Playing the Patent Lottery

Cameron M. Ellis*, Johannes G. Jaspersen† & Valentin Luz‡

June 5, 2026

Abstract

We investigate whether the U.S. patent examiner lottery, which is intended to ensure random assignment and equal treatment of applicants, can be strategically manipulated. Using data on six million applications from 2001–2020, we derive the null distribution of examiner leniency under true randomization and show that many law firms and in-house counsel systematically obtain far more lenient examiners than chance would permit. We document multiple mechanisms consistent with strategic manipulation, including leveraging institutional knowledge and the exploitation of publicly observable examiner workloads. These deviations from randomness have economically meaningful consequences: sorting firms on examiner leniency at publication produces sizable and statistically significant return spreads, implying that markets only partially incorporate assignment-based variation in approval likelihood. We develop a simple model of manipulation investments that explains which firms are most likely to game the patent lottery, highlight potential welfare consequences, and show empirical support for its predictions. Our findings challenge the fairness of the patent lottery and call into question empirical designs that rely on examiner randomization for causal identification.

JEL Codes: **032** · **034** · **L51** · **G14**

Keywords: Patents · Innovation · Leniency Designs · Law Firms

*Tippie College of Business, University of Iowa (cameron-ellis@uiowa.edu)
†LMU Munich School of Management, LMU Munich (jaspersen@lmu.de)
‡LMU Munich School of Management, LMU Munich (luz@lmu.de)

1 Introduction

The patent system exists because competitive markets do not properly incentivize private investment into research (Budish et al., 2016). To reward firms for their innovation, and to share the idea of the innovation itself with the public, firms are granted exclusive monopoly power over their inventions for a limited time. This monopoly power is extremely valuable, with even "marginal" innovations generating substantial economic rents (Kogan et al., 2017; Farre-Mensa et al., 2020). To ensure that this value is allocated to true innovation, patents are evaluated for novelty and non-obviousness by patent examiners. This evaluation process is inherently noisy, reflecting heterogeneity in examiners' expertise and willingness to grant patents. Ensuring a level playing field therefore requires that applications be assigned to examiners as good as randomly, a process commonly referred to as the patent "lottery."

In this paper we ask, and answer, four questions about the patent lottery. (1) Can it be gamed? We derive the theoretical distribution of examiner leniency under random assignment and show that certain law firms and in-house counsel receive lenient examiners at a rate far beyond what could be attributed to chance. (2) How do you game the patent lottery? We identify two mechanisms through which sophisticated applicants can influence examiner assignment: strategic timing and leveraging institutional knowledge. (3) When can we expect companies to game the patent lottery? We adapt the models of Nordhaus (1969) and Budish et al. (2016) and show that firms with lower cost of capital and larger potential gains from a patent are most likely to manipulate the system. Both predictions are supported by empirical evidence on the patent level when using relevant proxies. (4) Is there value in getting assigned a more lenient patent examiner? By implementing a trading strategy that exploits variation in examiner leniency at the publication date, we show that investors can earn sizable and statistically significant abnormal returns.

As is common in this setting, we do not directly observe manipulative behavior by any individual firm. Instead, we document a set of statistical patterns that are each consistent with manipulation and that are difficult to reconcile jointly with benign alternatives. Any single result in isolation may potentially admit another explanation. For example, entity-level differences in assigned examiner leniency could reflect residual specialization at a finer technological level than our fixed effects capture. Our empirical strategy therefore combines several distinct tests across different data margins. We show not only that certain counseling entities systematically receive more lenient examiners than expected under random assignment, but also that these entities submit patent applications at days when particularly lenient examiners are available to be assigned. Combined with evidence that hiring ex-USPTO staffers and having applications for high-value patents increases the assigned examiner leniency while high cost of capital decreases it, we provide convergent evidence that examiner assignment is at least partly driven by strategic behavior of sophisticated counseling entities.

The USPTO receives around 500,000 patent applications annually and employs approximately 8,000 examiners to evaluate them. Each examiner specializes in a specific technology area—called an "art unit"—and within each art unit, applications are distributed to examiners through an assignment system. This system, which is intended to eliminate favoritism and ensure fairness, operates as a lottery: applications are assigned based on examiner availability and relevant technical expertise.[1] Whom an application gets assigned to matters. Examiners differ substantially in their grant rates, with the median art unit having a 40 percentage point difference between the strictest and most lenient examiners. This variation in examiner leniency, combined with the high value of patent rights, creates powerful incentives to try to subvert the randomization process at the core of the patent lottery.

Can the patent lottery be gamed? To answer this question, we construct the null distribution of examiner leniency under true randomization. Using data on 6 million patent applications from 2001-2020, we calculate each examiner's grant rate relative to their annual art unit average. Under random assignment, the distribution of examiner leniency received by any entity providing legal counsel should converge to the population distribution as the number of applications increases. We analyze 1,427 entities with at least 1,000 applications, of which 983 are law firms. We find that the empirical distribution of entity fixed effects differs significantly from what would be expected under true randomization. Many entities show patterns of examiner assignment that cannot be reconciled with random allocation, translating to substantial advantages in patent approval rates.

How do you game the patent lottery? We demonstrate that strategic timing of patent submissions can predict favorable examiner assignment. Using public information on recently disclosed examiner outcomes and workload signals, we construct a set of timing predictors at the art-unit level and estimate rolling out-of-sample prediction models. The prediction exercise shows that realized examiner leniency is predictable from these pre-filing signals: out-of-sample R^2 is positive for 78% of all art units and for 96% of large art units. This suggests that sophisticated applicants with access to public USPTO data on examiner grant rates and workloads can identify favorable filing windows. We then provide evidence that higher-ranked entities make use of this channel by submitting more applications when predicted examiner leniency is high. It is also possible that firms leverage insider knowledge on the application assignment process. We find evidence

[1] This system which has recently been changed to an automated, computerized process is described in more detail in Section 2.

for this hypothesis by analyzing the revolving door between the USPTO and the private patent bar (Tabakovic and Wollmann, 2024). Here we show that hiring a patent attorney who has previous work experience with the USPTO increases the expected leniency for subsequently submitted patent applications. A back-of-the-envelope calculation suggests that, for the average firm in this sample, the first such hire is worth approximately $12M per year in additional patent grants after accounting for the zero-sum nature of examiner capacity. Former USPTO employees apparently provide a channel of influencing the patent lottery favorably.

When can we expect companies to game the patent lottery? To address this question, we adapt the theoretical model by Nordhaus (1969) and Budish et al. (2016) to introduce heterogeneous examiner quality and a costly manipulation technology. Firms compete over lenient examiners using manipulation investments. Firms' investments are determined by their idiosyncratic circumstances. Based on the unique Nash equilibrium solution to the model, we derive two empirically testable predictions. First, firms with lower marginal cost of capital will invest more to manipulate the patent lottery. Because larger firms often have lower cost of capital, this implies that smaller firms will spend less money on manipulation and thus receive examiners with a lower leniency. Second, firms with higher value patent applications have more to gain from a positive grant decision and thus invest more into manipulation. Our empirical analysis on patent-level data reveals support for these claims: small entities receive examiners who are 1.13 percentage points less lenient on average, while applications with more claims—a proxy for patent value—systematically receive more lenient examiners.

Is there value in getting assigned a more lenient patent examiner? We examine whether an investor could earn abnormal returns by trading on information revealed during the patent application process. We find that sorting firms into portfolios on the day after patent-application publication based on examiner leniency yields a pronounced return spread across value-weighted portfolios, and a long–short strategy delivers positive and statistically significant alphas of around 35 basis points per month. This pattern implies that an investor who conditions on this information could earn economically meaningful abnormal returns by following such a strategy. At the same time, this finding highlights the economic value inherent in getting assigned a more lenient examiner.

We discuss the welfare implications of our findings using our theoretical model as a basis. We show that welfare consequences are non-trivial. If manipulation does not induce or deter innovation, it likely detriments social welfare due to the resource waste from manipulation efforts. If, however, the ability to manipulate induces innovation for some firms because it increases the probability of obtaining a patent, then it might be welfare increasing. We provide a discussion of the likelihood of the individual scenarios as well as an argument why the fairness of the process itself might influence social welfare irrespective of its efficiency consequences.

Our findings also have important implications for existing empirical research on patents and innovation. A widely-used identification strategy in the literature relies on examiner leniency as an instrumental variable for patent grants, assuming that examiner assignment is random (Lemley and Sampat, 2012; Sampat and Williams, 2019).[2] This approach has been employed to study the causal effects of patents on topics such as follow-on innovation (Sampat and Williams, 2019), firm performance and entrepreneurial outcomes (Farre-Mensa et al., 2020), inventor mobility (Melero et al., 2020), and various other economic outcomes. We revisit this approach and demonstrate both analytically and in Monte Carlo simulations that it may produce spurious results. The problem arises because unobserved patent quality influences both examiner assignment and innovation outcomes. Even small correlations between quality and examiner leniency, precisely what we document when the lottery is gamed, can generate false positives in two-stage least squares estimations. This calls into question not only the fairness of the patent system but also the validity of a research design that has been employed across dozens of studies examining innovation spillovers, patent value, and firm performance.

This paper makes three contributions to the literature on patent systems and institutional design. First, we provide the first systematic evidence that the patent examiner lottery—long assumed to ensure equal treatment—can be gamed by sophisticated players. We show the assumption of random assignment of examiners fails for a substantial, non-random subset of applicants. Considered jointly, our diverse set of statistical tests reveal patterns impossible to reconcile with true randomization, calling into question a fundamental premise of patent system fairness. These findings highlight the role of law firms and in-house counselors. As of January 2024, there were 1.32 million lawyers in the United States, constituting slightly less than 1 percent of the US labor force (American Bar Association, 2024). Yet, with few exceptions (e.g., Briscoe and Rogan, 2016; Azmat and Ferrer, 2017; Henderson et al., 2025), lawyers and law firms are empirically understudied in finance and economics research.[3] We show that the choice of law firm (or the quality of the in-house counsel) has real economic consequences and has the potential to create significant value for innovators.

[2]The patent lottery is a canonical example in guides for using leniency designs (e.g., Goldsmith-Pinkham et al., 2025).

[3]This is in contrast to the extensive literature on the role of judges (e.g., Rehavi and Starr, 2014; Arnold et al., 2018; Cohen and Yang, 2019). In parallel to what we analyze for patent examiners here, this literature has documented cases of so-called "judge shopping" where parties strategically manipulate case assignments to gain an advantage in litigation (Botoman, 2017; Kahan and McKenzie, 2021).

Second, we develop a theory of when gaming randomized government programs becomes profitable. While the industrial organization literature has extensively studied regulatory capture and lobbying, less attention has focused on manipulation of supposedly random allocation mechanisms. Our model shows that even small advantages in examiner assignment can generate large returns when patent values are high and competition is concentrated. The framework applies beyond patents to any setting where government randomly allocates valuable rights or assigns cases to heterogeneous decision-makers, including immigration courts, disability determinations, and regulatory approvals. While most theoretical work on the patent system has focused on constructing institutions to maximize societal welfare under an assumption of no manipulation, a separate literature examines opportunities for exploitation, commonly through so-called "patent trolls" (e.g., Cohen et al., 2019; Appel et al., 2019; Dayani, 2023). We add to this literature by identifying a new avenue for exploiting inefficiencies in the system: strategic manipulation of examiner assignment within the patent lottery. We also show that welfare effects due to strategic manipulation can be ambiguous and depend on the specific application.

Third, our findings have immediate policy implications for patent system integrity and innovation incentives. Gaming the examiner lottery redistributes innovative rewards from small inventors to large corporations, potentially reducing the incentive for breakthrough innovation by startups and independent inventors. Our evidence shows that small entities systematically receive less lenient examiners, while high-value patents (those with more claims) receive more lenient treatment. Such patterns advantage sophisticated repeat players over individual inventors. Simple fixes, such as better randomization in the automated assignment system, regular audits of assignment patterns, and restrictions on examiner-applicant communications, could restore fairness at minimal cost. More broadly, our results suggest that any high-stakes government lottery requires active monitoring to prevent sophisticated actors from subverting randomization.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background and data. Section 3 examines whether the patent examiner lottery can be gamed. Section 4 investigates the mechanisms underlying such manipulation, while Section 5 analyzes which firms are most likely to engage in it. Section 6 evaluates the economic value of examiner leniency, and Section 7 discusses implications for social welfare and empirical identification. Section 8 concludes.

2 Institutional Background and Data

2.1 The Patent Application Process

Patents on mechanical, electronic, and chemical technologies are called *utility patents*. Such patents consist of a number of *claims*, which define the scope of the asserted invention. After submission, the application is given a technology classification and based on this classification is assigned to an *art unit*. These art units are organized by technology type of the patent application, each covering clusters of related technology subject-matter. By virtue of differences in technologies, art units differ in their size, the length of the review process and the rate of granted patents. The technological classification as well as the art unit assignment depends on the language used in an application's claims and specification. This allows for strategic use of specific language in applications with the goal of steering art unit assignment. Such behavior has limits, however, because an art unit assignment can be changed in the examination and search process after the initial assignment.

Within the art group, the applications are then assigned to individual patent examiners. How this assignment is handled has changed recently, but is constant throughout our observation period. The procedure is codified in the Manual of Patent Examining Procedure (MPEP). The eighth edition of the MPEP from 2001 puts the assignment under the purview of the examiner supervisors, allowing them relatively free discretion in the assignment. This led to a debate in the literature on whether the assignment is random or which of the characteristics of the patent application are used in the assignment process. Lemley and Sampat (2012) and later Sampat and Williams (2019) argue that the assignment is more or less random, conditional on the current availability of examiners. They base this on their own conversations with examiner supervisors. While Lemley and Sampat (2012) concede that there may be some systematic aspect to the assignment process, they state that supervisors only take examiner familiarity with a technology into account, not the quality of an application.[4]

A somewhat different opinion is taken by Righi and Simcoe (2019) who argue that supervisors pay more attention to examiner technical knowledge and specialization. They find some evidence supporting this, but only in a limited selection of art units. However, it is clear that even if the individual examiner's knowledge base is taken into account, assignment has to take availability into account, as well.[5] They further show that proxies for patent quality and scope as well as the

[4]See also Merges (1999) who argues that there is a tradition in the USPTO to treat all patents as equal.

[5]The most current version of the MPEP, instituted in February 2023, changes the assignment procedure and relegates it to an automated system. This system considers both examiner specialization and availability in the assignment, strengthening the hypothesis that both factors played a roll in the previous system, as well.

identity of the applicant are not randomly distributed across examiners. This is in line with our own results. However, while they claim that "there is no evidence that the broadest or most important applications are assigned to specific examiners" (p. 138), we show evidence to the contrary in Section 5.2.

After assignment to an examiner, the examiners starts a process of search and examination. Based on this process, if they deem the patent sufficient, it is granted. If not, the patent is rejected initially and the applicant, often represented through an attorney or agent, starts a negotiation process with the examiner through a series of submitted documents. This process can end with a granted patent that includes a reduced number of claims or abandonment of the application by the applicant.

The examiners have more or less sole discretion on granting or rejecting the patent application.[6] They work within relatively strict guidelines and are, in fact urged to grant the patent unless they can find a reason for refusal (Graham et al., 2018). Nevertheless, examiners differ quite substantially in their tendency to grant patents. This has lead to the patent system sometimes being called the patent *lottery* (Farre-Mensa et al., 2020). The argument is that the assignment of the patent examiner is decisively influential for the probability of getting a patent granted. Having a lenient examiner makes a grant more likely and constitutes a "win" in the lottery.

2.2 Data

Our main source of data is the USPTO Patent Examination Research Dataset (PatEx) database provided directly from the USPTO (for details see Graham et al., 2018). We use the 2022 release of these data that covers all patent applications up to the year 2022 inclusively. The data were drawn from the Public Patent Application Information Retrieval (PublicPAIR) system, a system that was put into place following the US Congress passing the American Inventors Protection Act (AIPA) in 2000.[7] AIPA is relevant because its passing drastically increased the number of observable patents and, crucially, made information about abandoned patent applications public starting November 2000. PatEx thus covers (almost) all patent applications from the year 2001 until 2022 inclusively.[8] The data does not cover those applications for which the applicants sought the opt-out provision

[6] In theory, the supervising patent can overrule decisions. In practice, this happens very rarely due to workloads and strategic behavior by the examiners (Tabakovic and Wollmann, 2024).

[7] PublicPAIR has since been replaced by the USPTO Patent Center. For our data, however, it is the relevant system as PublicPAIR was retired in July 2022 and all data used in the analyses below stems from the period before its retirement.

