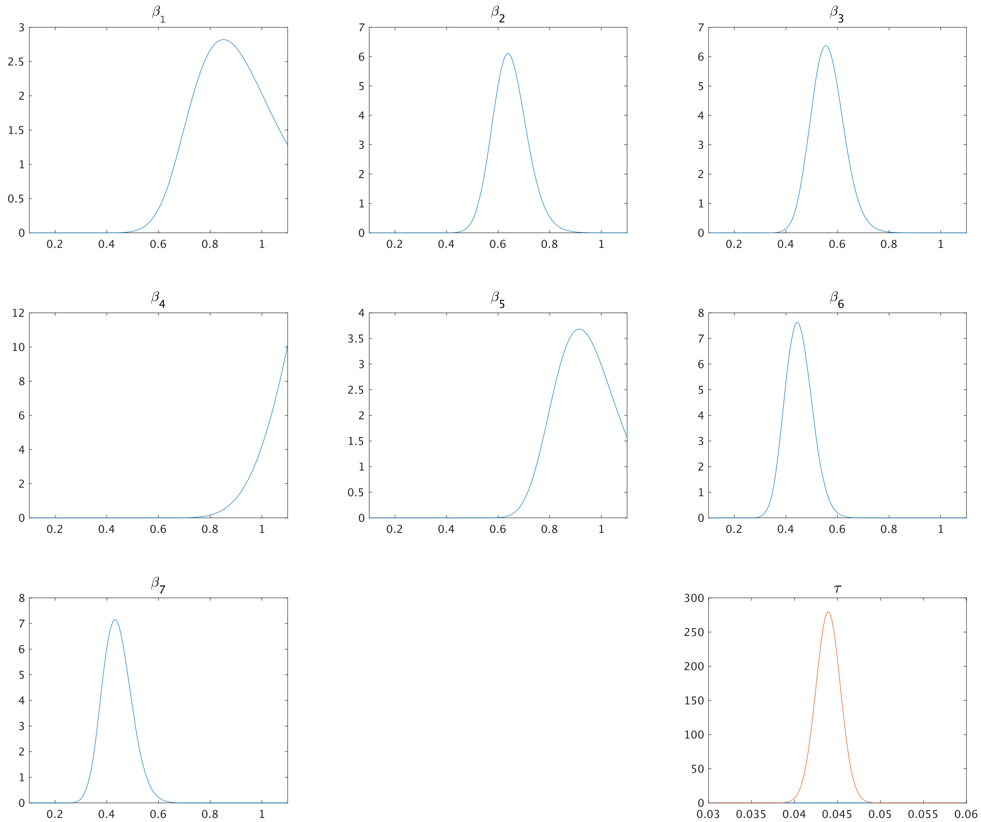

Figure C.1: Marginal posterior densities of the present-bias factors for the first seven examiners and the admissible failure rate estimated by grid computation. The marginal posterior density of the discount factor is not presented because it is not estimated.

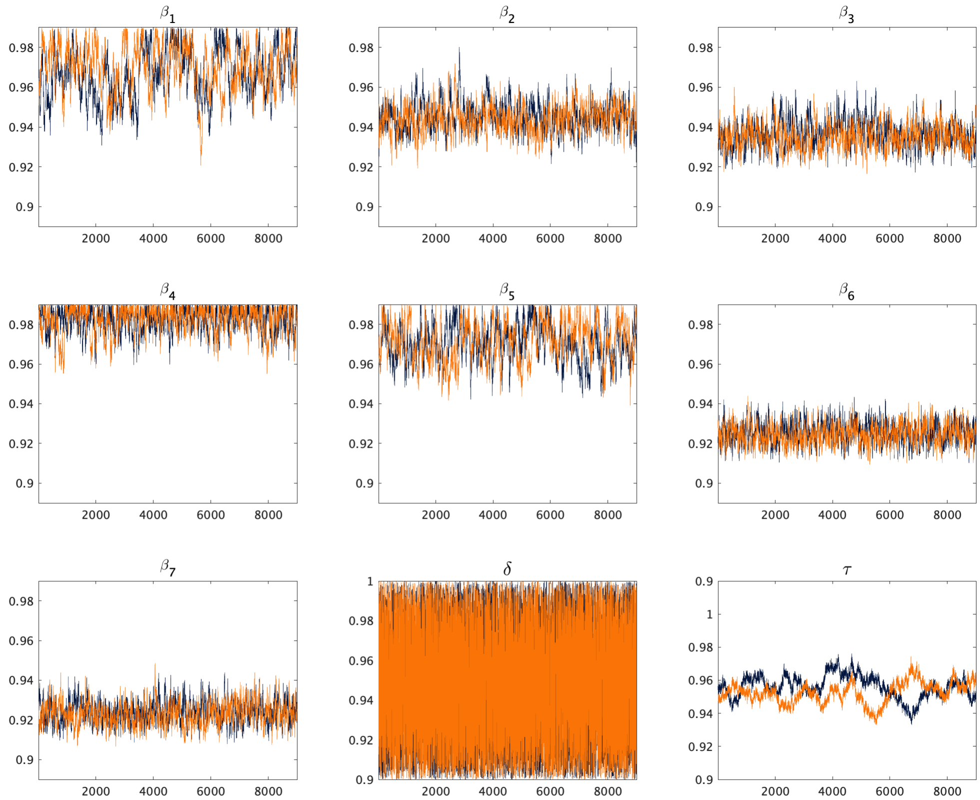

Figure C.2: Trace plots of the parameters obtained via Bayesian estimation and MCMC simulation. For the individual-specific present-bias factors, those of the first seven patent examiners are reported. Different colors in each figure correspond to independent MH chains.