[8] An additional data source that we use is PatentsView which also includes data on patent applications and extends to more recent years. However, it does not list the assigned examiner before a patent is granted, thus we use PatEx for our primary analysis and merge in additional data from PatentsView.

under AIPA. How exactly such exceptions should be evaluated is discussed in Graham et al. (2018) and they constitute a inevitable omission for the data in our case (see e.g., Raffiee et al., 2023, for an analysis with the same data restrictions).

The initial data from PatEx contains information on slightly over 14 million applications. The data from the 2022 version of PatEx does not track a unique patent examiner ID, but merely identifies the examiner by name. The 2014 version of Patex was the last release that included a unique examiner ID. We use this release to connect the 2022 data to examiner IDs and then use the examiner name to associate the examiner names of the newer applications with an examiner ID.

The next step in our data preparation involves linking patent applications to the legal entities that represent them. We begin by using USPTO application-level practitioner records from PatEx to identify the attorneys and agents associated with each application. These practitioners are then matched via their USPTO registration numbers to historical practitioner roster records maintained by the USPTO.[9] Because practitioner affiliations change over time, we assign each practitioner to the counsel entity observed closest to, but preceding, the application filing date; when no earlier observation is available, we use the closest subsequent record. We then aggregate these practitioner-level matches to the application level by counting how many linked practitioners map to the same entity and retaining the top three entities for each application. This procedure is intended to capture the entities most relevant to the prosecution process, while any fourth or lower-ranked entity is likely to make only a marginal contribution.

The resulting entity names contain substantial variation in how firms are recorded—a single law firm may appear under dozens of different text representations due to punctuation differences, abbreviations, typos, OCR errors, and individual attorney listings. To address this challenge, we developed a multi-stage disambiguation pipeline that combines deterministic text normalization, fuzzy matching with inverse-document-frequency weighting, large language model assistance (GPT-5-mini) for typo correction and ambiguous-case classification, and manual overrides for known entities. A second language-model pass audits every proposed merge before it is finalized. This procedure links over 6 million applications to identified counsel entities while maintaining high confidence in the matches. Full details are provided in Appendix E.

We then make a series of data selection steps. We only keep data with a valid filing date after 2000 and before 2021. With the former restriction, we consider the time period after AIPA and with the latter, we allow patent applications at least two years to obtain a grant decision. We further only consider data with a valid examiner name and art unit and focus on utility patents. After applying these restrictions and requiring successful law firm linkage as described above,

[9]See https://www.uspto.gov/learning-and-resources/patent-and-trademark-practitioners/historical-rosters

Table 1: Data Selection Procedure

Step	Description	Applications
0	Raw application data (PatEx)	14,100,378
1	Keep data with valid filing date	14,043,233
2	Keep applications filed after 2000	10,775,875
3	Keep data with valid examiner name	8,738,337
4	Keep data with valid art unit	7,995,530
5	Merge with attorney data, remove applications without law firm record	6,722,923
6	Keep only utility patents	6,258,185
7	Keep applications filed before 2021	6,047,474
Base Sample		**6,047,474**

Note: The table describes the steps in the data selection procedure and indicates the number of applications retained after every step.

we arrive at a base sample of 6,047,474 patent applications. The full data selection procedure is summarized in Table 1.

2.3 Examiner Leniency

The key variable that we analyze is the leniency for patent applications exhibited by the patent examiner. The advantage of this variable is its supposed randomness, conditional on art unit and potentially other patent characteristics. This implies that, after controlling for art unit and other relevant characteristics, any statistical significance of applicant or agent characteristics is indicative of non-randomness.

For every observation in the data, we calculate the average leniency as a leave-out instrument. That is, we consider the average leniency of an examiner for all applications assigned to them in a given filing year except for the current observation. The leniency of application i, analyzed by examiner e and filed in year t is calculated as

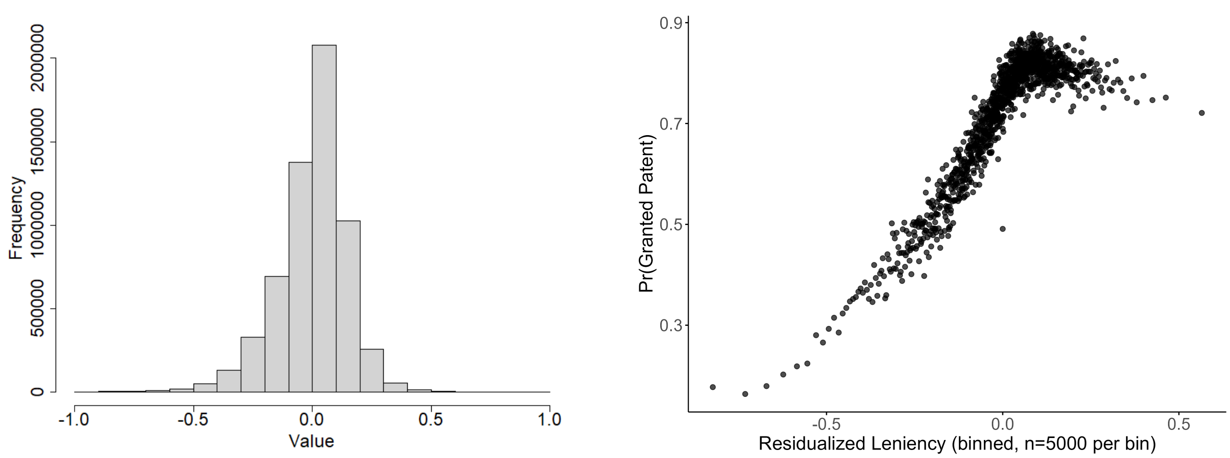
$$\text{Leniency}_{i,e,t} = \frac{\sum_{j=1}^{n_{e,t}} \mathbf{1}(\text{Success}_{j,e,t}) - \mathbf{1}(\text{Success}_{i,e,t})}{n_{e,t} - 1}. \tag{1}$$

Here, $n_{e,t}$ is the number of patents examined by examiner e in year t and $\mathbf{1}(\text{Success})$ is an indicator function for whether an application is ultimately successful. Calculating the examiner leniency in this way removes data from examiners who only handled a single application. This likely includes data entries with misspelled examiner names. We can calculate examiner leniency based on around 6 million applications.

There are some notable aspects to the definition given in (1). First, we calculate the leniency for each examiner on an annual basis. This allows examiners to change their idiosyncratic leniency over time. Secondly, we calculate the examiners' annual leniency as an average over all art units the examiner operates in. Examiners operate in more than one art unit for slightly less than 40% of all examiner-year observations. In these cases, we make the assumption that the idiosyncratic leniency of the examiner does not change between art units. Our analyses nevertheless allow for an art-unit-level effect on the leniency. We simply assume that this effect is not examiner specific. To see the variation induced by the examiner leniency, we plot a histogram of the leniency for all patent applications considered in our main analysis in panel (a) of Figure 1. Data in the histogram is residualized by art-unit $\times$ year fixed effects. Even absent these effects, examiners vary significantly in their leniency. Moreover, this variation matters. Panel (b) plots the average success probability for binned values of examiner leniency. There is a clear positive relationship between both variables.

Figure 1: Examiner Leniency: Variation and Prediction Value

(a) Residual leniency: application level (b) Relationship leniency and success probability



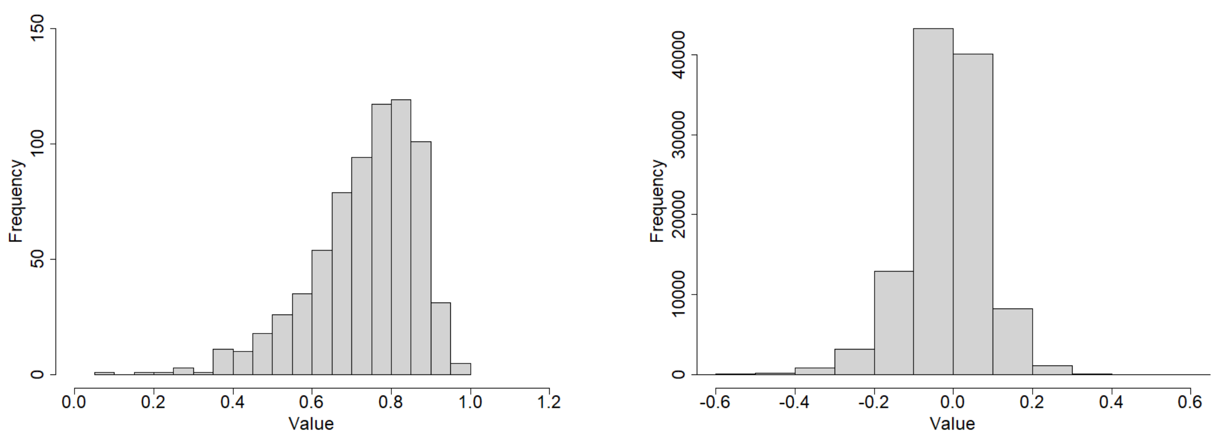
Note: Panel (a) shows a histogram of examiner leniency, calculated according to Equation (1), for n = 6,038,044 patent applications. The values are residualized using art-unit $\times$ year fixed effects. Panel (b) shows a binned scatter plot of the examiner leniency and the probability of getting the patent granted. There are 5000 observations per bin.

As mentioned above, it is generally accepted that patent grant rates differ by different technologies and thus differ by art units. Because examiner leniency is calculated from grant probabilities, this effect is reflected in the leniencies, as well. To visualize how much different art units differ in their leniency, we show a histogram of the average leniency per art unit (residualized by year fixed effects) in panel (a) of Figure 2. The strong variation across art units makes clear why any analysis of examiner leniency has to control for art unit assignment. Righi and Simcoe (2019) argue that examiner specialization exists even below the art-unit level and the most recent version of the MPEP even codifies this. To visualize how much leniency varies by US patent classification (USPC) subclass, we show a histogram over the average residual leniency across the 110,118 subclasses in panel (b) of Figure 2. Data is residualized by art-unit × year fixed effects but nevertheless shows some variation on the subclass level. This variation is, however, significantly smaller than the variation on the examiner level, shown in panel (a) of Figure 1.

Figure 2: Leniency by Art Unit and and USPC Subclass

(a) Average leniency by art unit

(b) Average residual leniency by USPC subclass



Note: Panel (a) shows a histogram of average examiner leniency, calculated according to Equation (1), for n = 707 art units. The values are adjusted by year fixed effects. Panel (b) shows a histogram of average examiner leniency for n = 110,118 USPC class by subclass combinations. The values are residualized using art-unit × year fixed effects.

3 Can the patent lottery be gamed?

A first test of whether the entity providing legal counsel to a patent applicant can influence the patent lottery is to see whether the choice of counsel has an effect on the examiner leniency. By virtue of a random process, different entities will have different average examiner leniencies in their patent applications. We are concerned with the question of whether the idiosyncratic effects of the different entities exceed the level implied by a random process. For this purpose, we first derive the distribution of the entities' effects under the reference model of a random process and then compare this distribution with the distribution observed in the data. To have a sufficient number of observations per entity, we focus our analysis on those that provide counsel for at least 1,000 applications. This retains 93% of the sample, covering 5.62 million applications across 1,427 entities.

We calculate the reference model using a numeric approach. Intuitively, if examiner assignment is already truly random, then re-scrambling may change which entities receive better-than-average and worse-than-average examiners, but the difference in average leniency for the newly lucky and newly unlucky entities should remain constant. Specifically, we randomly scramble the residualized examiner leniencies across the 5.62 million patent applications. We then use this randomly allocated dataset to estimate fixed effects for the reassigned law firms. We repeat this process 1,000 times. The replications allow us to calculate the average value for all 1,427 ordered fixed effects.

To obtain the actual effects of entities on residual examiner leniency, we estimate

$$\text{Residual Examiner Leniency}_i = \sum_{\lambda=1}^{L} \beta_\lambda \mathbf{1}(\lambda \in \mathcal{L}(i)) + \varepsilon_i. \tag{2}$$

We start at the examiner leniency as calculated in equation (1) for patent application i. We then residualize it by projecting out art-unit $\times$ year fixed effects. We aim to observe the fixed effects for the different law firms, which are given by the different β_λ. $\mathcal{L}(i)$ represents the set of all entities that provide legal counsel for application i. Because multiple entities can be associated with the same patent, we estimate Equation (2) through a sparse matrix approach. For comparison with the reference model, we order the fixed effects by their size.

The two different distributions of fixed effects are shown in panel (a) of Figure 3. Here we can see that the empirical differences in examiner leniency between the considered entities are considerably larger than would be predicted under the reference model where the allocation of examiners is random conditional on the art unit and the filing year. A Kolmogorov-Smirnoff test validates the visual assessment. It shows statistically significant differences between the empirical distribution and the distribution derived from the reference model ($D = 0.16$, $p < 0.001$). Importantly, the difference in the distributions is not driven by the extreme tails, where such differences could be attributable to chance.[10] Rather, there is also a difference between the curves

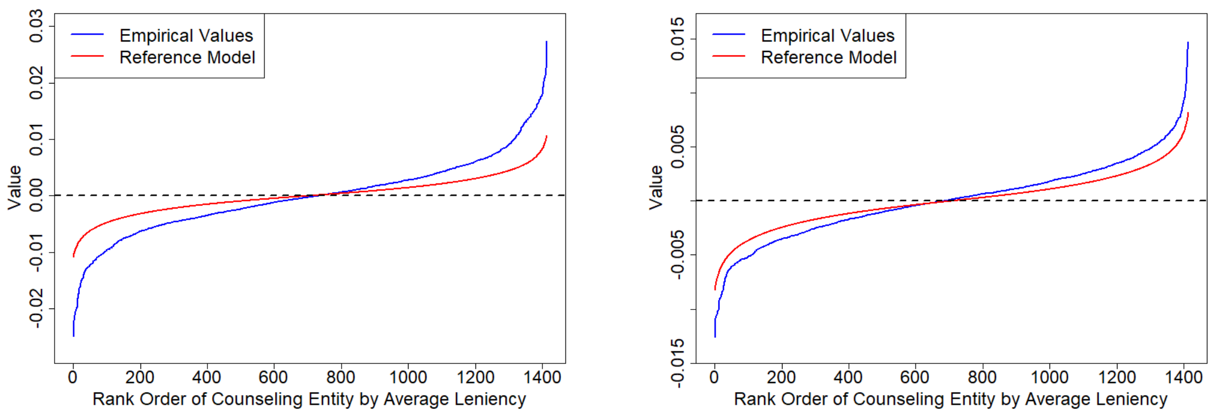
[10]Note that the blue line in Figure 3 is the sum of two effects: the actual entity's effect and random sampling effect. Since the latter is, by definition, independent from the former, extreme values in the tails of the (blue) empirical distribution can appear when a large random effect is paired with a large actual entity effect. We thus truncate the sample by removing the upper and lower 0.5% of observations for better visibility. Graphs with all entities are given in Appendix B.

for medium-high and medium-low ranked entities. This shows a the quality of the legal counsel affects the leniency of the patent examiner across the full sample of legal entities.

Figure 3: Comparison of Empirically Observed Counseling Entity Effects with Reference Model of Random Examiner Leniency

(a) Art-unit × Year Residual

(b) Art-unit × Year × USPC Subclass Residual



Note: The figure considers data from n = 5,621,374 observations from 1,413 entities proving legal counsel. The original sample contains 1,427 observations, but we remove the upper and lower 0.5% of observations for better visibility. Graphs with all entities are given in Appendix B. We only consider patent applications associated entities associated with at least 1000 applications in the data. The reference model is calculated as the average of 1,000 replications of scrambling examiner leniency across applications. The empirical values come from estimating Equation (2). The reference model is calculated for the full sample of 1,427 entities and then truncated in the same way as the empirical values.

Righi and Simcoe (2019) raise the point that patent examiners specialize into specific technological sub-classes of their art units and that these sub-classes have different grant rates associated with them (see panel (b) of Figure 2). In theory, it is possible that counseling entities also specialize on such sub-classes and that what we see in panel (a) of Figure 3 reflects this specialization. To understand whether such a specialization effect drives our results, we repeat the analysis using examiner leniency residualized by art-unit × year × USPC subclass effects.[11] The result can be seen in panel (b) of Figure 3. Even under this extremely granular fixed effect structure, where we demean over 1.68 million categories, the result of our analysis remains the same: entities seem to have a more extreme effect on examiner leniency than can be explained by simply random allocation of patent applications. These visual results are again validated by Kolmogorov-Smirnoff tests (a$D = 0.10$, $p < 0.001$).

[11]For example, private companies, such as LexisNexis, offer to "Draft your patent application to better target a favorable tech center group." (LexisNexis Intellectual Property Solutions, 2026).

While there still is a visual difference between the empirical values and the reference model in panel (b), it is not as large as in panel (a). It is, however, unclear which of the two panels should be considered the better analysis for manipulation. After all, while it might be hard to influence which art unit an application gets assigned to, the sheer number of USPC subclasses suggests that some of them are quite close to each other and that counseling entities can thus influence subclass assignment. This would imply that that residualizing by USPC subclass already removes some of the idiosyncratic effects of counseling entities that we aim to find.

There are two possible concerns. The first is that the empirical values calculated in Equation (2) assign each counseling firm the same weight, regardless of their placement on the application as first, second, or third. This could dilute the effect of the most important counseling entities or our analysis could pick up a mechanical effect of shared representation rather than an entity-specific effect. To address this concern, we repeat the analysis, but only use the counseling entity with the most lawyers present on the application, which we refer to as the "first listed" entity. Second, there could be a cross-contamination effect from average quality of applications counseled by a given entity to our leniency measure when entities get assigned the same patent examiner repeatedly. To avoid this, we also calculate leniency using a leave-one-entity-out (LEO) design where the grant rates of all applications from the counseling entity get excluded when calculating the leniency of the examiner assigned to an application. In this case, we can again only consider the first listed counseling entity, because otherwise, we would have multiple leniency values per application.

Table 2: Robustness of Non-random Effects of Counseling Entities

	Art-unit $\times$ Year	Art-unit $\times$ Year $+$ USPC Subclass	Art-unit $\times$ Year $\times$ USPC Subclass
All entities, LOO Leniency	0.16	0.12	0.10
	$p < 0.001$	$p < 0.001$	$p < 0.001$
First entity, LOO Leniency	0.17	0.15	0.13
	$p < 0.001$	$p < 0.001$	$p < 0.001$
First entity, LEO Leniency	0.17	0.16	0.13
	$p < 0.001$	$p < 0.001$	$p < 0.001$

Note: The table repeats the analysis of comparing empirically observed counseling entity effects with the reference model of random examiner leniency. The first row reports the test statistic and p-values of the main analysis with the addition of Art-unit $\times$ Year $+$ USPC Subclass residuals. The second row considers only the first listed entity as providing counsel. The third row also limits the counseling entities to the first listed ones and calculates the leniency using a leave-one-entity-out (LE0) design. The first row considers data from n = 5,621,374 applications counseled by 1,427 entities proving legal counsel. The second and third row consider only 1,406 counseling entities and the third row reduces the data to 5,973,302 applications.

The results of both robustness analyses and the main analysis are given in Table 2. We report them for the two fixed-effect structures already displayed in Figure 3 and add an Art-unit $\times$ Year $+$ USPC Subclass effect structure, as well. We can see that using only the first-listed entity increases the test statistic across all levels of residualization. Calculating the leniency using a LEO design increases the statistics further. We can thus see that the entity effects are unlikely to be driven by mis-attribution across co-counsel or by contamination of the examiner-leniency measure. Instead, they are consistent with entity-specific differences in examiner assignment.

Who are these entities that produce particularly good (or bad) draws from the patent lottery? To answer this question, we consider the observable characteristics of the 1,427 entities with 1000 or more applications in our data. Let X be a set of averaged application-level control variable. We estimate the equation

$$\begin{aligned}\text{Rank}_i = \beta_0 + \beta_1 \cdot \#\text{Apps.}_{\cdot i} + \beta_2 \cdot \mathbf{1}(\text{IP Stars}_i) + \beta_3 \cdot \text{HHI}_i + \\ \beta_4 \cdot \mathbf{1}(\text{In-House}_i) + \beta_5 \cdot \mathbf{1}(\text{Other}_i) + X_i \gamma' \varepsilon_i. \end{aligned} \tag{3}$$

For robustness against outliers, the dependent variable is the entity's rank. It is the result of the fixed effect analysis displayed in Figure 3. We consider both the ranks pertaining to the leniency residualized with Art-unit $\times$ Year fixed effects and to that residualized with Art-unit $\times$ Year $\times$ USPC subclass fixed effects.[12] Note that for ease of interpretation, we estimate Equation (3) with OLS even though the rank is ordinal in nature. Thus, the marginal effects in our estimation represent an average ordinal shift across the distribution rather than linear changes in the underlying fixed effects.

Summary statistics of all variables are given in Table A1 in the Appendix. 61 of the 983 law firms in our sample are listed in the 2020 ranking of the IP Stars specialists guide. According to the guides own information, the ranking is "based on a weighted review of information submitted by firms, publicly available information, and market feedback."[13] The Art Unit HHI is a concentration measure for the applications of a given entity. A higher value implies that the entities focus on a more narrow set of art units, implying greater familiarity with both the subject matter and the bureaucratic idiosyncrasies of the art units in question. In order to make the results more interpretable, we normalize this variable to have a mean of 0 and a standard deviation of 1. When all entities are considered, law firms (n=983) are the reference category in the estimation. In ad-

[12] Unsurprisingly, these ranks are highly correlated. See Figure A2 in Appendix B.1.

[13] See https://www.ipstars.com/ for futher details. We accessed the 2020 data of the ranking using the wayback machine (https://web.archive.org/). It was the earliest ranking accessible in this fashion.

dition, there are 367 in-house counsels in our data. Other entities includes individuals (n=34), the government (n=10) and unclassified entities (n=33).

Table 3: Determinants of Entity Fixed Effect Rank

	Dependent Variable: Rank of Entity Fixed Effect			
	(1)	(2)	(3)	(4)
Constant	784.2***	774.4***	731.1***	731.0***
	(20.13)	(23.57)	(20.64)	(23.75)
No. Applications	-2.959	-1.675	-11.32	-9.579
	(7.831)	(9.211)	(8.023)	(9.949)
IP Stars Ranked (2020)	126.6***	123.7**	128.3***	118.3**
	(47.00)	(48.37)	(48.41)	(50.55)
Art Unit HHI	44.82***	49.19	27.17**	30.83
	(16.96)	(39.84)	(12.96)	(35.11)
In-House Counsel	-13.76		-25.85	
	(29.42)		(29.13)	
Other Entity	52.98		-4.399	
	(48.46)		(51.22)	
Entity Types	All	Law Firms	All	Law Firms
Application Controls	Yes	Yes	Yes	Yes
Residualized	Art Unit × Year	Art Unit × Year	Art Unit × Year × Subclass	Art Unit × Year × Subclass
R^2	0.10294	0.08861	0.07311	0.09290
Observations	1,427	983	1,427	983

Note: Table shows the results of estimating Equation (3) either with data on all entities (columns (1) and (3)) or with data for law firms only (columns (2) and (4)). The rank is the result of the fixed effect analysis displayed in Figure 3. For columns (1) and (2) we consider the fixed effects from leniency residualized with Art-unit × Year fixed effects. For columns (3) and (4) we additionally add USPC subclass fixed effects. 61 of the 983 law firms are listed in the 2020 IP Stars ranking. The Art Unit HHI is a concentration measure, normalized to mean 0 and standard deviation 1. In columns (1) and (3), law firms are the reference category. Other entities includes individuals (n=34), the government (n=10) and unclassified entities (n=33). Application controls are averages of the number of application claims, the scope and the small entity status of the applicant. Standard errors are heteroscedasticity robust. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Results of the estimation are given in Table 3. Throughout all specifications, we can see that the influence of the number of applications is limited. By contrast, the IP Stars ranking has a large, positive, and statistically significant effect, irrespective of the specification. The ranking either already takes the positive outcomes into account or better reputed firms are better at influencing the patent lottery. The concentration measure is interesting. When only residualizing by art unit and year, it is positive and significant. This shows that concentration on a few technological fields increases the effectiveness at influencing the patent lottery. This effect becomes smaller, when the USPC subclass is added to the residualization process of the leniency. We know from Righi and Simcoe (2019) that examiners (and thus their leniency) differ by these subclasses, at least for some art units. Strategic use of language when writing claims in the patent application can influence the subclass allocation of the application. Comparing the coefficient on the Art Unit HHI between the first column and the third column indicates that the institutional knowledge necessary for such strategic behavior might be what part of the effect of art unit concentration on entity rank. Lastly, we can see no significant difference between entity types at gaining an advantage in the patent lottery.

4 How do you game the patent lottery?

We now ask how the patent lottery can be manipulated. An immediate thought is that entities must be using illicit means to influence patent examiner assignment. We instead focus on two, legal, channels: strategic application timing based on publicly available data and insider knowledge relevant to the examiner assignment process.[14]

4.1 Strategic timing of applications

Real-time publishing of patent decisions, through a system called PublicPair and its successor system, the Patent Public Search Basic (PPUBS), provide for a different, completely legal, way to get a favorable assignment.[15] Because there are a limited number of examiners in any art unit and because applications are assigned taking examiner availability into account, an application can be timed strategically so that only more lenient examiners are available at the time of submission. Below, we use a relatively simple prediction model to show that such strategic submission timing is in fact possible.

We approach the issue as follows. If strategic submission timing was not feasible, we would be unable to predict the examiner-assignment environment faced by a newly filed application based on timing alone. The null hypothesis would thus be that, conditional on the art unit, the leniency of the examiner assigned to an application filed on day t could not be predicted based

[14] We also note that insider knowledge may also help law firms use publicly available data for strategic timing.

[15] PublicPAIR was officially retired on August 1, 2022, and replaced by the Patent Center system. See https://www.uspto.gov/patents/public-pair-be-retired for the official transition details. Additionally, the USPTO Open Data Portal's Patent File Wrapper search (https://data.uspto.gov/patent-file-wrapper/search) provides daily-refreshed access to the complete transaction history and bibliographic data of all public applications. This allows sophisticated applicants, such as law firms, to follow live updates on examiner workloads and decision patterns within specific art units.

on previously publicly available information. We thus test whether time-t conditioning variables predict the subsequent examiner leniency received by a newly filed patent application. These predictors are constructed strictly prior to filing, use no information of the focal application other than its art unit, and summarize contemporaneous examiner availability and leniency. Evidence of predictability rejects the null hypothesis and speaks against independence of the patent process. The setup mirrors market timing: an investor uses a time-t signal observed today to forecast future returns. Likewise, a law firm or in-house counsel observes expected examiner leniency before filing and forecasts whether the assignment window is favorable.

Let i index applications, e examiners, a art units, and t_i the filing date (day) of application i. To ensure the feasibility of the model, we account for the institutional information barriers of the patent system. An examiner's identity and their specific track record are generally not public until the later of the application's 18-month publication date or the first substantive office action (First Office Action Date, FOAD). We define the disclosure date d as the specific day an application's outcome (grant or abandonment) becomes public knowledge. For grants, this corresponds to the patent issue date; for abandonments, it is the mailing date of the notice of abandonment, provided the application is already visible to the public. All "as-of" quantities use end-of-day logic, that is, events on the filing day t_i are excluded. We write $x(d^-)$ for "strictly before day d" and t^- for "strictly before the filing day." Figure 4 illustrates this logic and shows the timeline of a patent application and its information disclosure process.

Examiner leniency. To capture the public reputation of an examiner, we define a rolling public grant-rate ("leniency") step function for each (e, a) based only on outcomes that have already crossed the visibility barrier as

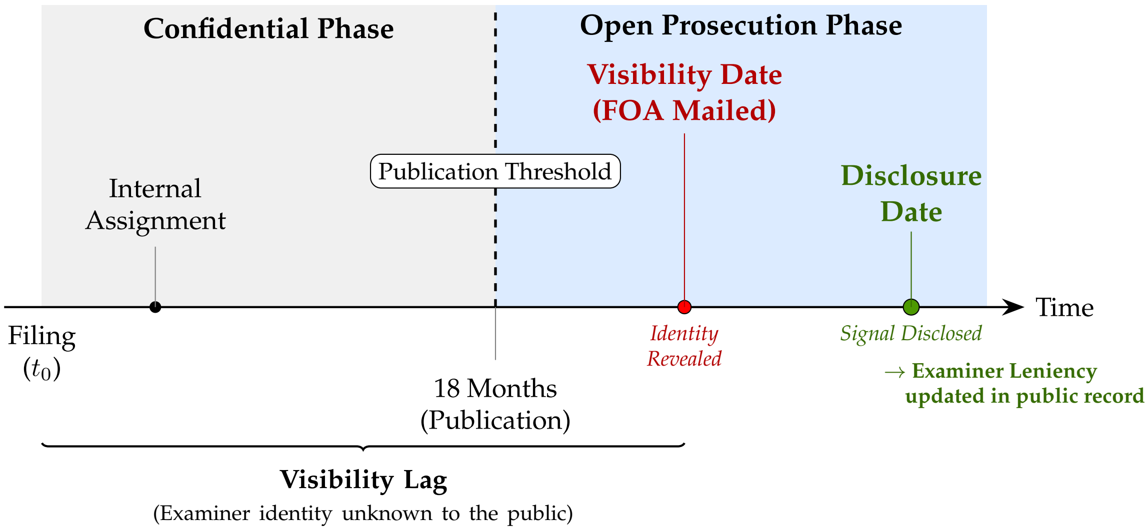
$$\ell_{e,a}(t^-) \;=\; \frac{\#\{\text{patent grants for } (e,a) \text{ with disclosure date } < t\}}{\#\{\text{public final outcomes for } (e,a) \text{ with disclosure date } < t\}}\,, \tag{4}$$

updated only on disclosure dates d.[16]

Final outcomes in the denominator represent the total number of public disposals and exclude technical administrative events, such as Requests for Continued Examination (RCE), which do not constitute a final decision. We require at least $K_{\min} = 15$ prior visible disclosures for $\ell_{e,a}(t^-)$ to be defined; otherwise it is missing. Due to the combination of the institutional visibility lag

[16]While the numerator is defined as patent grants, we follow a behavioral logic and include all public signals where the examiner authorized a grant. This encompasses cases where a notice of allowance was issued but the patent did not reach final issuance due to applicant choice, such as a failure to pay the required issue fee or an express abandonment after allowance.

Figure 4: Timeline of the Patent Application and Information Disclosure Process



Note: This figure illustrates the transition of application data from the confidential phase until it is visible to the public in the open prosecution phase. The **Visibility Date** marks the first moment an examiner's identity is public, defined as $\max(\text{Publication}, \text{FOAD})$. The **Disclosure Date** is the day a final outcome (patent grant or notice of abandonment) is disclosed. At this point, the examiner's public rolling leniency score is updated, providing an actionable signal for future applicants. To ensure feasibility, our predictive model for a new filing at time T only conditions on signals where $d < T$.

and this minimum history requirement, we find that a valid leniency score can be calculated for approximately 61% of all applications in our sample. This figure highlights the information barrier inherent in the patent system: for the remaining 39% of filings, sufficient public information does not yet exist to form an actionable signal at the time of filing. Note that the leniency of an examiner is calculated for every art unit separately. Thus, the approximately 40% of examiners who handle applications in multiple art units have more than one leniency value. By only including publicly disclosed outcomes, we ensure the measure reflects the information set available to a law firm's algorithm at the time of filing.

Expected leniency and availability. For application i in art unit $a(i)$ with filing day t_i, we construct timing predictors from the 21 calendar days prior to filing, excluding the filing day itself. We split this lookback period into three windows: days 1–7, 8–14, and 15–21 before filing. In each window, we identify distinct examiners who had at least one public disclosure in the focal art unit and for whom $\ell_{e,a(i)}(t^-)$ is defined. If an examiner appears multiple times within the 21-day lookback period, we use that examiner's latest disclosure. For each window, we then compute three variables: the mean leniency of available examiners, the log number of available examiners, and the share of all 21-day available examiners appearing in that window. This gives nine predictors, all constructed from information available strictly before filing.