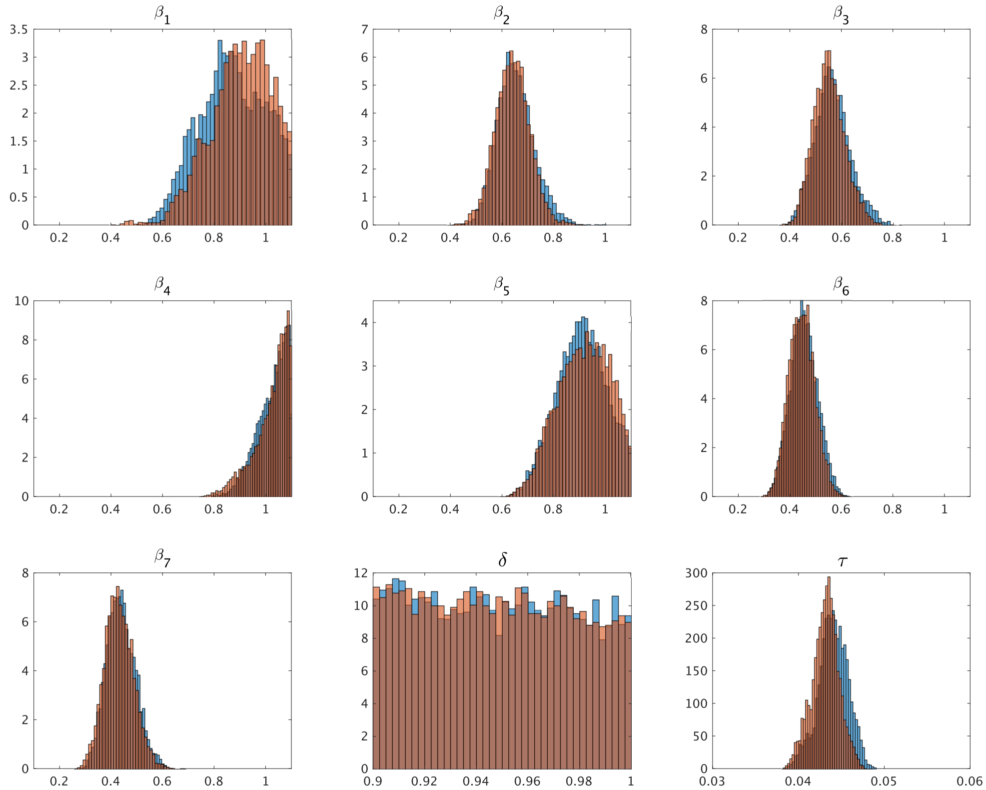

Figure C.3: Estimated marginal posterior densities for a portion of the parameters. The densities are for the present-bias factors of the first seven examiners, the discount factor, and the admissible failure rate, respectively. Different colors in each figure correspond to independent MH chains.

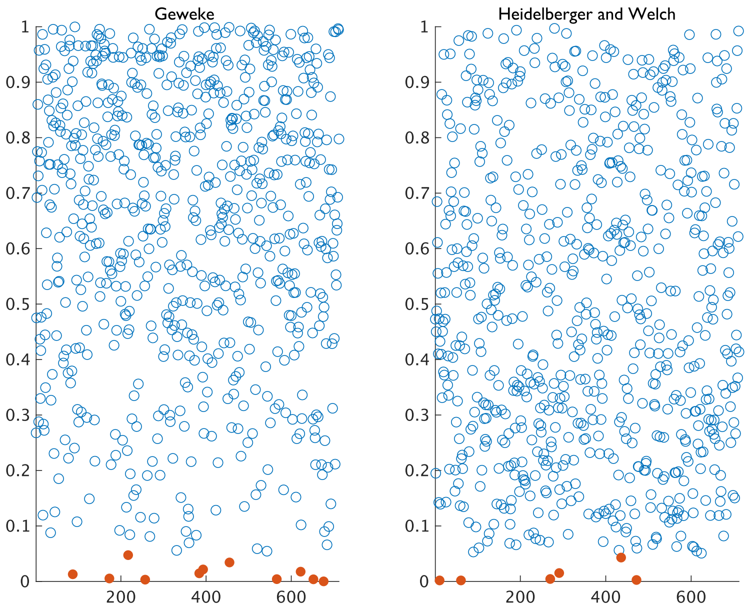

Figure C.4: The results of the Geweke test (left) and the Heidelberger-Welch test (right). The p-values from the tests performed on two independent samplers are plotted against the parameter set. Solid circles represent the cases where the p-value is less than five percent, indicating that the null hypothesis is rejected.

Appendix D Calibration Results for Exponential Discounters

The calibration results reported in Section 5 show that the assumption on the patent examiner's time preferences is consonant with the empirical data. However, an alternative assumption might be possible. One can think that heterogeneity in the standard discount factor δ, not in the present-bias factor β, is responsible for the observed heterogeneity in the empirical log-failure odds across examiners. In this appendix, we perform another calibration run to validate the alternative explanation, but this time, we assume exponential discounting examiners with various discounting factors.

The calibration procedure is the same as that in the main text. We modify a part of the assumptions concerning the reward scheme presented in Appendix A while keeping the core intact. Specifically, let us assume that the patent office has an *incorrect* belief in the discounting factor of the examiner. To distinguish it from the discount factor δ that the patent examiner has, we denote it by ζ. We assume that the patent office determined the target reward r^* under the belief that the examiner is an exponential discounter with discount factor $\zeta = 0.95$.

Figure D.1 provides scatterplots of the empirical and theoretical log-failure odds, which are parallel with those presented in Figure 3. In this case, hypothetical patent examiners differ in the discount factor: the individual-specific discount factor, δ, ranges from 0.9 to 1.0 by intervals of 0.1.

Remarkably, the calibrated theoretical log-failure odds exhibit distributions far narrower than those of the empirical log-failure odds. Given the poor fit of the calibrated model to the empirical pattern, the time-consistent hypothesis does not appear to be plausible.