Out-of-sample estimations. Let L_i denote the realized leniency of the examiner assigned to application i. As described in Section 2, we restrict the analysis to applications filed before 2021 to ensure that the application process has been completed. We predict realized examiner leniency from the pre-filing timing predictors,

$$L_i = f(\mathbf{X}_{i,t_i^-}) + \varepsilon_i, \tag{5}$$

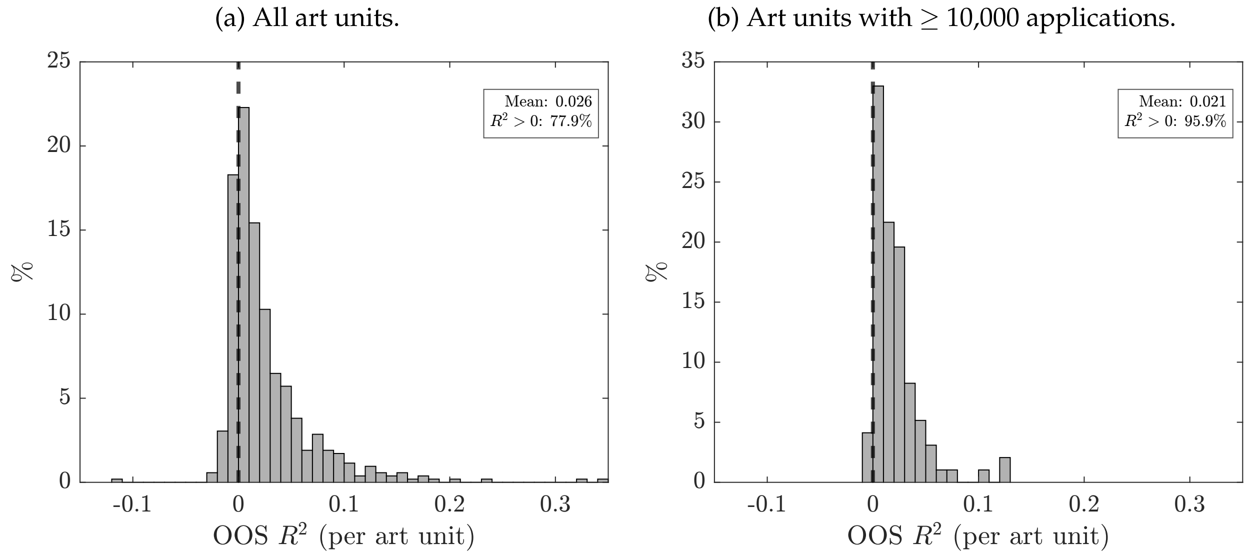
where $\mathbf{X}_{i,t_i^-}$ contains the nine timing predictors constructed from information available before the filing day.

To simulate a real-world prediction problem, in which a law firm would use only public disclosure information to predict the realized assignment environment, we estimate this relation in a rolling out-of-sample (OOS) procedure. The model is estimated separately by art unit. This prevents the model from mechanically predicting leniency from persistent differences across technology areas rather than from within-art-unit variation in recent examiner availability. For each art unit and prediction month, it is trained on applications filed during the previous rolling three-year window (1,095 days) and then applied to applications filed in that month. All training observations precede the predicted observations. We use XGBoost for $f(\cdot)$ because it allows the relationship between recent availability signals and realized examiner leniency to be nonlinear while keeping the information set fixed at the filing date. The use of shallow trees and a small, economically interpretable predictor set limits model complexity.[17] Predictive performance is measured by out-of-sample R^2 relative to a rolling art-unit mean benchmark. We also repeat the exercise for application success, using AUC as the performance measure; those results are reported in the appendix.

Figure 5 reports the results of this estimation. We run the prediction exercise separately by art unit to avoid a pooled model that would mainly identify the application's art unit rather than predict favorable assignment windows. The left panel includes all eligible art units. The right panel restricts the sample to the large art units used in the timing regressions, defined as art units with at least 10,000 applications. Across all art units, the mean out-of-sample R^2 is 0.026 and 77.9% of art units have positive predictive performance relative to the rolling art-unit mean. For

[17] Hyperparameters are selected within each rolling training window using an internal validation split. The full tuning procedure and hyperparameter grid are reported in Appendix B.2.1.

Figure 5: Distribution of Predictive Quality across Art Units



Note: This figure shows the distribution of Art Unit-level out-of-sample R^2 values from rolling XGBoost predictions of realized examiner leniency. The model uses nine latest-examiner predictors constructed from public information available before filing: mean expected leniency, log number of available examiners, and the share of available examiners in the 1–7, 8–14, and 15–21 day windows. The left panel includes all eligible Art Units; the right panel restricts the sample to the large-Art-Unit timing sample, defined as Art Units with at least 10,000 applications. The dashed vertical line marks an out-of-sample R^2 of zero, so values to the right indicate improvement relative to a rolling Art Unit mean benchmark. The boxes report the mean out-of-sample R^2 and the share of Art Units with positive out-of-sample R^2.

large art units, the mean out-of-sample R^2 is 0.021 and 95.9% of art units have positive predictive performance.

These findings indicate that even with public information alone, realized examiner leniency can be predicted from recent availability and leniency signals. The result is particularly strong in large art units, where disclosures are frequent enough for the public signal to be more precisely measured. Appendix Figure A3 shows that the same information also predicts application success: 91.1% of all art units and 99.0% of large art units have AUCs above the random-assignment benchmark of 0.5. Taken together, these results suggest that sophisticated applicants can identify favorable disclosure windows in the assignment queue. The patent process is thus shown to be non-independent; instead, strategic timing allows for manipulating the lottery by targeting periods in which predicted examiner leniency is high.

It has to be noted that the model employed here uses only public disclosure dates and a limited set of availability-based predictors. In a real-world setting, the information asymmetry between the general public and a filing law firm likely provides the latter with an even greater advantage. While external observers remain blind during the 18-month confidentiality phase, law firms may receive information regarding their own extensive dockets before such data enters the public record. If a law firm receives a notification for an action on its own application, such as a First Office Action (FOAD) or a notice of allowance, it identifies the assigned examiner and their current availability status well before the rest of the public. By monitoring these private signals across thousands of concurrent applications, sophisticated repeat players can potentially detect when a lenient examiner clears a slot on their docket months before that information is disclosed via official public channels. Consequently, our results likely represent a conservative lower bound of predictability; with access to such granular, private data, sophisticated models could likely achieve even higher accuracy across a broader range of art units.

4.2 Do Higher-Ranked Entities Time Applications Around Expected Leniency?

Section 4.1 shows that public timing signals contain predictive information about realized examiner leniency. That analysis asks whether favorable examiner-assignment windows are predictable per se: conditional on the available information at the filing date, do applications filed in more favorable windows receive more lenient examiners? A separate question is whether sophisticated applicants or their representatives appear to act on this information. If firms understand that some filing windows are more favorable, then higher-ranked entities should be more likely to submit applications when predicted examiner leniency is high.

To test this, we construct a daily panel at the entity–art unit level. We estimate the regressions both for all art units and for the 97 large art units with at least 10,000 applications. For each entity, art unit, and filing day t, the dependent variable is either the log number of applications filed on that day, $\log(1 + \text{Submissions})$, or an indicator equal to one if the entity files at least one application in that art unit on that day. The main explanatory variable is the fitted value from the XGBoost leniency prediction model, $\widehat{\text{Leniency}}_{\text{XGB},a,t^-}$, standardized within art unit. The notation t^- emphasizes that the signal is assigned to filing day t but constructed only from public information available strictly before that day. Thus, a one-unit increase corresponds to a one-standard-deviation increase in predicted examiner leniency within the art unit on that filing day.

The key heterogeneity variable is the entity rank. Throughout the main analysis, entity rank is measured using the annual seven-year rolling rank. Let $y(t)$ denote the calendar year of filing day t. For applications filed in year $y(t)$, we estimate entity quality using only applications filed during the previous seven calendar years, from $y(t) - 7$ through $y(t) - 1$. The rank is therefore strictly predetermined relative to the current filing decision. Within each rolling window, we regress realized examiner leniency on entity indicators, art-unit fixed effects, and filing-year fixed effects.

Entities are then ranked based on the estimated entity fixed effects. The rank is normalized to lie between zero and one, where zero denotes the lowest-ranked entity in the rolling window and one denotes the highest-ranked entity. This construction allows entity ranks to vary over time, while accounting for persistent differences across art units when estimating the rank. The regression specification is

$$\begin{aligned} Y_{i,a,t} = {} & \beta_1 \widehat{\text{Leniency}}_{\text{XGB},a,t^-} + \beta_2 \text{Rank}_{i,y(t)-1} + \beta_3 \widehat{\text{Leniency}}_{\text{XGB},a,t^-} \times \text{Rank}_{i,y(t)-1} \\ & + \gamma_i + \delta_t + \alpha_a + \varepsilon_{i,a,t}. \end{aligned}$$

Here, γ_i are entity fixed effects, δ_t are day fixed effects, and α_a are art unit fixed effects. Standard errors are two-way clustered by entity and day. The coefficient of interest is β_3. It measures whether the same filing-day predicted-leniency signal is associated with a stronger filing response among higher-ranked entities.

Table 4: Strategic Filing into Favorable Examination Windows

	All art units		Large art units	
	Log sub. (1)	Sub. 1/0 (2)	Log sub. (3)	Sub. 1/0 (4)
$\widehat{\text{Leniency}}_{\text{XGB}}$	−0.0001	−0.0003	−0.0000	−0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Entity Rank	0.0012	0.0012	−0.0004	−0.0008
	(0.0012)	(0.0015)	(0.0010)	(0.0012)
$\widehat{\text{Leniency}}_{\text{XGB}}$ × Entity Rank	0.0010***	0.0012***	0.0018***	0.0022***
	(0.0003)	(0.0004)	(0.0004)	(0.0005)
Clustering	Entity + Day	Entity + Day	Entity + Day	Entity + Day
Within R^2	0.00001	0.00001	0.00005	0.00004
Observations	128,331,881	128,331,881	53,079,931	53,079,931
Entity FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Art Unit FE	Yes	Yes	Yes	Yes

Note: Table reports timing regressions using fitted values from the XGBoost leniency prediction model. The dependent variables are log submissions and an indicator for any submission by an entity on a given art-unit day. Columns (1)–(2) use all art units; columns (3)–(4) use large art units. Fitted leniency is standardized within art unit. Entity ranks are based on backward-looking seven-year windows, adjusted for art-unit and filing-year fixed effects, and lagged by one year. All specifications include entity, day, and art unit fixed effects. Standard errors are two-way clustered by entity and day. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

The results are given in Table 4. The interaction term is positive and statistically significant in all four specifications. This means that higher-ranked entities react more strongly to favorable predicted-leniency windows. This implies that at least part of the ranking of law firms comes from active decisions these firms make about the timing of applications. In the all-art-unit sample, moving from the lowest-ranked to the highest-ranked entity increases the response to a one-standard-deviation increase in predicted leniency by approximately 0.10 percent in the log-submissions specification and by 0.12 percentage points in the submission-indicator specification.

The effects are larger in the large-art-unit sample. There, moving from the lowest-ranked to the highest-ranked entity increases the response to a one-standard-deviation increase in predicted leniency by approximately 0.18 percent in the log-submissions specification and by 0.22 percentage points in the submission-indicator specification. Relative to the baseline daily submission probabilities of 1.95% in the all-art-unit sample and 2.50% in the large-art-unit sample, these probability effects correspond to increases of approximately 6.2% and 8.8%, respectively. The larger effect in high-volume art units is consistent with the interpretation that timing is easier when public disclosure signals are more frequent and the assignment environment can be measured more precisely.

Although these economic magnitudes are modest, they are likely to be conservative. Both variables entering the interaction are estimated rather than directly observed. The fitted leniency measure is generated from an out-of-sample prediction model based only on public disclosure information. Sophisticated entities may have access to richer information sets and more refined prediction models, so the variation captured by our measure may understate the information actually used when timing applications. The entity-rank variable is also estimated from past applications and therefore contains measurement error. Measurement error in either component of the interaction attenuates the estimated heterogeneous timing response and makes it harder to detect sorting by entity quality. The estimates should therefore be interpreted as evidence of strategic timing based on noisy proxies for both the relevant filing-window advantage and entity sophistication.

Overall, the evidence suggests that predicted examiner leniency is not only informative about realized examiner assignment, but is also related to filing behavior. The positive interaction indicates that higher-ranked entities are more likely to shift submissions toward favorable predicted-leniency windows. Appendix Table A3 shows that the conclusion is similar when we use the raw split-window expected-leniency predictors directly rather than the XGBoost-fitted values. This is consistent with strategic timing by more sophisticated or better-performing entities.

4.3 The revolving door between USPTO and patent firms

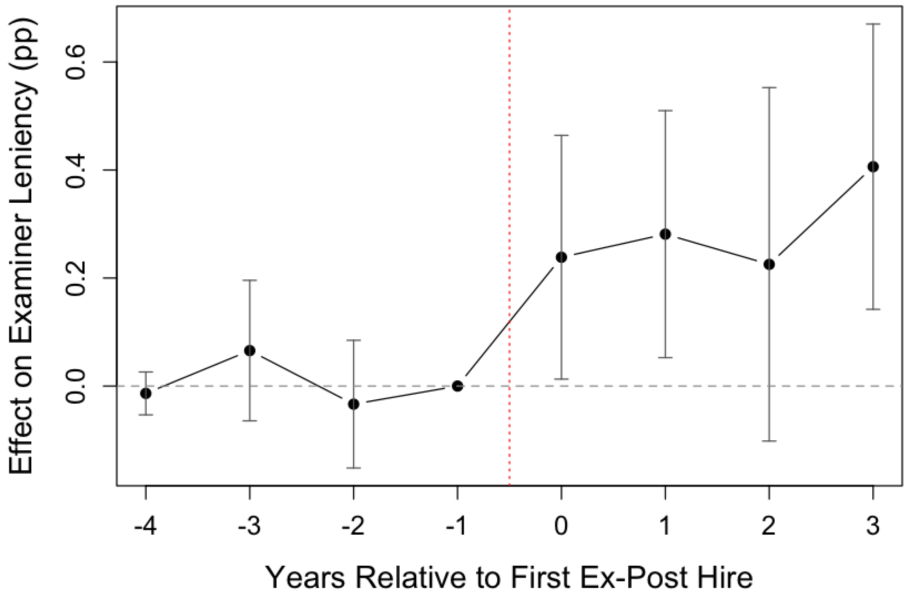
One channel through which firms may gain the institutional knowledge necessary to manipulate examiner assignment is the "revolving door" between the USPTO and the private patent bar. Attorneys who have worked as patent examiners and subsequently join private law firms bring detailed knowledge of internal procedures, examiner workloads, and assignment practices. We investigate whether firms that accumulate such revolving door hires receive more lenient examiners over time.

We identify revolving door connections from LinkedIn career history data. We process 110 million US LinkedIn profiles, extracting the full employment history for each. To identify former USPTO employees, we match organization names against USPTO name variants using Jaro-Winkler string distance (≤ 0.08), then apply three successive filters: validating matched organizations against a patent office whitelist (excluding commonly confused agencies), retaining only patent-relevant job titles (examiners, administrative patent judges, trademark examining attorneys), and tightening firm-level string distance thresholds. After filtering, we identify 2,585 unique revolving door individuals connected to 1,363 canonical firms. The revolving door operates in both directions, but transitions from the USPTO to private firms (ex-post) are the most relevant for our institutional knowledge hypothesis. Additional details on the data processing funnel and summary statistics are provided in Appendix B.2.

Our primary specification is an event study around the timing of each firm's first ex-post hire, estimated with the Borusyak et al. (2024) imputation estimator (see also Gardner, 2022) to avoid the contamination of treated-on-treated comparisons that biases standard two-way fixed-effects event studies under staggered adoption with heterogeneous treatment effects. If revolving door connections causally improve examiner assignment, the leniency effect should appear only after the hire. Figure 6 presents the results. All pre-hire coefficients are small and statistically insignificant, while post-hire coefficients ($t = 0$ through $t = +3$) are uniformly positive, with the effect peaking at $t = +3$ at 0.41 percentage points. The absence of pre-trends supports a causal interpretation: firms do not receive more lenient examiners before they hire from the USPTO.

To gauge the economic magnitude of this effect, we combine the event-study ATT with two external benchmarks. Averaged over the four years immediately following the first ex-post hire, the post-hire coefficients imply that a treated firm's per-application grant probability rises by approximately 0.29 percentage points. Among firms with at least 1,000 lifetime patent applications—the population plausibly hiring ex-USPTO staff—the mean such firm files 212 applications per year, so the first hire generates approximately 0.61 additional patent grants per year. Valuing each

Figure 6: Event Study: Examiner Leniency Around First Ex-Post Hire



Note: Coefficients from a Borusyak et al. (2024) two-stage imputation event study (see also Gardner, 2022) of examiner leniency on indicators for years relative to the firm's first ex-post USPTO hire, matching the imputation columns of Table 5. The reference period is $t = -1$ and the leftmost bin collects $t \leq -4$. The first stage fits leniency on firm-by-art-unit and art-unit-by-year fixed effects together with application controls (small entity status, number of claims, scope) on untreated applications only; the second stage regresses the residual on event-time indicators interacted with treatment. Standard errors are from a firm-level fractional random-weight (Dirichlet) block bootstrap with $B = 200$ draws, and bands are ± 1.96 bootstrap standard errors. The dashed line marks zero.

marginal grant at the mean market-implied patent value of Kogan et al. (2017)—$10.36M in 1982 dollars, or approximately $35M in today's dollars after CPI adjustment—yields a naive annual value of roughly $21M. Because examiner capacity within an art-unit-by-year cell is approximately fixed, however, the patent lottery is zero-sum: a treated firm's gain is offset by lower leniency for other applicants in the same cell.[18] After this SUTVA correction the realized annual value is approximately $12M per year. Two caveats apply. First, the event-study coefficients are estimated with non-trivial uncertainty, so these dollar magnitudes should be read as orders of magnitude rather than point estimates. Second, the Kogan et al. (2017) valuations are estimated on public

[18]Formally, with examiner capacity fixed within an art-unit-by-year cell, total leniency assigned in the cell is conserved: any increase in mean leniency for treated applicants is offset by a decrease for untreated applicants. Let s_c denote the share of applications in the cell coming from already-treated firms (those that have made an ex-post USPTO hire by the filing year). The naive ATT compares treated firms to a hypothetical zero-treatment baseline; the realized advantage of a treated firm relative to the cell average is $(1 - s_c)$ times the naive ATT, so the implied dollar value scales by the same factor. We compute s_c as the application-weighted share of treated observations in the regression sample (equivalently, the simple share of treated applications), which yields $s_c \approx 0.43$.

firms and likely overstate the average applicant's marginal grant value; the median KPSS patent (approximately \$10M in today's dollars) would scale these numbers down by roughly a factor of three.

Table 5: Revolving Door Stock and Examiner Leniency

	Imputation		TWFE	
	Ext. (1)	Int. (2)	Ext. (3)	Int. (4)
Has RD Stock (extensive margin)	0.2872***		0.1372***	
	(0.0986)		(0.0503)	
log(RD Stock) (intensive margin)		0.1305		0.0799
		(0.1489)		(0.0766)
Estimator	Imputation	Imputation	TWFE	TWFE
Std. errors	Firm-level Dirichlet block bootstrap ($B = 200$)			
R^2	—	—	0.49772	0.49181
Observations	3,791,164	140,019	5,639,790	2,426,282
Entity FE	✓	✓	✓	✓
Art Unit × Year FE	✓	✓	✓	✓

Note: Dependent variable is examiner leniency (percentage points). Columns (1)–(2) report the Borusyak et al. (2024) two-stage imputation difference-in-differences on a staggered-adoption (switcher) sample, where treatment is a firm's first ex-post USPTO hire: the extensive margin is the static ATT (event time $t \in [0, 3]$) and the intensive margin regresses the residualized outcome on $\log(\text{RD Stock})$ among treated applications; the first stage uses firm-by-art-unit and art-unit-by-year fixed effects. Columns (3)–(4) report the two-way fixed-effects (TWFE) extensive and intensive margins on the full sample, the latter restricted to firm-years with positive stock and using $\log(\text{RD Stock})$ without the $+1$. All specifications control for small entity status, number of claims, and patent scope (coefficients suppressed). Standard errors in parentheses are from a firm-level fractional random-weight (Dirichlet) block bootstrap with $B = 200$ replications, re-running both imputation stages under each draw. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

We corroborate the event-study finding with the panel specifications in Table 5, which place the Borusyak et al. (2024) imputation estimator alongside two-way fixed-effects (TWFE) on both the extensive and intensive margins. The dependent variable throughout is examiner leniency (leave-one-out), measured in percentage points; all specifications include firm and art-unit-by-year fixed effects (the imputation first stage uses firm-by-art-unit fixed effects) and control for small entity status, number of claims, and patent scope, and standard errors are obtained from a firm-level fractional random-weight (Dirichlet) block bootstrap with $B = 200$ draws. Columns (1) and (3) report the extensive margin: an indicator for whether the firm has any ex-post hire by the filing year. The imputation estimate of 0.287 percentage points is roughly twice the TWFE estimate of 0.137, with both significant at the 1% level; the larger imputation coefficient is consistent with the downward bias that contaminates TWFE under staggered adoption with heterogeneous treatment effects. Columns (2) and (4) report the intensive margin using $\log(\text{RD Stock})$ (without the $+1$), restricted to firm-years with at least one hire. The intensive-margin coefficient is positive under both estimators (0.131 and 0.080 percentage points, respectively) but noisy—not statistically distinguishable from zero—indicating that the effect operates primarily through the extensive margin: whether the firm has any revolving-door connection at all, rather than the number of such connections. The flat pre-hire coefficients in Figure 6 serve as a future-hires placebo, ruling out that these connections merely proxy for time-invariant firm quality. Additional specifications, including alternative revolving-door measures and a mediation analysis of the leniency-to-grant-probability channel, are reported in Appendix B.2.

5 When can we expect companies to game the patent lottery?

The previous sections have shown that the patent lottery can be gamed and how it is gamed. We now consider which applicants are particularly likely to invest into such manipulation. We first provide theoretical predictions for this question using a simple game-theoretic model in Section 5.1. The comparative statics from this model provide hypotheses which are then analyzed empirically in Section 5.2. We use the same model as a basis for our welfare discussion in Section 7.

5.1 Model

We use a simplified version of the Nordhaus (1969) model in the notation of Budish et al. (2016). The model considers the societal trade-off between incentives given to firms for innovating through a period of time in which their innovations are protected by patents and the welfare losses due to the monopoly rents induced by patents in that time period. For this, the model considers a collection of firms that differ on their cost of innovation, the chance that innovation succeeds and the monopoly rents they can obtain from a patent. Other than is typically the case, we are not interested in finding a societal equilibrium for the optimal patent protection length. Rather, we use the model's setting to analyze the effect of a random element in the patent system and the possibility to manipulate this element.

To focus on this aspect, we use a series of simplifying assumptions to keep the analysis tractable and the exposition intuitive. First, we assume that there are only two firms that innovate, $i \in \{1, 2\}$, and each firm has a single possible innovation to patent. Because it is inessential for the purpose of the model, we assume that if a firm attempts to innovate, it is always successful. If a patent application is successful, the firm receives increased profits due to the patent protection. If the application fails, there is free firm entry to the technology and competitive pressure results in reduced profits for the firm. We denote the net difference between these two profits as π_i. When using this term, we abstract from the exact length of the patent protection period. π_i can be seen as the discounted total net benefit to the firm of a patent protected innovation.[19]

Cost of innovation for a firm are given by c_i. We consider firms with innovations which can differ in costs and potential benefits, but do not differ in their *ex-ante* probability of succeeding in the patent lottery. Absent manipulation, both patent applications have probability $p_i = \bar{p}$ of succeeding. Because there is no manipulation, the probability of succeeding in the patent lottery is not influenced by firm or application characteristics. While independence from application characteristics is a simplification in the model, it is not material for the empirical test, in which we simply use examiner leniency to measure p_i.[20] Given the probability of succeeding in the manipulation free patent lottery, a risk neutral firm i innovates if $\bar{p}\pi_i \geq c_i$. Since the model uses $p_i = \bar{p}$, there are no comparative statics of p_i with regard to firm characteristics.

Comparative statics do appear if firms can manipulate the patent lottery. We model this by letting them spend expenses e_i at marginal cost k_i. A higher level of e_i increases the probability of a granted application, while spending of the other firm (e_{-i}) decreases it. We assume that $p(e_i, e_{-i})$ is twice continuously differentiable and that expenses have decreasing benefits such that $p_1 > 0, p_2 < 0, p_{11} < 0$ and $p_{22} > 0$. As argued in the introduction, the nature of the manipulation is to spend ressources such that the application is assigned to a more lenient patent examiner. Because the pool of examiners is fixed (at least in the short term), manipulating the lottery in this form is a zero-sum game. As such, it has to hold that

$$p(e_1, e_2) + p(e_2, e_1) = 2\bar{p}. \tag{6}$$

If a firm expends money to get a more lenient patent examiner, the other firm must get the less lenient examiner. Since the expenses are targeted at the selection of the examiner rather than at influencing the examiners themselves (for example through bribes), the average leniency in the system does not change.

[19]Note that this benefit is private for the firm and does not represent the societal benefit of the innovation. It appears because of monopoly protection for the firm and thus comes at the expense of the rest of society. Given the deadweight loss inherent to monopolies, total welfare is larger if the application fails than if it succeeds. Nevertheless, a probability of success has to be present, because otherwise the firms would have no incentive to innovate.

[20]Alternative model approaches in which $\bar{p}$ is a baseline probability that is then adjusted based on the characteristics of the innovation are possible but will only complicate the exposition without any further information. The advantage of using examiner leniency as a measure for the probability of patent lottery success is that it is independent from the characteristics of applicant and application unless there is manipulation, which is the exact hypothesis our empirical analysis aims to study.

We can now see that firms innovate if

$$\pi_i p(e_i^*, e_{-i}^*) - k_i e_i^* \geq c_i \tag{7}$$

where asterisks indicate the equilibrium solution to the manipulation game played by the two firms. The equilibrium of the game is given in the following proposition.

Proposition 1. *Conditional on both firms innovating and* $p_1(0,0) > \max\left\{\frac{k_1}{\pi_1}, \frac{k_2}{\pi_2}\right\}$*, the unique Nash equilibrium of the patent lottery manipulation game is given by*

$$p_1(e_1^*, e_2^*) = \frac{k_1}{\pi_1} \qquad \textit{and} \qquad p_1(e_2^*, e_1^*) = \frac{k_2}{\pi_2}. \tag{8}$$

All proofs are provided in the appendix. An interior solution implies that both firms spend money on manipulating the patent lottery. While we cannot show this claim explicitly in the empirical data, we can show evidence that is consistent with it. For this, we derive the comparative statics of the Nash equilibrium. Specifically, we are interested in how the grant probability of a patent changes when the parameters of the firm or its application change. Showing support for these results in the data will give credibility to the claim that manipulation does indeed exist.

Proposition 2. *Under the conditions of the interior Nash equilibrium,* $\frac{\partial p(e_i^*, e_{-i}^*)}{\partial k_i} < 0$ *and* $\frac{\partial p(e_i^*, e_{-i}^*)}{\partial \pi_i} > 0$ *if* $p_{i,-i}(e_i^*, e_{-i}^*)$ *is not too negative.*

This proposition provides us with two testable hypotheses. $\frac{\partial p(e_i^*, e_{-i}^*)}{\partial k_i} < 0$ states that firms with higher marginal costs of influencing the patent lottery will invest less into the patent lottery and are thus less likely to have a patent examiner with a high leniency assigned to them. In the real world, differences in marginal costs will likely appear due to differences in liquidity or access to informal networks. Both are more likely to be present in large firms. Proposition 2 thus implies that smaller firms will, on average, have less lenient examiners. $\frac{\partial p(e_i^*, e_{-i}^*)}{\partial \pi_i} > 0$ states that firms with patent applications that have a higher value once they are granted will invest more in the patent lottery and are thus more likely to have a lenient patent examiner. We test this hypothesis using the number of claims and the scope in a patent application as a proxy for the patent's quality.

5.2 Empirical Evidence

We test the hypotheses derived from the theoretical model using proxies of firm size and patent quality that are included in the USPTO data. For the cost of capital, we already argued that individual inventors or smaller firms will have a harder time raising sufficient funds to influence the patent lottery. The USPTO tracks the size of applicant through a small entity indicator. When an individual, a nonprofit organization or a firm with fewer than 500 employees applies for a patent, the application fee is reduced by 50%. We use this indicator as a proxy for cost of capital, building on the well-established result that the cost of capital is decreasing in firm size (e.g. Petersen and Rajan, 1994; Fama and French, 1995; Bao et al., 2011). The USPTO grants an even greater discount of 75% if applicants meet more strict requirements on the number of previous applications and, when applicable, the revenue of a company. In this case, they distinguish between "small" and "micro" applicants. However, since only around 70,000 applications meet the criteria for "micro" in our data, we report results both for a summarizing category of discounted applicants and for the full differentiation.

For the quality of the patent, we follow Sampat and Williams (2019). They argue that the quality measure needs to be known at the time of the patent application. This disqualifies *ex-post* measures such as forward citations (Trajtenberg, 1990), patent renewals (Schankerman and Pakes, 1986), patent litigation (Harhoff et al., 2003), or excess stock returns (Kogan et al., 2017). They settle on two quality indicators, one of which is immediately available from the USPTO: the number of claims made in a patent. This is also the quality indicator utilized by Farre-Mensa et al. (2020).

Another commonly considered characteristic of patent applications is the applications' scope. Kuhn and Thompson (2019) propose the length of the first claim in a patent application as an inverse measure of scope. The argument is that shorter claims imply fewer qualifiers or conditional statements, thus broadening the scope of the potential patent. Note that both in their own work and in the works of others, it is argued that this measure is different from application quality (e.g. Righi and Simcoe, 2019). However, scope might still increase the value of a patent in the sense of π_i in the model. We are thus uncertain whether π_i is captured only by the number of claims, or by the number of claims and the scope measure.

We estimate the equation

$$\text{Leniency}_{i,a,t} = \beta_1 \cdot \mathbf{1}(\text{Discounted}_{i,a,t}) + \beta_2 \cdot \text{No. Claims}_{i,a,t} + \beta_3 \cdot \text{Scope}_{i,a,t} + \nu_{a,t} + \varepsilon_{i,a,t}. \qquad (9)$$

Here, the leniency of the examiner of patent application i in art unit a and year t is a function of the small entity indicator, the number of claims and the applications scope, measured inversely by the length of the application's first claim. The estimation features art-unit-by-year fixed effects and the standard errors are clustered on the level of the art unit. Summary statistics for the dependent and independent variables across the 5.7 million analyzed applications are provided in Table A8 of the Appendix.

Table 6 shows the results of the estimation. The estimates show clear and highly statistically significant support for the hypotheses derived in the theoretical model. On average, an application made by a small entity gets assigned a 1.13 percentage point less lenient patent examiner. Similarly, the number of claims increases the leniency of the examiner such that an additional claim leads to a 0.003 percentage point higher leniency. While both point estimates are small, it has to be kept in mind that these are one-dimensional proxies for marginal cost of influencing the lottery and quality of the potential patent. Both concepts are, however, highly complex and multi-dimensional. Attenuation bias thus likely decreases the point estimates provided here. It should also be noted that the effect of a one standard-deviation shift in the number of claims on examiner leniency is about comparable to the necessary effect we calculate in Appendix D.2 for reproducing the coefficients of Sampat and Williams (2019) without any causal result.

The scope of the application is not statistically significantly associated with the leniency of the examiner. However, when controlling for the USPC subclass of the patent application, it becomes significant with the hypothesized negative sign (see Table A9 for details). This difference in the results emphasizes the ambiguity about the ability of scope to proxy for patent quality.

Less than 7% of the applicants who receive a discount are micro entities by the definition of the USPTO. These micro entities are even smaller than the typical small entities receiving a discount. Amon others, one condition to qualify is that the gross income of the entity cannot be more than three times the median household in the previous year. Such entities likely have even larger costs of capital than mere small entities. Our model would thus predict a lower leniency for their application. This prediction is supported in the data, both in the univariate and the multivariate analysis, as can be seen in columns (2) and (6) of Table 6. This lends further support for our model and the ultimate conclusion that the patent lottery can be manipulated by applicants (at a given cost).

From the argument of Righi and Simcoe (2019) and the descriptive analyses of Section 2, we know that the USPC subclass explains variability in the examiner leniency. For the purpose of our analysis, however, it is unclear whether this is variability we aim to explain using application characteristics or whether it should be projected out beforehand. If applicants (or their counsel) manipulate the patent lottery by strategically choosing claim language such that the patent application is given a USPC subclass with more lenient examiners, then we should not control for the effect of the subclass in our analysis. If the manipulation uses a different channel, then controlling for subclasses should not affect the results. As can be seen in Table A9 of the Appendix, including art unit $\times$ year $\times$ subclass fixed effects also leads to results consistent with our theoretical predictions. Because the coefficients for discounted applications are smaller than in Table 6, but those of

Table 6: Application-level Evidence for Model Hypotheses

	Dependent Variable: Examiner Leniency (in 100%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Discounted Appl.	-1.130^{***}				-1.132^{***}	
	(0.104)				(0.104)	
Small Entity		-1.089^{***}				-1.092^{***}
		(0.096)				(0.096)
Micro Entity		-1.837^{***}				-1.834^{***}
		(0.311)				(0.311)
No. Claims			0.002^{***}		0.003^{***}	0.003^{***}
			(0.001)		(0.001)	(0.001)
Application Scope				0.00001	0.00001	0.00001
				(0.00003)	(0.00003)	(0.00003)
Fixed effects	Art unit × Year	Art unit × Year	Art unit × Year	Art unit × Year	Art unit × Year	Art unit × Year
Clustered st. err.	Art unit	Art unit	Art unit	Art unit	Art unit	Art unit
Observations	5,708,617	5,708,617	5,708,617	5,708,617	5,708,617	5,708,617

Note: Table shows the results of estimating Equation (9) either in full (columns (5) and (6)) or with each explanatory variable in isolation (columns (1) through (4)). Examiner leniency is defined as the average probability of an application's examiner to grant a patent in the applications' year and art unit. The application is discounted if it comes from an individual, a nonprofit organization or a company with fewer than 500 employees. Micro entities get a larger discount than small entities. No. Claims is the number of claims listed at the time of the application. Scope is the number of characters in the first claim of the application at the initial time of application. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

quality indicators are larger, we can conclude that likely both strategic subclass assignment and alternative channels are at work.