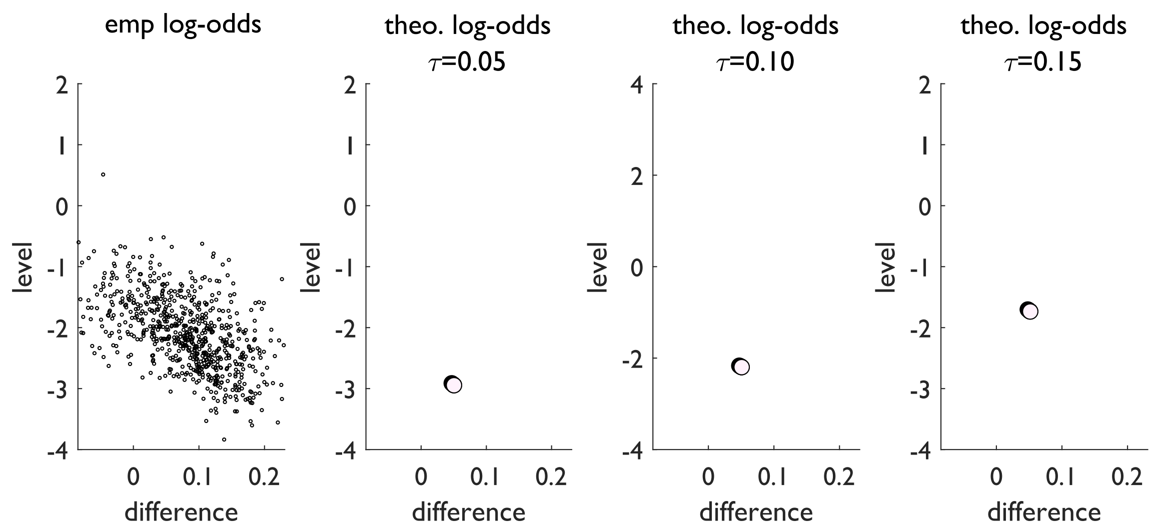

Figure D.1: The empirical log-failure odds and the simulated theoretical log-failure odds for the exponential discounting model. The vertical axis represents the log-failure odds in the level, and the horizontal axis represents the log-failure odds in the difference.

Appendix E Count Data Model for PTA Period

The generic regression model for the PTA period is given by

$$\begin{aligned}&\mathrm{E}\left(PTA_{ait}|X_{at},W_{it}\right)=\\&\exp\left(\alpha_{1t}+\alpha_{1X}X_{at}+\alpha_{1W}W_{it}+\rho_1 PresentBiasMeasure_i\right)=\mu_{ait},\end{aligned}\tag{E.1}$$

where a indexes the patent application, i indexes the patent examiner, and t indexes the year. We denote by μ_{ait} the conditional mean of the PTA period presented above and assume it is linked to the application and patent examiner characteristics. The left-hand side variable, PTA_{ait}, is the actual days of the extended patent term. The control variables include a measure of present-biased preferences, $PresentBiasMeasure_i$, the application characteristics X_{at}, and the patent examiner characteristics W_{it}.

The Poisson model assumes that the conditional probability of PTA_{iat} is governed by a Poisson process. The probability function at nonnegative integer values $k=0,1,2,\cdots$ is given by

$$f_{PO}(k|\mu_{ait})=\frac{\exp(-\mu_{ait})\mu_{ait}^k}{k!}$$

The negative binomial regression model incorporates an additional random term $\exp(\varepsilon_{ait})$ into the Poisson model in a multiplicative manner

$$\begin{aligned}&\mathrm{E}\left(PTA_{ait}|X_{at},W_{it}\right)=\\&\exp\left(\alpha_{1t}+\alpha_{1X}X_{at}+\alpha_{1W}W_{it}+\rho_1 PresentBiasMeasure_i\right)\exp(\varepsilon_{ait})\end{aligned}$$

where the distribution of $\exp(\varepsilon_{ait})$ follows a gamma distribution with parameter α_ε. The probability function at nonnegative integer values k is given by

$$f_{NB}(k|\mu_{ait})=\frac{\Gamma(y+\alpha_\varepsilon^{-1})}{\Gamma(\alpha_\varepsilon^{-1})\Gamma(k+1)}\left(\frac{1}{1+\alpha_\varepsilon\mu_{ait}}\right)^{\alpha_\varepsilon^{-1}}\left(\frac{\alpha_\varepsilon\mu_{ait}}{1+\alpha_\varepsilon\mu_{ait}}\right)^k$$

The zero-inflated models account for excess zeros by allowing two separate models—a binary model generating zero values and a zero truncated model for the remaining counts. In the context of patent examination, we assume that patent applications are classified into two categories—easy applications and difficult applications. We assume that the PTA is enforced only for difficult applications. Let ϕ_{ait} represent the probability that patent application a reviewed by patent examiner i is easy, and assume that it is given by

$$\phi_{ait}=\Lambda\left(\alpha_{1t}+\alpha_{1X}X_{at}+\alpha_{1W}W_{it}+\rho_1 PresentBiasMeasure_i\right)$$

where Λ represents the logistic distribution function. The probability function of PTA_{ait} at nonnegative integer values k for the zero-inflated Poisson (ZIP) regression model is

$$f_{ZIP}(k|\mu_{ait},\phi_{iat})=\begin{cases}\phi_{ait}+(1-\phi_{ait})f_{PO}(0|\mu_{ait}) & \text{if } k=0\\(1-\phi_{ait})f_{PO}(k|\mu_{ait}) & \text{if } k>0\end{cases}$$

On the other hand, the probability of PTA_{ait} at nonnegative integer values k for the zero-inflated negative binomial (ZINB) regression model is

$$f_{ZINB}(k|\mu_{ait}, \phi_{iat}) = \begin{cases} \phi_{ait} + (1-\phi_{ait})f_{NB}(0|\mu_{ait}) & \text{if } k = 0 \\ (1-\phi_{ait})f_{NB}(k|\mu_{ait}) & \text{if } k > 0 \end{cases}$$

The estimation results of the count-data regression models are presented in Table E.1. All the reported values are marginal effects with robust standard errors. According to AIC and BIC, the ZINB regression has the best fit for the data. Nonetheless, it should be noted that in all the specifications, the estimated coefficients of the present-bias measure exhibit qualitatively the same features as those from the tobit regression model. Specifically, while the magnitude of the marginal effect varies from specification to specification for the count data regression models, they all are positive and statistically significant. It is therefore suggested that the conclusions drawn from the tobit regression do not change.