6 Is there value in getting assigned a more lenient patent examiner?

Is there value in being assigned a more lenient patent examiner? Prior research establishes that obtaining a patent increases firm value (Pakes, 1985; Kogan et al., 2017) and that successful patent filings are associated with higher market valuations (Farre-Mensa et al., 2020). At the time an application and its examiner is published, information about the assigned examiner therefore has potential valuation implications. A more lenient examiner increases the probability of eventual patent grant and may thus raise expected firm value.

To operationalize this hypothesis, we construct a measure of *residualized examiner leniency* that isolates the idiosyncratic component of examiner assignment. We first compute each examiner's raw grant rate as the proportion of patent applications approved, as described in Section 4.1. To account for technology-specific heterogeneity, we demean each examiner's grant rate by the average grant rate within the corresponding Art Unit over the previous calendar year. This leads to comparably demeaned values as in the ranking analysis of Section 3. This residualized measure ensures that examiners are compared only against their immediate peers facing a similar quality of patent applications. For days or months in which a firm has multiple patent application disclosures, we aggregate the signal by taking the mean residualized leniency across all applicable examiners.

We merge these patent-level signals with firm identifiers for applicants using the matching tables provided by Arora et al. (2017) and Arora et al. (2021). Identifying the precise timing of information access is critical to ensuring that our trading strategy is strictly implementable and free from forward-looking bias. Consequently, we align our analysis with the *Visibility Date* established in Section 4.1 (see Figure 4), which corresponds to the later of the application's 18-month publication date or the First Office Action Date (FOAD). This date represents the first moment an investor can identify the examiner and compute an associated leniency score using prior disclosure dates. By conditioning the strategy on the Visibility Date, we ensure that the results represent a tradable process based on information that was fully available in the public domain before the return measurement period begins.

For the equity universe, we obtain daily and monthly stock returns from CRSP and restrict the sample to common equities (share codes 10 and 11) listed on the NYSE, NYSE MKT, NASDAQ, or ARCA. We apply standard liquidity and microstructure filters by excluding observations with stock prices below $1, daily trading volume below 100 shares, or arithmetic returns below -100%. Returns are adjusted for delistings following Shumway (1997).

The disclosure of a patent application provides the market with multiple layers of complex information that are subject to varying degrees of investor attention. Following Hirshleifer et al. (2018), we recognize that the technological content of an innovation, specifically its originality and potential for long-term profitability, is inherently difficult to evaluate and frequently leads to an initial underweighting of a firm's innovative strength. This informational friction is further compounded by the cognitive and strategic biases identified by Fitzgerald et al. (2021). They argue that investors exhibit a consistent preference for novelty, paying excessive attention to unfamiliar explorative patents while neglecting incremental exploitative innovations that build on a firm's existing competencies.

We argue that the assigned examiner's identity constitutes an additional, even less salient layer of administrative metadata that sits at the intersection of these frictions. While technological originality and search strategy determine the potential value of the underlying resource, the examiner's identity represents a leading indicator of the probability that this resource will successfully clear the regulatory hurdle and receive legal protection. Interpreting an examiner's identity as a signal requires comparing individual grant rates to contemporaneous Art Unit averages. If this comparison does not attract investor attention, the resulting change in perceived grant probability will be incorporated into prices only gradually over time.

To test whether the market fully incorporates this identity signal, we implement a monthly portfolio sorting strategy. At the end of each month $t-1$, we assign firms to terciles based on the mean residualized examiner leniency disclosed during that month. In month t, we form value-weighted portfolios and track their returns. This approach relies exclusively on publicly available metadata with a lag, ensuring that the results represent a strictly implementable trading strategy based on information available prior to the return measurement period. To ensure reliable portfolio averages, we require a minimum of 100 stocks per tercile-month, resulting in an average of around 420 stocks per month.

Appendix Table A10 reports average firm characteristics across the residualized-leniency terciles. Firms in the highest-leniency tercile exhibit somewhat higher book-to-market ratios, gross profitability, and investment, and somewhat lower cash holdings and idiosyncratic volatility, than firms in the lowest-leniency tercile. At the same time, the magnitudes of these differences are generally modest, suggesting that the return spread documented below is unlikely to be explained entirely by extreme differences in standard firm characteristics. At the beginning of the sample, the value-weighted mean market capitalization of firms in our sample corresponds to the 89^{th} value-weighted percentile of the CRSP universe, indicating that the sample is tilted toward relatively large firms.

Table 7 reports the results. Monthly excess returns increase monotonically across the leniency terciles. The portfolio of firms assigned to the strictest examiners (T1) earns a mean monthly return of 106.60 basis points ($t = 3.62$), while the most lenient tercile (T3) earns 149.65 basis points ($t = 4.81$). The resulting T3 minus T1 long-short portfolio generates a mean excess return of 43.05 basis points per month ($t = 2.89$). Because the long-short portfolio contrasts two groups of firms that both experience the publication of patent applications, it removes the general effect of publication and helps isolate the return differential associated specifically with examiner leniency.[21]

[21] A concern is that visibility may coincide with favorable prosecution news rather than merely examiner disclosure. However, 78.3% of first visibility events are non-final rejections and 19.8% are restriction requirements, a procedural request to elect claims for examination; thus, about 98.1% of first events are adverse or procedural rather than favorable.

Table 7: Value-weighted mean returns and alphas of trading strategy

	T1	T2	T3	T3-T1
Mean Returns (bps)	106.60***	117.34***	149.65***	43.05***
	(29.45)	(33.94)	(31.10)	(14.91)
CAPM Alpha (bps)	10.95	12.17	46.05***	35.10**
	(12.15)	(11.52)	(10.73)	(15.15)
FF3 Alpha (bps)	6.78	5.63	42.06***	35.28**
	(10.60)	(8.71)	(11.30)	(15.20)
FF3 + MOM Alpha (bps)	4.56	6.67	40.36***	35.80**
	(10.39)	(8.56)	(11.54)	(15.27)

Note: This table reports value-weighted average monthly excess returns and risk-adjusted alphas for portfolios sorted by residualized examiner leniency. At the end of each month t, firms are assigned to terciles based on the mean residual leniency of their assigned patent examiners. Residual leniency is defined as the examiner's grant rate demeaned by Art Unit over the previous calendar year. LS denotes a zero-cost portfolio that is long the most lenient tercile (T3) and short the strictest tercile (T1). Alphas are reported in basis points. Risk adjustments include the CAPM, Fama-French three-factor (FF3), and FF3 plus momentum (FF3+MOM) models. Newey and West (1986) standard errors with a bandwidth of 5 are reported in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Crucially, this return spread is not explained by standard risk factors. After adjusting for market, size, value, and momentum exposures, the strategy yields a statistically significant alpha of 35.80 basis points per month ($t = 2.34$). The existence of this significant alpha suggests that the market underreacts to the implications of examiner assignment at the time of disclosure. These findings connect our setting to a broader literature on limited attention and information-processing constraints. Hirshleifer et al. (2018) show that innovative originality elicits delayed price reactions. Our evidence suggests that even administrative metadata can generate similar dynamics when extracting its value implications requires costly aggregation and benchmarking. In this sense, examiner assignment represents another difficult-to-process signal that requires extensive expertise to evaluate. More broadly, our results echo evidence from settings where investors focus on salient top-line news while ignoring subtle, value-relevant details, leading to predictable post-event drifts (e.g., Bernard and Thomas, 1989; Sloan, 1996; Loughran and McDonald, 2011).[22]

Our finding of a positive return spread at disclosure contrasts with the negative relation of leniency and stock returns documented by Shu et al. (2022) following patent issuance. The two findings, however, do not contradict each other. Shu et al. (2022) consider stock returns conditional on patent issuance, while we consider stock returns for all patent applications. The primary driver in their study is that granted patents by lenient examiners are, on average, of lower qual-

[22] Appendix Figure A5 shows that portfolio membership varies substantially from month to month. This degree of turnover makes it less likely that the return spread is driven solely by a fixed set of firms with persistently different risk exposures. Instead, the evidence is more consistent with mispricing of time-varying information in examiner assignment that is only gradually incorporated into stock prices.

ity because, given a fixed set of applications, lenient examiners approve a larger share of them ("rubber-stamping"). In contrast, our analysis emphasizes the increased probability of obtaining a patent in the first place as the main valuation channel. Appendix C shows formally that both effects can coexist and are therefore not contradictory. If anything, our findings complement those of Shu et al. (2022) by suggesting that markets initially underreact to information about examiner leniency. This is also in line with other studies such as Hirshleifer et al. (2009), which indicate that markets may not immediately incorporate the implications of complex information, such as patent applications, examiner leniency, and the current stage of the application process. Consistent with this interpretation, Table A11 in the Appendix shows that the abnormal returns are concentrated in the first post-disclosure month and attenuate quickly at longer holding horizons. Taken together, the portfolio results suggest that examiner assignment provides a setting in which complex, non-salient administrative information is only gradually impounded into financial markets. A strategy based on public patent metadata can therefore generate significant risk-adjusted returns before this information is fully reflected in prices.

7 Discussion

7.1 Welfare implications

Our results naturally invite the question of how a manipulable patent lottery affects social welfare. The theoretical considerations in Section 5.1 only model the private consequences of the patent application for the firm. In order to identify the societal consequences, we require two additional concepts. Let v_i be the gross societal value for developing innovation i. Further, let d_i be the deadweight loss associated with a patent protection for this innovation. It appears with probability p_i which is the probability of obtaining a patent for the innovation. As before, let c_i be the cost of developing the innovation. For the two firms in our model, we can then describe the social welfare of innovation as

$$W = \sum_{i \in 1,2} \mathbf{1}(i \text{ innovates}) \left(v_i - c_i - p_i d_i - k_i e_i \right). \tag{10}$$

Here, $v_i - c_i$ is the social surplus of the innovation and $p_i d_i$ is the expected monopoly distortion. As in the standard Nordhaus (1969) framework, some expected monopoly distortion is the price society pays for providing innovation incentives. Lastly, $k_i e_i$ is pure rent-seeking waste.

Absent manipulation it holds that $p_i = \bar{p}$ and $e_i = 0$ for both firms. If we assume that both firms innovate under both types of regimes, then welfare without manipulation is given as $W^0 = \sum_{i\in\{1,2\}}(v_i - c_i - \bar{p}d_i)$ and welfare with manipulation is given by

$$W^M = \sum_{i\in\{1,2\}} (v_i - c_i - p_i^* d_i - k_i e_i^*). \tag{11}$$

Here, $p_i^* = p(e_i^*, e_{-i}^*)$ is the result of the Nash equilibrium from Proposition 1. By the zero sum condition, we have $p_1^* + p_2^* = 2\bar{p}$ which provides us with the key result that

$$W^M - W^0 = -(p_1^* - \bar{p})(d_1 - d_2) - (k_1 e_1^* + k_2 e_2^*). \tag{12}$$

Equation (12) shows that, when both firms innovate under both regimes, manipulation has two effects. First, it reallocates patent rights across firms while leaving the average grant probability unchanged. Second, it generates resource waste through equilibrium manipulation expenditures. If dead-weight loss is similar across firms, or larger for the firms that are more successful at manipulation, welfare falls unambiguously. This is plausible if innovations with larger private patent rents also tend to generate larger monopoly distortions. In such cases, manipulation moves grants toward firms that are better at extracting monopoly rents, not toward socially better innovations. Nevertheless, there is the chance that welfare can be affected positively by manipulation. This will happen, if the manipulation efforts lead to a higher chance of a patent for the innovation with less dead-weight loss due to patent protection and this effect is larger than the joint cost of the manipulation efforts of both firms.

A more plausible route through which manipulation could increase welfare arises at the extensive margin. Suppose one firm would not innovate in the absence of manipulation because $\bar{p}\pi_i < c_i$. If the possibility of manipulation raises its effective probability of obtaining a patent, it may induce that firm to innovate. Welfare could then rise if the additional innovation generates enough social surplus and if the resulting manipulation costs remain limited. By contrast, welfare falls if manipulation instead deters socially valuable but manipulation-disadvantaged firms from innovating. How realistic these scenarios are depends on the individual applications. However, it should be kept in mind that because equilibrium effort responds to π_i and k_i, the manipulatable system favors firms with large private patent rents and low manipulation costs, not necessarily firms with the largest social surplus.

An alternative channel for welfare improvement could appear through risk aversion of applicants. In the model, applicants are risk neutral and make their decisions to innovate and manipulate the patent lottery in an effort maximize the expected value of their outcomes. If they were risk averse, the potential to influence the outcome distribution of the patent lottery through manipulation could represent additional value. However, this need not actually be the case. The effect of risk aversion on the decision to influence outcome distributions is not monotonic (Peter, 2021). In the present case, risk aversion could also lead to less manipulation rather than to more. We thus opt to model applicants as risk neutral in order to not complicate welfare considerations further.

Patent lottery manipulation may reduce welfare through an additional channel not captured by the standard incentive-distortion trade-off, namely fairness. A large literature shows that individuals value fair outcomes and fair procedures (Fehr and Schmidt, 1999; Konow, 2003; Brock et al., 2013). It is thus likely that they value fairness in public administration systems, as well. Whether the patent lottery is fair in its current, manipulable form depends on the normative viewpoint. Under a strictly egalitarian point of view, it is certainly not fair, because different applicants have different chances of obtaining a patent. One might argue from a more libertarian point of view that the procedure is fair insofar as all applicants formally face the same rules. This defense is weakened, however, when applicants differ systematically in their practical ability to influence the lottery. Not all applicants have the same possibility to influence the system, because they might differ in their ability to hire a high-quality law firm, pay for a sophisticated prediction system or build up a stock of revolving door employees. From such a liberal egalitarian perspective, a manipulable patent lottery reduces welfare because it undermines the legitimacy of the allocation process by making success depend partly on unequal capacities to influence the system.[23] Cappelen et al. (2007) estimate that over 80% of their subjects are either strict or liberal egalitarians. To the extent that individuals derive utility from fair procedures in a similar way, a manipulable patent lottery reduces welfare even apart from its efficiency consequences, because it allocates patent rights in a way that depends on unequal capacities to influence the system.

7.2 Implications for empirical identification

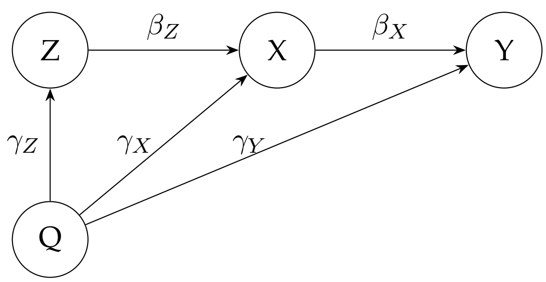
What do the results of this study imply for research using the patent lottery as a source of exogenous variation? Such research uses the patent examiner leniency (henceforth Z) as an instrument for the probability of a granted patent (henceforth X). This is done to examine the effect of a patent grant on some outcome Y. This can be a diverse set of outcomes such as follow-up innovation (Sampat and Williams, 2019), economic outcomes of a firm (Farre-Mensa et al., 2020) or

[23] While we do not model this channel in explicitly in Equation (10), it is possible to do so. One could add an additional term $-\Phi$ after the sum where $\Phi(|p_i^* - \bar{p}|, |k_1 - k_2|)$ is a function of both outcomes of the manipulation process and the difference in the ability to influence the system. We refrain from doing so for the sake of brevity.

the career development of inventors (Melero et al., 2020). The instrument is used, because it is generally acknowledged that unobserved characteristics of the patent application (henceforth Q) influence the grant rate X as well as the outcome Y. The results in this paper make clear that Q can have an influence on Z as well. This calls the identification strategy into question.

The typical data-generating process of studies using an examiner leniency IV is depicted in Figure 7. Studies are interested in the value of β_X and use Z as an instrument to estimate it. They do so being mindful of the existence of Q as well as taking non-zero values for γ_X and γ_Y into account. Crucially, though, the assumption that examiner assignment and thus examiner leniency is random: they assume that $\gamma_Z = 0$. This is at odds with our results which imply that $\gamma_Z \neq 0$.

Figure 7: Causal Influence Diagram for Patent Examiner Leniency IVs



Note: The diagram shows the causal influences considered in a typical patent examiner randomness designs, but adds the quality dimension Q and its potential influence on examiner leniency (Z), the probability of a granted patent (X) and the final outcome (Y).

To analyze the effect of γ_Z on the identification strategy analytically, we can formalize this data-generating process under the assumption that other than the influences depicted in Figure 7, all randomness is mutually independent and mean zero. Specifically, we can set

$$Z = \gamma_Z Q + \varepsilon_Z \tag{13}$$

$$X = \gamma_X Q + \beta_Z Z + \varepsilon_X \tag{14}$$

$$Y = \gamma_Y Q + \beta_X X + \varepsilon_Y \tag{15}$$

where $Q, \varepsilon_Z, \varepsilon_X$, and ε_Y are mutually independent, mean zero random variables with finite variances.

For the just-identified 2SLS regression of Y on X using Z as an instrument, the population IV estimand is

$$\beta^{IV} = \text{plim } \hat{\beta}_{2SLS} = \frac{Cov(Z, Y)}{Cov(Z, X)}. \tag{16}$$

As we show in Appendix D.1, the DGP implies that

$$\text{plim } \hat{\beta}_{2SLS} = \beta_X + \frac{\gamma_Y \gamma_Z \text{Var}(Q)}{\gamma_X \gamma_Z \text{Var}(Q) + \beta_Z(\gamma_Z^2 \text{Var}(Q) + \text{Var}(\varepsilon_Z))}. \tag{17}$$

Equation (17) shows how the common factor leads to a violation of the exclusion restriction and thus results in a bias of the estimated coefficient for β_X. The bias disappears with $\gamma_Z = 0$ or $\gamma_Y = 0$, but as long as the common factor influences both examiner leniency and the outcome of interest, there will be a bias. Importantly, the bias is additive to the value of β_X. As such, patent examiner leniency IVs can potentially find statistically significant results even in absence of an actual effect.

We next explore numerically under which circumstances a faulty rejection of the null hypothesis can appear. For the sake of simplicity, we assume that $Q, \varepsilon_Z, \varepsilon_X$, and ε_Y are distributed $N(0,1)$. We also set $\beta_Z = 0.8$, which is a commonly reported magnitude for the influence of leniency on grant rate. Further, we set $\gamma_X = \gamma_Y = 0.4$ for the (in the literature undisputed) relationship of the common factor with grant rate and outcome. Lastly, we assume $\beta_X = 0$. Thus, any statistically significant effect detected in a 2SLS estimation will be due to the bias, the hypothesized effect of patent grant on outcome does not exist. Under these circumstances we ask how big γ_Z or the sample size N have to be in order to erroneously reject the null hypothesis that $\beta_X = 0$ at the conventional significance level of 5%.

Results of our analysis are shown in Figure 8. Panel (a) displays the rejection rate in 1,000 simulations as a function of the ratio γ_Z/γ_X for three different sample sizes. When the effect of Q on leniency is only one hundredth of the effect of Q on the grant rate, there is only a small chance of erroneously rejection the null hypothesis for small and medium-sized samples. However, as γ_Z increases, this becomes increasingly more likely such that at for medium sample sizes of 100,000 observations (remember that the USPTO provides data on over 8 million patents since 2001), even one twentieth of the effect on the grant rate is sufficient to falsely claim a nonzero β_X more than half the time.[24] If γ_Z is half the size of γ_X, the null hypothesis gets rejected virtually always for

[24]While the USPTO records cover millions of patent applications, leniency-IV studies typically operate on substantially smaller analysis samples once the patent data is merged with the external datasets needed to construct the outcome of interest. Sampat and Williams (2019) use 293,652 application-gene observations after restricting to patents claiming human-gene DNA sequences and linking to PubMed, Pharmaprojects, and clinical-trials records. Farre-Mensa et al. (2020) analyze 34,215 first-time patent applications filed since 2001 by stand-alone, for-profit U.S. small-business entities, with outcomes drawn from Dun & Bradstreet's NETS, the USPTO patent assignment database, VentureXpert, and Thomson-Reuters SDC. Melero et al. (2020) use 129,740 inventor-application-year observations restricted to early-career inventors at firms covered by Standard & Poor's Capital IQ. Each of these sample sizes falls within the range where Figure 8 indicates non-trivial false-rejection rates for moderate values of γ_Z.

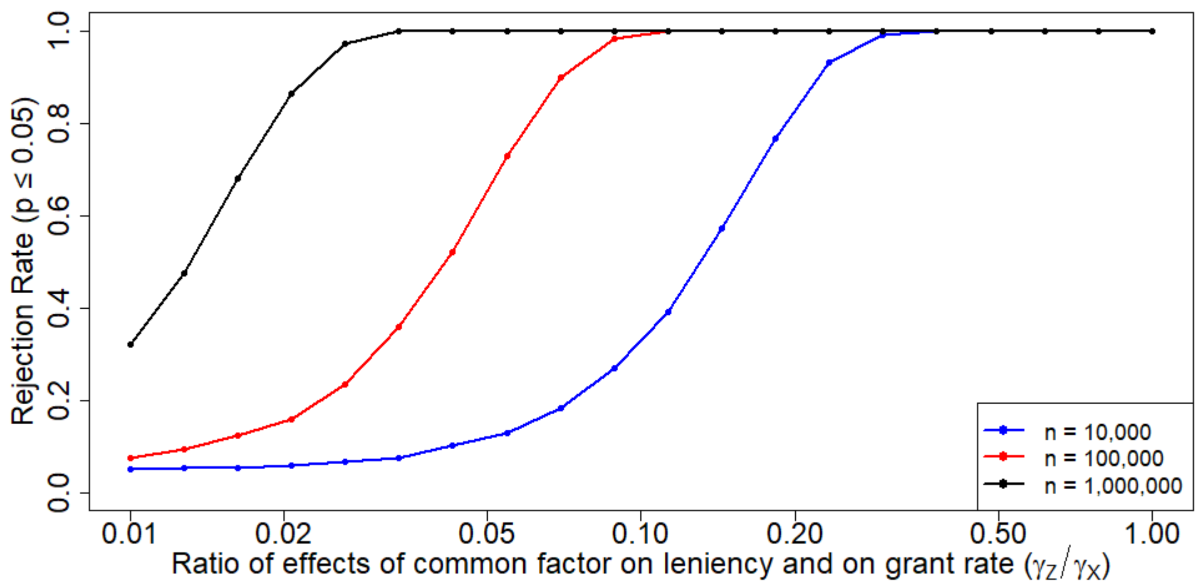
all analyzed sample sizes. It should also be noted that for samples with one million observations, even the smallest values of γ_Z already lead to notable rejection rates.

Panel (b) of Figure 8 considers the effect of sample size in more detail. When γ_Z and γ_X are equally larges (black line), even small sample sizes of 500 observations lead to rejection rates over 80%. If the effect on leniency is one tenth of that on the grant rate (red line), a sample size of 20,000 is sufficient for a rejection rate over 50%. For very small effects (blue line), the sample size has to become large for high rejection rates, but the threshold is still below the number of patent applications in the USPTO data since 2001.

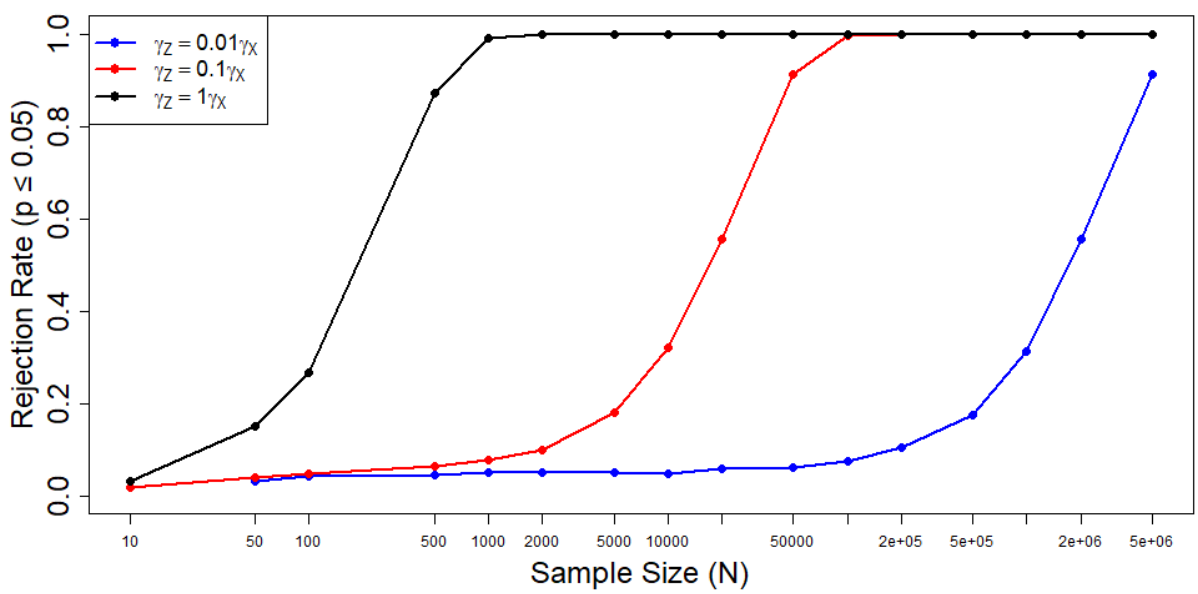
The bias derived in Equation (17) and the numerical results from Figure 8 show the vulnerability of patent examiner leniency IVs to manipulation in the patent lottery. If firms can strategically invest into manipulating the patent lottery, then any empirically observed effect of a patent being granted could instead represent unobserved characteristics of the invention that drive firms' strategies. Sampat and Williams (2019), who were the first to implement such a design, already note an influence of ex-ante application quality on the leniency of the examiner. They provide a test of the quality dimension and argue that based on this test, its importance can be discarded – a conclusion at odds with our results. In Appendix D.2 we reconsider their empirical set-up, their test of the quality dimension, and their empirical identification strategy. We argue that given the values provided in the paper, one can construct a data generating process in which there is no influence of a patent grant on follow-up innovation but the identification strategy nevertheless results in the same coefficients they report in their paper.

Figure 8: Distribution of Predictive Quality across Art Units

(a) Effect of γ_Z as a multiple of γ_X.



(b) Effect of sample size (N).



The figure shows the rejection rate of the correct null hypothesis that $\beta_X = 0$. Data is simulated according to the DGP given in Equations (13) though (15) and illustrated in Figure 7. In the simulations, we set $\beta_X = 0$, $\beta_Z = 0.8$, $\gamma_X = \gamma_Y = 0.4$ and $Q, \varepsilon_Z, \varepsilon_X$, and ε_Y as iid $N(0, 1)$. Rejection rate is reported at the 5% level based on 1,000 simulations.

8 Conclusions

We examine whether the ostensibly random assignment of patent examiners – the "patent lottery" – can be gamed and what this implies for innovation and empirical research. Using 5.6 million U.S. applications from 2001–2020, we construct the null distribution of examiner leniency and show that assignment patterns at the law firm level deviate significantly from randomization. Larger entities receive examiners who are, on average, 1.13 percentage points more lenient, applications with more claims are matched to more lenient examiners, and publicly available information enables predictive timing. A simple trading strategy based on examiner leniency earns positive abnormal returns. A model with heterogeneous examiner quality and costly manipulation rationalizes these facts.

These findings matter for both policy and research. Fair and independent assignment is a cornerstone of the patent system; systematic deviations shift approval probabilities toward sophisticated repeat players and away from small innovators, altering the distribution of innovative rents. As we show, this can affect social welfare in non-trivial ways. Our results also challenge a common identification strategy that instruments patent grants with examiner leniency under the assumption of random assignment, highlighting the risk of spurious inference when assignment is partly endogenous.

Our analysis has limitations. While we document convergent patterns consistent with manipulation and provide strong predictive evidence, we do not directly observe all channels of influence or the assignment algorithm's internal logic. Measures of examiner leniency and attorney identity, while carefully constructed, are subject to measurement error. The return predictability we document is based on historical backtests and may be attenuated in real time.

The broader implications are nevertheless important. If assignment can be influenced at scale, then the patent system may inadvertently reward influence rather than invention, with consequences for entry, competition, and the geography of innovation. Transparent, auditable randomization and periodic assignment audits could restore confidence in the system and improve the external validity of research designs that rely on examiner heterogeneity.

Several directions merit future work. First, evaluate policy reforms – for example, stricter automated randomization, throttling of filing-time discretion, or routine anomaly detection – using pilot programs or staggered rollouts. Second, develop designs that exploit plausibly exogenous shocks (e.g., staffing or procedural changes) to pin down causal channels. Third, quantify welfare and distributional effects across firm size, technology, and regions. Fourth, extend the analysis to other jurisdictions and to downstream outcomes such as litigation, licensing, and market entry. Finally, explore assignment mechanisms that preserve workload balancing while hardening randomness against strategic behavior.

References

American Bar Association (2024). 2023–24 national lawyer population survey. Survey report, American Bar Association, Center for Bar Leadership.

Appel, I., J. Farre-Mensa, and E. Simintzi (2019). Patent trolls and startup employment. Journal of Financial Economics 133(3), 708–725.

Arnold, D., W. Dobbie, and C. S. Yang (2018). Racial bias in bail decisions. The Quarterly Journal of Economics 133(4), 1885–1932.

Arora, A., S. Belenzon, and L. Sheer (2017). Back to basics: Why do firms invest in research? Working paper 23187, National Bureau of Economic Research.

Arora, A., S. Belenzon, and L. Sheer (2021). Matching patents to compustat firms, 1980–2015: Dynamic reassignment, name changes, and ownership structures. Research Policy 50(5), 104217.

Azmat, G. and R. Ferrer (2017). Gender gaps in performance: Evidence from young lawyers. Journal of Political Economy 125(5), 1306–1355.

Bao, J., J. Pan, and J. Wang (2011). The illiquidity of corporate bonds. The Journal of Finance 66(3), 911–946.

Bernard, V. L. and J. K. Thomas (1989). Post-earnings-announcement drift: Delayed price response or risk premium? Journal of Accounting Research 27, 1–36.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: Robust and efficient estimation. Review of Economic Studies 91(6), 3253–3285.

Botoman, A. (2017). Divisional judge-shopping. Columbia Human Rights Law Review 49, 297.

Briscoe, F. and M. Rogan (2016). Coordinating complex work: Knowledge networks, partner departures, and client relationship performance in a law firm. Management Science 62(8), 2392–2411.

Brock, J. M., A. Lange, and E. Y. Ozbay (2013). Dictating the risk: Experimental evidence on giving in risky environments. American Economic Review 103(1), 415–437.

Budish, E., B. N. Roin, and H. Williams (2016). Patents and research investments: Assessing the empirical evidence. American Economic Review 106(5), 183–187.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007). The pluralism of fairness ideals: An experimental approach. American Economic Review 97(3), 818–827.

Cohen, A. and C. S. Yang (2019). Judicial politics and sentencing decisions. American Economic Journal: Economic Policy 11(1), 160–191.

Cohen, L., U. G. Gurun, and S. D. Kominers (2019). Patent trolls: Evidence from targeted firms. Management Science 65(12), 5461–5486.

Dayani, A. (2023). Patent trolls and the market for acquisitions. Journal of Financial and Quantitative Analysis 58(7), 2928–2958.

Fama, E. F. and K. R. French (1995). Size and book-to-market factors in earnings and returns. The Journal of Finance 50(1), 131–155.

Farre-Mensa, J., D. Hegde, and A. Ljungqvist (2020). What is a patent worth? Evidence from the us patent "lottery". The Journal of Finance 75(2), 639–682.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. The Quarterly Journal of Economics 114(3), 817–868.

Fitzgerald, T., B. Balsmeier, L. Fleming, and G. Manso (2021). Innovation search strategy and predictable returns. Management Science 67(2), 1109–1137.

Gardner, J. (2022). Two-stage differences in differences. arXiv preprint arXiv:2207.05943.

Goldsmith-Pinkham, P., P. Hull, and M. Kolesár (2025). Leniency designs: An operator's manual.

Graham, S. J., A. C. Marco, and R. Miller (2018). The uspto patent examination research dataset: A window on patent processing. Journal of Economics & Management Strategy 27(3), 554–578.

Harhoff, D., F. M. Scherer, and K. Vopel (2003). Citations, family size, opposition and the value of patent rights. Research policy 32(8), 1343–1363.

Henderson, M. T., I. Hutton, D. Jiang, and M. Pierson (2025). Lawyer ceos. Journal of Financial and Quantitative Analysis 60(2), 580–616.

Hirshleifer, D., P.-H. Hsu, and D. Li (2018). Innovative originality, profitability, and stock returns. The Review of Financial Studies 31(7), 2553–2605.

Hirshleifer, D., S. S. Lim, and S. H. Teoh (2009). Driven to distraction: Extraneous events and underreaction to earnings news. The Journal of Finance 64(5), 2289–2325.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance 48(1), 65–91.