Table E.1: The effect of the present-bias measure on the PTA period using count data regression models

	Poission	Neg. Bin.	Zero-Inflated Poisson	Zero-Inflated Neg. Bin.
((A) Marginal effect estimates				
small_entityr	-13.62	-12.82	-18.55	-14.16
	[0.000]	[0.000]	[0.000]	[0.000]
parent	-12.32	-9.987	-6.825	-8.978
	[0.000]	[0.000]	[0.070]	[0.000]
cip	26.6	28.17	36.94	29.79
	[0.000]	[0.000]	[0.000]	[0.000]
cont	-64.26	-51.54	-46.8	-44.41
	[0.000]	[0.000]	[0.000]	[0.000]
div	-83.22	-69.55	-69.49	-64.66
	[0.000]	[0.000]	[0.000]	[0.000]
foreign_priority	4.231	0.229	-8.037	-1.47
	[0.058]	[0.916]	[0.036]	[0.520]
ctrs	87.18	80.31	82.6	75.84
	[0.000]	[0.000]	[0.000]	[0.000]
accel	-158.1	-124.1	-152.6	-122.1
	[0.000]	[0.000]	[0.000]	[0.000]
fam_root	23.27	24.25	27.49	22.15
	[0.000]	[0.000]	[0.000]	[0.000]
PresentBiasMeasure	15.06	15.56	22.45	15.67
	[0.000]	[0.000]	[0.000]	[0.000]
Failure	88.04	89.54	110.7	92.81
	[0.000]	[0.000]	[0.000]	[0.000]
(B) Measures of goodness-of-fit				
AIC	307.942	145.930	10.155	9.380
BIC	307.953	145.951	10.166	9.402
(C) No. of observations	97951	97951	97951	97951

Notes: This table reports the marginal effects on the PTA period of a patent application based on the regression model (E.1). The sample consists of applications that were granted patents because unapproved patent applications do not have PTA records. The specifications are (i) Poisson, (ii) negative binomial (Neg. Bin.), (iii) zero-inflated Poisson, and (iv) zero-inflated negative binomial (Zero-Inflated Neg. Bin.). The p-values in brackets are calculated based on robust standard errors clustered at the art unit level. Akaike information criteria (AIC) and Bayesian information criteria (BIC) are given for each specification. The AIC and BIC are normalized by the number of observations.

Appendix F A Survival Analysis of Patent Examiner Job Duration

To add quantitative evidence of the relationship between job separation and present-biased preferences, we perform a survival analysis of patent examiner job duration using a Cox proportional hazard regression.

Let h_{it} be the hazard ratio that patent examiner i left the patent office at time t since joining. We estimate regression models of the form:

$$h_{it} = h_{0t} \exp\left(ArtUnit_i + EntryYear_i + \psi PresentBiasMeasure_i\right), \tag{F.1}$$

and of the form:

$$h_{it} = h_0 \exp\left(ArtUnit_i + EntryYear_i + \sum_{k=2}^{5} \psi_k PresentBiasDummy_k\right), \tag{F.2}$$

where h_{0t} represents the baseline hazard, $ArtUnit_i$ represents an art unit dummy variable, and $EntryYear_i$ represents an entry year dummy variable. We denote by $PresentBiasDummy_k$ the k- th quintile group dummy based on $PresentBiasMasure$ for $k = 2, 3, 4, 5$

Table F.1 reports the estimation results of Cox regression models, where $\hat{\psi}$ from the former specification is in column 1, and $(\hat{\psi}_2, \hat{\psi}_3, \hat{\psi}_4, \hat{\psi}_5)$ from the latter specification are in column 2.

Table F.1: The relationship between patent examiners' present-biased preferences and their tendency to separate from the patent office

	(1)	(2)
(A) Parameter estimates		
$\hat{\psi}$	-0.885	
	[0.001]	
$\hat{\psi}_2$		0.048
		[0.922]
$\hat{\psi}_3$		-1.177
		[0.046]
$\hat{\psi}_4$		-1.704
		[0.044]
$\hat{\psi}_5$		-2.262
		[0.003]
(B) Measures of goodness-of-fit		
AIC	2.352	2.213
BIC	2.967	2.614
(C) No. of observations	269	269

Notes: This table reports the estimated coefficients of the Cox regression models (F.1) and (F.2). The p-values in brackets are calculated based on robust standard errors clustered at the art unit level. Akaike information criteria (AIC) and Bayesian information criteria (BIC) are given for each specification. The AIC and BIC are normalized by the number of observations.

Appendix G Appendix Tables and Figures

Table G.1: Definition of variables

variable name	type	no of dummies	granted data	description
art_unit	dummy	55		The art unit to which the examiner who issued the first office action on the patent application belongs, as a dummy variable.
uspc_class	dummy	31		The technology class to which the patent application belongs, as a dummy variable for United States Patent Classification(USPC) system.
year	dummy	9		The year in which the first office action was issued on the patent application, as a dummy variable.
qx_str	dummy	4		A dummy variable that takes 1if the week in which the first office action was issued on the patent application is the start of the x th quarter (for x=1,2,3,4).
qx_end	dummy	4		A dummy variable that takes 1 if the week in which the first office action was issued on the patent application is the end of the x th quarter (for x=1,2,3,4).
month	dummy	12		The month in which the first office action was issued on the patent application, as a dummy variable.
biweek	dummy	26		The biweekly period in which the first office action was issued on the patent application, as a dummy variable.
bw_thnks	dummy	1		A dummy variable that takes 1 if the biweekly period in which the first office action was issued on the patent application is the Thanksgiving holiday period.
exam_exp_year	numeric	30		The number of experience years of the examiner who issued the first office action for the patent application
exam_rank	dummy	1		A dummy variabe that takes 1 if the examiner who issued the first office action for the patent application is a primary (senior) examiner, and zero if she is an assistant (junior) examiner.
small_entity	dummy	1		A dummy variable that takes 1 if the patent application claims small-entity status.
foreign_priority	dummy	1		A dummy variable that takes 1 if the patent application claims foreign priority.
parent	dummy	1		A dummy variable that takes 1 if the patent application is a parent application. In other words, a related continuing application is filed.

Table G.1: Definition of variables (cont'd)

fam_root	dummy	1		A dummy variable that takes 1 if the patent application is a family root of any other patent applications.
cip	dummy	1		A dummy variable that takes 1 if the patent application is a continuation-in-part (CIP) application
con	dummy	1		A dummy variable that takes 1 if the patent application is a continuation (CON) application
div	dummy	1		A dummy variable that takes 1 if the patent application is a divisional (DIV) application
accel	dummy	1		A dummy variable that takes 1 if the patent application is granted accelerated examination status.
ctrs	dummy	1		A dummy variable that takes 1 if the patent application is requested restriction/election.
*bw_order*10	dummy	10		The order in which the first office action was issued in the biweekly period in which it was issued for the patent application, as a dummy variable up to the tenth (censored for the tenth or more).
log_bclog_new	numeric	-		The number of unexamined new patent applications (backlogs) at the beginning of the biweekly period in which the first office action was issued for the patent application (in logarithmic value).
log_bclog_amend	numeric	-		The number of unexamined amended patent applications (backlogs) at the beginning of the biweekly period in which the first office action was issued for the patent application (in logarithmic value).
ln_fwd_exa	numeric	-	x	The number of examiner's forward citations on the patent application (in logarithmic value).
ln_fwd_other	numeric	-	x	The number of non-examiner's forward citations on the patent application (in logarithmic value).
ln_nclm	numeric	-	x	The number of claims of the patent application (in logarithmic value).
ln_draws	numeric	-	x	The number of drawings in the patent application (in logarithmic value).
ln_wtypes_clm	numeric	-	x	The number of word types in the claims of the patent application (in logarithmic value).
ln_wtokens_clm	numeric	-	x	The number of word tokens in the claims of the patent application (in logarithmic value).
ln_wtypes_abs	numeric	-	x	The number of word types in the abstract of the patent application (in logarithmic value).
ln_wtokens_abs	numeric	-	x	The number of word tokens in the abstract of the patent application (in logarithmic value).

Notes: The sing 'x' in the "granted data" column represents the variable is included in the granted patent data.

Table G.2: Reduced-form regression evidence for present-bias-induced procrastination (cont'd)

	(1)	(2)	(3)	(4)	(5)	(6)
exam_rank	-0.174	-0.172	-0.176	-0.146	-0.249	-0.174
	[0.003]	[0.003]	[0.003]	[0.028]	[0.000]	[0.003]
small_entity	0.015	0.015	0.01	0.015	0.058	0.016
	[0.387]	[0.387]	[0.615]	[0.408]	[0.011]	[0.368]
foreign_priority	0.009	0.009	0.017	0.011	0.048	0.01
	[0.605]	[0.607]	[0.378]	[0.540]	[0.036]	[0.572]
parent	0.104	0.104	0.11	0.105	0.124	0.104
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
fam_root	-0.021	-0.021	-0.028	-0.024	-0.013	-0.02
	[0.543]	[0.542]	[0.461]	[0.501]	[0.769]	[0.557]
cip	0.049	0.049	0.039	0.051	0.038	0.049
	[0.036]	[0.036]	[0.131]	[0.032]	[0.206]	[0.034]
cont	-0.09	-0.09	-0.094	-0.089	-0.04	-0.09
	[0.000]	[0.000]	[0.000]	[0.000]	[0.151]	[0.000]
div	-0.074	-0.074	-0.058	-0.07	-0.184	-0.073
	[0.005]	[0.005]	[0.046]	[0.008]	[0.000]	[0.005]
accel	0.053	0.053	0.077	0.044	0.03	0.058
	[0.543]	[0.541]	[0.431]	[0.615]	[0.775]	[0.507]
ctrs	0.269	0.269	0.27	0.265	0.257	0.27
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
log_bclog_new	-0.001	-0.001	0	-0.013	0.009	0
	[0.974]	[0.972]	[0.991]	[0.566]	[0.751]	[0.997]
log_bclog_amend	-0.044	-0.044	-0.044	-0.06	0.034	-0.045
	[0.048]	[0.048]	[0.049]	[0.009]	[0.233]	[0.042]
ln_fwd_exa					-0.079	
					[0.001]	
ln_fwd_other					-0.076	
					[0.000]	
ln_nclm					0.193	
					[0.000]	
ln_ndrw					-0.04	
					[0.007]	
ln_wtypes_clm					0.125	
					[0.026]	
ln_wtokens_clm					-0.003	
					[0.944]	
ln_wtypes_abs					0.05	
					[0.498]	
ln_wtokens_abs					-0.024	
					[0.657]	

Notes: This table reports the supplementary estimation results for the regression model (2). The estimated coefficients of the covariates on the first office action failure are reported. The p-values in brackets are calculated based on robust standard errors clustered at the art unit level.

Table G.3: Estimated standard deviations and correlation matrix of the parameters for the regression model with the lead and lag effects of the shortened deadline

σ_0	σ_1	σ_2	σ_3
0.696 [0.000]	0.204 [0.000]	0.210 [0.000]	0.126 [0.004]
σ_{01}	σ_{02}	σ_{03}	
-0.343 [0.001]	-0.307 [0.002]	-0.281 [0.114]	
σ_{12}	σ_{13}	σ_{23}	
0.676 [0.002]	0.556 [0.124]	0.616 [0.089]	

Notes: This table reports the estimated standard deviations and correlation matrix of the regression model (3). The p-values in brackets are calculated based on robust standard errors clustered at the art unit level.

Table G.4: The estimation results for the key parameters of the present-bias-induced procrastination model under wider prior distributions

Priors	$\beta \in [0.1, 1.1]$ $\delta \in [0.1, 1.0]$ $\tau \in (0.0, 0.2]$	$\beta \in [0.1, 1.2]$ $\delta \in [0.9, 1.0]$ $\tau \in (0.0, 0.2]$	$\beta \in [0.1, 1.2]$ $\delta \in [0.9, 1.0]$ $\tau \in (0.0, 0.3]$	$\beta \in [0.1, 1.2]$ $\delta \in [0.1, 1.0]$ $\tau \in (0.0, 0.3]$
(A) Marginal posterior means				
Present bias factors $\bar{\hat{\beta}}_i$				
1st Qt.	0.432	0.469	0.434	0.429
	(0.312, 0.570)	(0.312, 0.661)	(0.368, 0.509)	(0.306, 0.582)
2nd Qt.	0.599	0.643	0.613	0.601
	(0.495, 0.714)	(0.325, 1.090)	(0.502, 0.739)	(0.435, 0.793)
3rd Qt.	0.792	0.847	0.797	0.789
	(0.574, 1.045)	(0.658, 1.096)	(0.652, 0.964)	(0.601, 1.037)
Discount factor $\bar{\hat{\delta}}$	0.321	0.948	0.949	0.319
	(0.104, 0.904)	(0.902, 0.997)	(0.902, 0.997)	(0.104, 0.902)
Admissible failure rate $\bar{\hat{\tau}}$	0.044	0.049	0.044	0.044
	(0.041, 0.046)	(0.046, 0.052)	(0.041, 0.047)	(0.041, 0.046)
(B) Measures of goodness-of-fit				
WAIC	0.756	0.756	0.756	0.756
AUC.ROC	0.692	0.692	0.692	0.692
(C) Present-biased patent examiners (%)	70.000	63.239	69.296	69.155

Notes: This table reports the results of the Bayesian estimation. The means of the marginal posterior distributions are presented for the key parameters. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. We use (i) either $[0.1, 1.0]$ or $[0.9, 1.0]$ as the support of the uniform prior for the discount factor δ; (ii) $[0.1, 1.1]$ or $[0.1, 1.2]$ as the support of the uniform prior for the present bias factor β, and (iii) $(0.0, 0, 2]$ or $(0.0, 0.3]$ as the support of the uniform prior for the admissible failure rate τ.

Table G.5: The estimation results for the key parameters of the present-bias-induced procrastination model under various utility curvature parameters

	$\gamma = 0.0$	$\gamma = 0.5$	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 2.0$	$\gamma = 2.5$	$\gamma = 3.0$	$\gamma = 4.0$	$\gamma = 5.0$	$\gamma = 6.0$
(A) Marginal posterior means										
Present bias factors $\bar{\bar{\beta}}_i$										
1st Qt.	0.353	0.420	0.430	0.422	0.407	0.386	0.358	0.322	0.295	0.284
	((0.286, 0.435)	(0.308, 0.552)	(0.282, 0.619)	(0.333, 0.529)	(0.316, 0.517)	(0.322, 0.462)	(0.274, 0.468)	(0.277, 0.375)	(0.247, 0.351)	(0.145, 0.529)
2nd Qt.	0.556	0.604	0.600	0.595	0.585	0.576	0.55	0.543	0.546	0.551
	(0.423, 0.715)	(0.307, 0.984)	(0.495, 0.716)	(0.410, 0.842)	(0.463, 0.748)	(0.430, 0.777)	(0.310, 0.952)	(0.373, 0.797)	(0.234, 0.998)	(0.280, 1.011)
3rd Qt.	0.790	0.791	0.792	0.797	0.802	0.799	0.786	0.79	0.805	0.803
	(0.395, 1.085)	(0.654, 0.951)	(0.642, 0.966)	(0.637, 0.996)	(0.584, 1.053)	(0.439, 1.084)	(0.538, 1.056)	(0.364, 1.087)	(0.484, 1.085)	(0.501, 1.076)
Discount factor $\bar{\bar{\delta}}$	0.950	0.949	0.949	0.949	0.949	0.950	0.95	0.95	0.95	0.95
	(0.903, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.998)	(0.903, 0.997)	(0.902, 0.998)	(0.903, 0.997)	(0.902, 0.998)
Admissible failure rate $\bar{\bar{\tau}}$	0.055	0.042	0.044	0.047	0.050	0.052	0.053	0.057	0.061	0.066
	(0.053, 0.057)	(0.039, 0.044)	(0.040, 0.047)	(0.044, 0.049)	(0.047, 0.052)	(0.049, 0.054)	(0.050, 0.055)	(0.055, 0.059)	(0.059, 0.063)	(0.065, 0.068)
(B) Measures of goodness-of-fit										
WAIC	0.756	0.756	0.756	0.756	0.756	0.756	0.756	0.756	0.757	0.758
AUC.ROC	0.692	0.692	0.692	0.692	0.692	0.692	0.692	0.691	0.691	0.689
(C) Present-biased (%)	62.113	71.268	69.014	65.775	63.380	62.958	62.676	59.437	56.901	55.493

Notes: This table reports the results of the Bayesian estimation. The means of the marginal posterior distributions are presented for the key parameters. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. We change the relative risk aversion coefficient γ from 0.0 to 6.0.

Table G.6: The estimation results for the key parameters of the present-bias-induced procrastination model under different beliefs of the patent office on the patent examiner's present-biased preferences

	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.3$	$\beta = 0.4$	$\beta = 0.5$	$\beta = 0.6$	$\beta = 0.7$	$\beta = 0.8$	$\beta = 0.9$
(A) Marginal posterior means									
Present bias factors $\bar{\bar{\beta}}_i$									
1st Qt.	0.140	0.252	0.354	0.444	0.522	0.456	0.442	0.451	0.465
	(0.110, 0.155)	(0.179, 0.343)	(0.268, 0.450)	(0.362, 0.525)	(0.419, 0.623)	(0.320, 0.599)	(0.381, 0.501)	(0.372, 0.536)	(0.377, 0.557)
2nd Qt.	0.262	0.362	0.476	0.572	0.643	0.577	0.573	0.583	0.603
	(0.248, 0.275)	(0.289, 0.430)	(0.426, 0.527)	(0.509, 0.637)	(0.578, 0.709)	(0.503, 0.647)	(0.376, 0.772)	(0.489, 0.680)	(0.317, 0.930)
3rd Qt.	0.596	0.694	0.768	0.806	0.828	0.695	0.694	0.708	0.732
	(0.250, 0.946)	(0.297, 1.075)	(0.479, 1.081)	(0.505, 1.084)	(0.486, 1.085)	(0.600, 0.782)	(0.476, 0.904)	(0.627, 0.789)	(0.567, 0.903)
Discount factor $\bar{\bar{\delta}}$	0.950	0.950	0.950	0.949	0.949	0.949	0.949	0.949	0.949
	(0.902, 0.998)	(0.903, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)
Admissible failure rate $\bar{\bar{\tau}}$	0.200	0.200	0.199	0.197	0.179	0.100	0.067	0.047	0.033
	(0.200, 0.200)	(0.199, 0.200)	(0.197, 0.200)	(0.192, 0.200)	(0.171, 0.187)	(0.085, 0.111)	(0.061, 0.074)	(0.042, 0.052)	(0.030, 0.037)
(B) Measures of goodness-of-fit									
WAIC	1.411	1.329	1.365	1.209	1.031	0.870	0.785	0.756	0.767
AUC.ROC	0.636	0.665	0.683	0.687	0.689	0.691	0.692	0.692	0.692
(C) Present-biased (%)	95.211	59.859	57.324	55.493	59.577	83.944	88.169	87.465	85.493

Notes: This table reports the results of the Bayesian estimation. The means of the marginal posterior distributions are presented for the key parameters. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. We assume that the patent examiner's present-bias factor that the patent office believes when determining the reward varies from 0.1 to 0.9.

Table G.7: The effect of the present bias measure on the PTA period using a tobit regression model (cont'd)

	(1)	(2)	(3)	(4)
small_entityr	-23.493	-24.497	-19.098	-11.792
	[0.000]	[0.000]	[0.000]	[0.003]
parent	-21.889	-23.635	-18.004	-35.975
	[0.000]	[0.000]	[0.000]	[0.000]
cip	55.858	55.153	44.796	-1.612
	[0.000]	[0.000]	[0.000]	[0.818]
cont	-111.822	-110.447	-111.886	-295.952
	[0.000]	[0.000]	[0.000]	[0.000]
div	-124.835	-121.432	-123.064	-278.093
	[0.000]	[0.000]	[0.000]	[0.000]
foreign_priority	0.113	-0.471	2.616	15.954
	[0.975]	[0.894]	[0.444]	[0.000]
ctrs	139.682	134.095	151.665	176.292
	[0.000]	[0.000]	[0.000]	[0.000]
accel	-289.357	-289.21	-293.836	-491.57
	[0.000]	[0.000]	[0.000]	[0.000]
fam_root	48.76	49.041	34.742	7.507
	[0.000]	[0.000]	[0.000]	[0.462]

Notes: This table provides the supplementary contents for the marginal effects on the PTA period of a patent application based on the regression model (8). The p-values in brackets are calculated based on robust standard errors clustered at the art unit level.

Table G.8: The estimation results for the key parameters of the present-bias-induced procrastination model using a sample of TC 2100

	$\gamma = 0$	$\gamma = 0.5$	$\gamma = 1.0$	$\gamma = 1.5$	$\gamma = 2.0$
(A) Marginal posterior means					
Present bias factors $\bar{\hat{\beta}}_i$					
1st Qt.	0.359	0.405	0.408	0.401	0.392
	(0.148, 0.765)	(0.141, 0.860)	(0.271, 0.577)	(0.283, 0.554)	(0.304, 0.501)
2nd Qt.	0.602	0.605	0.601	0.605	0.606
	(0.413, 0.818)	(0.329, 0.949)	(0.310, 0.986)	(0.435, 0.818)	(0.246, 1.060)
3rd Qt.	0.766	0.779	0.78	0.777	0.776
	(0.385, 1.081)	(0.599, 0.991)	(0.338, 1.081)	(0.496, 1.076)	(0.287, 1.087)
Discount factor $\bar{\hat{\delta}}$	0.951	0.949	0.949	0.949	0.949
	(0.902, 0.998)	(0.903, 0.997)	(0.902, 0.997)	(0.902, 0.997)	(0.902, 0.997)
Admissible failure rate $\bar{\hat{\tau}}$	0.041	0.03	0.03	0.032	0.035
	(0.039, 0.043)	(0.029, 0.032)	(0.028, 0.032)	(0.030, 0.034)	(0.032, 0.037)
(B) Measures of goodness-of-fit					
WAIC	0.661	0.660	0.660	0.661	0.661
AUC.ROC	0.715	0.717	0.717	0.717	0.716
(C) Present-biased patent examiners (%)	49.026	58.658	58.009	54.437	53.139

Notes: This table reports the replication results of the Bayesian estimation for the model parameters using the sample of patent applications reviewed by patent examiners who belong to TC2100 (Computer Architecture and Software). The means of the marginal posterior distributions are presented for the key parameters. For the individual-specific present bias factors, the first to third quartiles of the marginal posterior means' distribution are presented. The values in parentheses represent the lower and upper limits of the 95 percent highest density interval (HDI) of the marginal posterior distribution. A patent examiner is defined as present biased if the upper limit of 95 percent HDI of her present-bias factor does not reach the value of 1.0. As goodness-fit-measures, we report the widely applicable information criterion (WAIC), which is normalized by the number of observations, and the area under the ROC curve (AU.ROC) from the receiver operating characteristics (ROC) analysis.

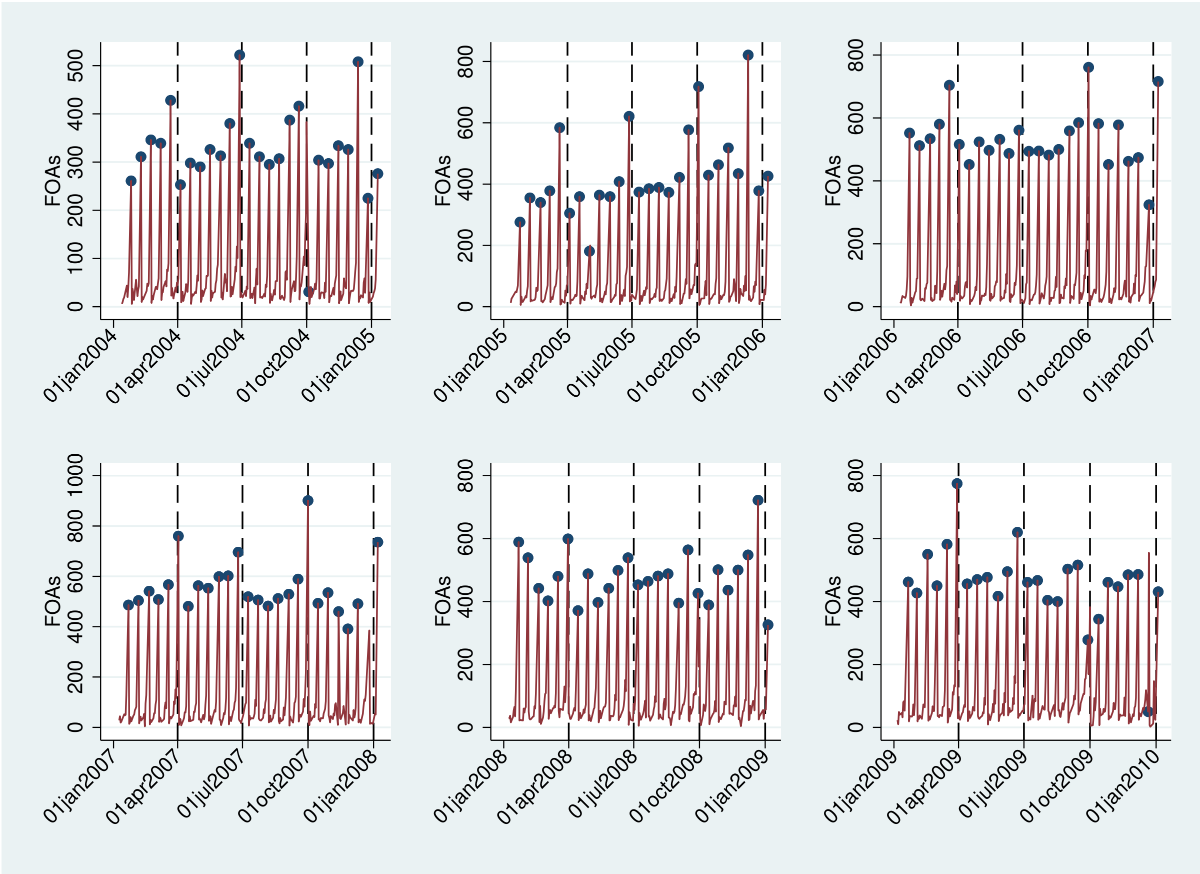

Figure G.1: The daily number of first office actions for TC1600 for the period 2004-2009. The solid circle represents the deadline day of a biweekly quota period. The dashed line represents the end of the quarterly period.

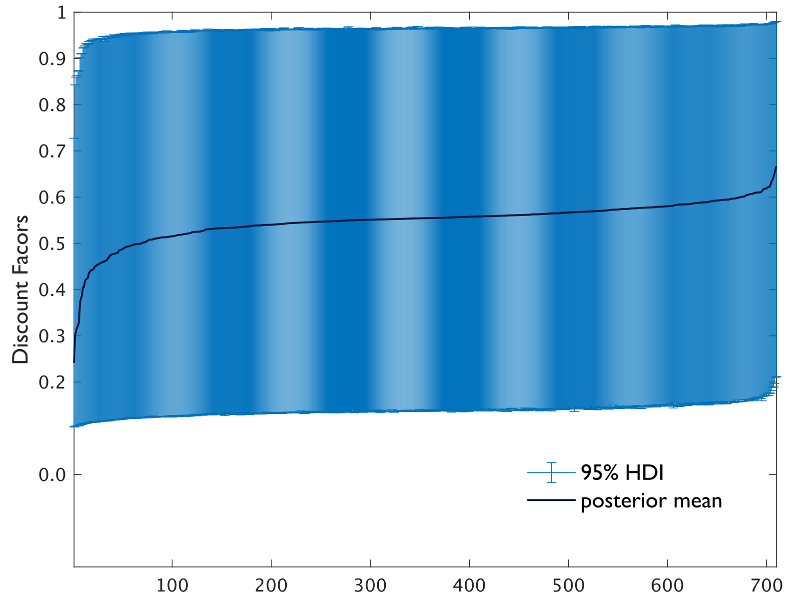

Figure G.2: Estimated marginal posterior distributions of the discount factor for patent examiners. The estimation is based on the exponential discounting model. The bar represents the 95% HDI of the marginal posterior distribution of a patent examiner's annual discount factor.

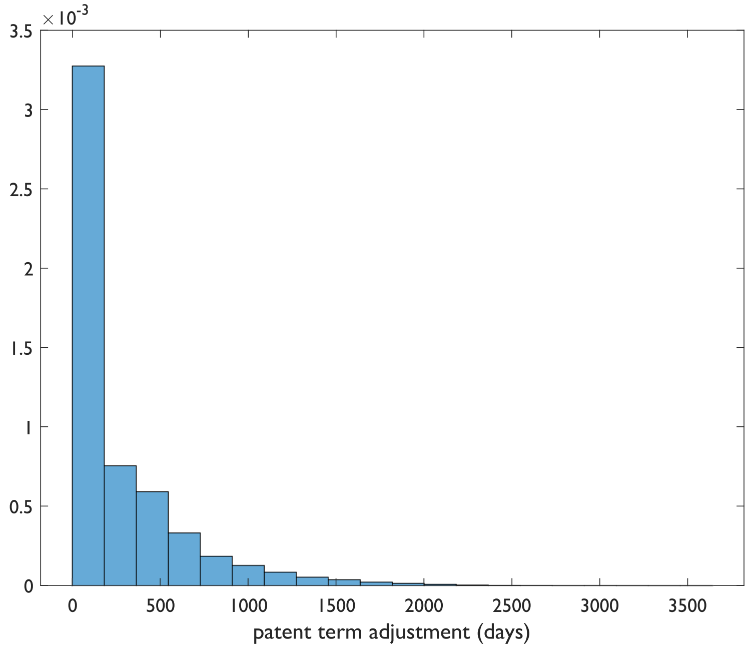

Figure G.3: A histogram of the PTA period. The sample consists of applications that were granted patents because unapproved patent applications do not have PTA records.