Kahan, M. and T. A. McKenzie (2021). Judge shopping. Journal of Legal Analysis 13(1), 341–379.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. The Quarterly Journal of Economics 132(2), 665–712.

Konow, J. (2003). Which is the fairest one of all? a positive analysis of justice theories. Journal of Economic Literature 41(4), 1188–1239.

Kuhn, J. M. and N. C. Thompson (2019). How to measure and draw causal inferences with patent scope. International Journal of the Economics of Business 26(1), 5–38.

Lemley, M. A. and B. Sampat (2012). Examiner characteristics and patent office outcomes. Review of Economics and Statistics 94(3), 817–827.

LexisNexis Intellectual Property Solutions (2026). Tech Center Navigator. https://www.lexisnexisip.com/solutions/patent-drafting/tech-center-navigator/. Accessed: 2026-05-09.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. The Journal of Finance 66(1), 35–65.

Melero, E., N. Palomeras, and D. Wehrheim (2020). The effect of patent protection on inventor mobility. Management Science 66(12), 5485–5504.

Merges, R. P. (1999). As many as six impossible patent before breakfast: Property rights for business concepts and patent system reform. Berkeley Tech. LJ 14, 577.

Nordhaus, W. D. (1969). Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change. MIT Press.

Pakes, A. (1985). On patents, R&D, and the stock market rate of return. Journal of Political Economy 93(2), 390–409.

Peter, R. (2021). Who should exert more effort? risk aversion, downside risk aversion and optimal prevention. Economic Theory 71(4), 1259–1281.

Petersen, M. A. and R. G. Rajan (1994). The benefits of lending relationships: Evidence from small business data. The Journal of Finance 49(1), 3–37.

Raffiee, J., F. Teodoridis, and D. Fehder (2023). Partisan patent examiners? Exploring the link between the political ideology of patent examiners and patent office outcomes. Research Policy 52(9), 104853.

Rehavi, M. M. and S. B. Starr (2014). Racial disparity in federal criminal sentences. Journal of Political Economy 122(6), 1320–1354.

Righi, C. and T. Simcoe (2019). Patent examiner specialization. Research Policy 48(1), 137–148.

Rosen, J. B. (1965). Existence and uniqueness of equilibrium points for concave n-person games. Econometrica, 520–534.

Sampat, B. and H. L. Williams (2019). How do patents affect follow-on innovation? Evidence from the human genome. American Economic Review 109(1), 203–236.

Schankerman, M. and A. Pakes (1986). Estimates of the value of patent rights in european countries during the post-1950 period. The Economic Journal 96(384), 1052–1076.

Shu, T., X. Tian, and X. Zhan (2022). Patent quality, firm value, and investor underreaction: Evidence from patent examiner busyness. Journal of Financial Economics 143(3), 1043–1069.

Shumway, T. (1997). The delisting bias in CRSP data. The Journal of Finance 52(1), 327–340.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? Accounting Review, 289–315.

Tabakovic, H. and T. G. Wollmann (2024). From revolving doors to regulatory capture? Evidence from patent examiners. Working paper 24638, National Bureau of Economic Research. First published May 2018.

Trajtenberg, M. (1990). Economic analysis of product innovation: The case of CT scanners, Volume 160. Harvard University Press.

Appendix A Proofs

A.1 Proof of Proposition 1

Proof. We can collect the first order conditions in a system $F(e_1, e_2) = 0$ with $F_1(e_1, e_2) = p_1(e_1, e_2) - \frac{k_1}{\pi_1} = 0$ and $F_2(e_1, e_2) = p_1(e_2, e_1) - \frac{k_2}{\pi_2} = 0$. The Jacobian is then given by

$$J = \begin{bmatrix} p_{11}(e_1, e_2) & p_{12}(e_1, e_2) \\ p_{12}(e_2, e_1) & p_{11}(e_2, e_1) \end{bmatrix}. \tag{A1}$$

Defining $S := \frac{1}{2}(J + J^\top)$ we can see that

$$S = \begin{bmatrix} p_{11}(e_1, e_2) & \frac{1}{2}(p_{12}(e_1, e_2) + p_{12}(e_2, e_1)) \\ \frac{1}{2}(p_{12}(e_1, e_2) + p_{12}(e_2, e_1)) & p_{11}(e_2, e_1) \end{bmatrix} \tag{A2}$$

$$= \begin{bmatrix} p_{11}(e_1, e_2) & 0 \\ 0 & p_{11}(e_2, e_1) \end{bmatrix}. \tag{A3}$$

The second line follows because the zero-sum nature of the manipulation game ensures that $p(e_1, e_2) + p(e_2, e_1) = 2\bar{p}$. Taking the derivative of both sides first with respect to e_1 and then with respect to e_2 renders

$$p_{12}(e_1, e_2) + p_{12}(e_2, e_1) = 0. \tag{A4}$$

From this we can see that S is own-concave by $p_{11} < 0$ and that it has a positive determinant

$$p_{11}(e_1, e_2)p_{11}(e_2, e_1) > 0. \tag{A5}$$

Thus, by Rosen (1965), (8) identifies the unique Nash equilibrium as long as optimal expenses are not given by the corner solution of $e_i^* = 0$. The latter can be guaranteed by $p_1(0, 0) > \max\left\{\frac{k_1}{\pi_1}, \frac{k_2}{\pi_2}\right\}$. □

A.2 Proof of Proposition 2

Proof. Due to Proposition 1, we know that we study an interior solution with the first order conditions given by (8). Define the system of first-order conditions as a mapping $F : \mathbb{R}^2 \times \mathbb{R}^4_{++} \to \mathbb{R}^2$:

$$F_1(e_1, e_2; k_1, k_2, \pi_1, \pi_2) := p_1(e_1, e_2) - \frac{k_1}{\pi_1} = 0, \tag{A6}$$

$$F_2(e_1, e_2; k_1, k_2, \pi_1, \pi_2) := p_1(e_2, e_1) - \frac{k_2}{\pi_2} = 0. \tag{A7}$$

Let $A := \partial F/\partial(e_1, e_2)$ evaluated at (e_1^*, e_2^*). This is equivalent to $J(e_1^*, e_2^*)$ from (A1). From the proof of Proposition 1 we know that $p_{12}(e_1^*, e_2^*) = -p_{12}(e_2^*, e_1^*)$. Thus

$$\det(A) = p_{11}(e_1^*, e_2^*)\, p_{11}(e_2^*, e_1^*) - p_{12}(e_1^*, e_2^*)\, p_{12}(e_2^*, e_1^*) \tag{A8}$$

$$= p_{11}(e_1^*, e_2^*)\, p_{11}(e_2^*, e_1^*) + (p_{12}(e_1^*, e_2^*))^2 > 0. \tag{A9}$$

Let $E = (e_1^*, e_2^*)^\top$. The Implicit Function Theorem (IFT) implies that, in a neighborhood where A is nonsingular, E is a continuously differentiable function of (k_1, k_2) and

$$\frac{\partial E}{\partial k_i} = -A^{-1}\frac{\partial F}{\partial k_i}, \qquad i \in \{1, 2\}. \tag{A10}$$

Focusing on k_1, the relevant parameter derivative is $\frac{\partial F}{\partial k_1} = (-1/\pi_1, 0)^\top$. Solving the system of linear equations in (A10) renders

$$\frac{\partial e_1^*}{\partial k_1} = \frac{p_{11}(e_2^*, e_1^*)}{\pi_1 \det A} \qquad \text{and} \tag{A11}$$

$$\frac{\partial e_2^*}{\partial k_1} = -\frac{p_{12}(e_2^*, e_1^*)}{\pi_1 \det A}. \tag{A12}$$

From this, we can consider the comparative static of interest

$$\frac{\partial p(e_1^*, e_2^*)}{\partial k_1} = p_1(e_1^*, e_2^*)\frac{\partial e_1^*}{\partial k_1} + p_2(e_1^*, e_2^*)\frac{\partial e_2^*}{\partial k_1} \tag{A13}$$

$$= \frac{p_1(e_1^*, e_2^*)p_{11}(e_2^*, e_1^*) - p_2(e_1^*, e_2^*)p_{12}(e_2^*, e_1^*)}{\pi_1 \det A} \tag{A14}$$

We can see that this is negative if $p_{12}(e_1^*, e_2^*)$ is not too negative, because that implies that $p_{12}(e_2^*, e_1^*)$ is not too positive. The analogous result can be shown for $\frac{\partial p(e_2^*, e_1^*)}{\partial k_2}$.

Treating the benefit parameter analogously renders $\frac{\partial E}{\partial \pi_i} = -A^{-1}\frac{\partial F}{\partial \pi_i}$ for $i \in \{1, 2\}$. The relevant parameter derivative is $\frac{\partial F}{\partial \pi_1} = \left(\frac{k_1}{\pi_1^2}, 0\right)^\top$. Solving the system of linear equations again renders

$$\frac{\partial e_1^*}{\partial \pi_1} = -\frac{k_1 p_{11}(e_2^*, e_1^*)}{\pi_1^2 \det A} \qquad \text{and} \tag{A15}$$

$$\frac{\partial e_2^*}{\partial \pi_1} = \frac{k_1 p_{12}(e_2^*, e_1^*)}{\pi_1^2 \det A}. \tag{A16}$$

And analogously $\frac{\partial e_1^*}{\partial \pi_2} < 0$ and $\frac{\partial e_2^*}{\partial \pi_2} > 0$. By a similar argument as before

$$\frac{\partial p(e_1^*, e_2^*)}{\partial \pi_1} = \frac{-p_1(e_1^*, e_2^*)k_1 p_{11}(e_2^*, e_1^*) + p_2(e_1^*, e_2^*)k_1 p_{12}(e_2^*, e_1^*)}{\pi_1^2 \det A} \tag{A17}$$

Which is positive if $p_{12}(e_1^*, e_2^*)$ is not too negative implying that $p_{12}(e_2^*, e_1^*)$ is not too positive. The analogous result can be shown for $\frac{\partial p(e_2^*, e_1^*)}{\partial \pi_2}$. $\square$

Appendix B Additional Details on Empirical Analyses

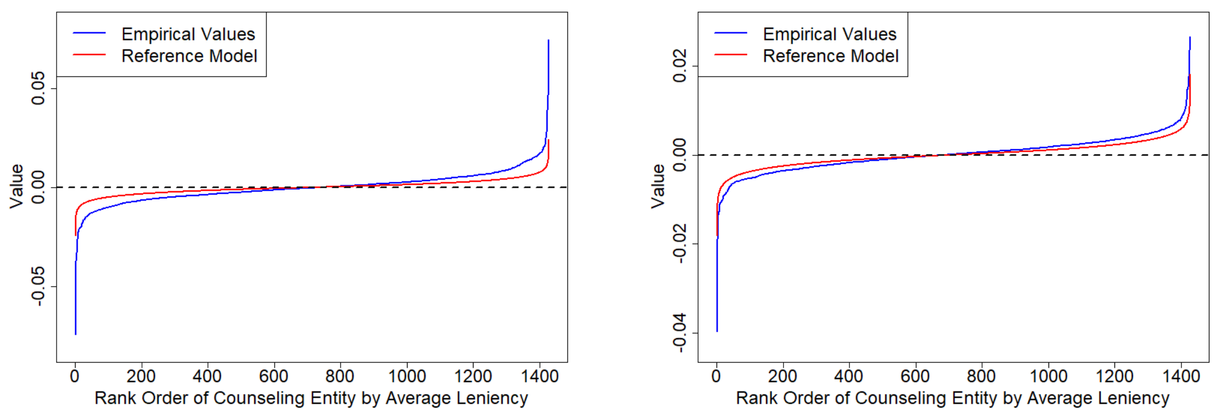
B.1 Additional Details for Section 3

Figure A1 shows the same data as Figure 3 but includes all 1,427 observations, including the seven most extreme observations in both tails. We can see that the tails in the empirical values are much more pronounced that in the reference models (particularly in panel (a)). They thus emphasize the conclusion that the empirical values are different from the reference models.

Figure A1: Comparison of Empirically Observed Law Firm Effects with Reference Model of Random Examiner Leniency

(a) Art-unit × Year Residual

(b) Art-unit × Year × USPC Subclass Residual

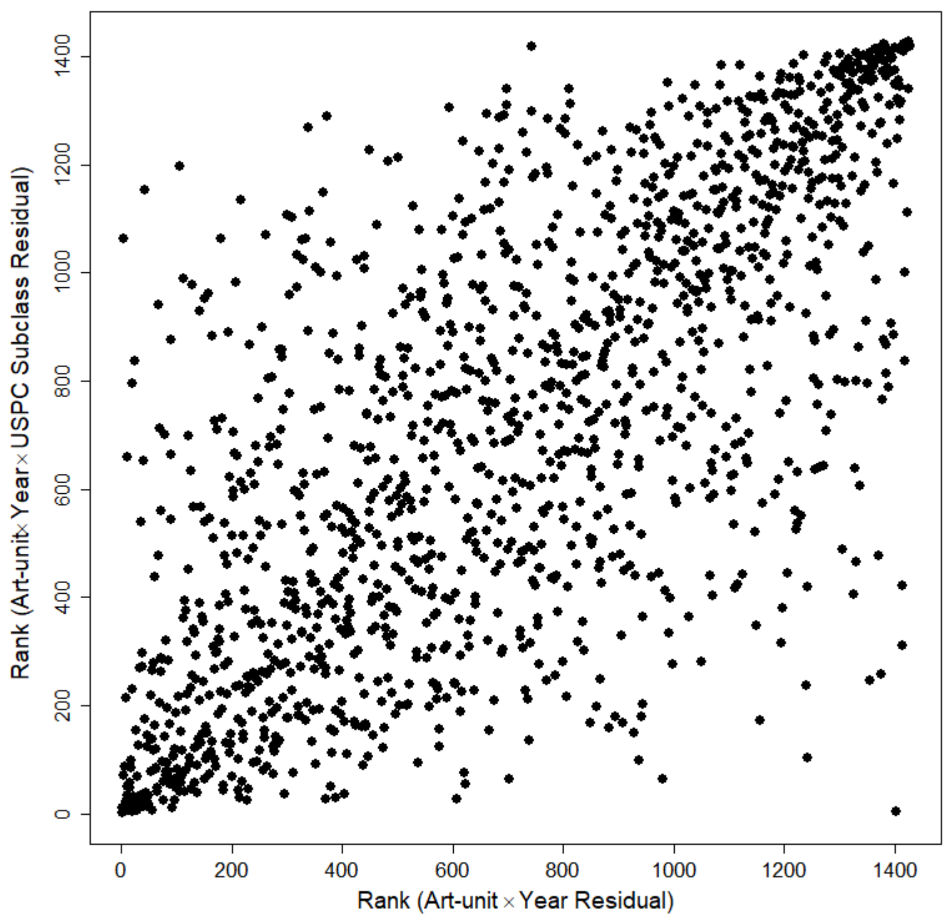


Note: The figure considers data from n = 5,621,374 observations from all 1,427 entities proving legal counsel. We only consider patent applications associated entities associated with at least 1000 applications in the data. The reference model is calculated as the average of 1,000 replications of scrambling examiner leniency across applications. The empirical values come from estimating Equation (2).

Figure A2 shows the scatter plot of ranks for firms under the two different residualization schemes used in Figure 3 and Table 3.

Table A1 shows the descriptive statistics for the variables used in the regression analysis reported in Table 3. It also includes the application-level controls number of claims, scope and small applicant indicator. For these controls, the analysis uses averages over all applications of the entity.

Figure A2: Scatter Plot of the Ranks from the Analyses Displayed in Figure 3



Note: We consider 1,427 counseling entities. The x-axis shows the rank according to average assigned examiner leniency residualized by Art-unit × Year fixed effects. The y-axis shows the rank according to average assigned examiner leniency residualized by Art-unit × Year × USPC subclass fixed effects.

Table A1: Summary Statistics on the Entity-level

	n	Mean	St.Dev.	1. Quartile	Median	3. Quartile
No. Applications	1,427	6813.98	12546.78	1500	2875	6336
HHI	1,427	0.016	0.028	0.006	0.009	0.016
Avrg. No. Claims	1,427	19.78	3.67	18	20	22
Avrg. Scope	1,427	772.41	171.81	686	750	827
Share Small Applicants	1,427	0.228	0.210	0.057	0.178	0.340
IP Stars Ranked (2020)	61	-	-	-	-	-
Law Firm	983	-	-	-	-	-
In-house Counsel	367	-	-	-	-	-
Individual	34	-	-	-	-	-
Government	10	-	-	-	-	-
Other	33	-	-	-	-	-

Note: No. Applications is the total number of applications counseled by the entity. Avrg. No. Claims is the average number of claims listed at the initial time of application. HHI is the concentration of applications by one counseling entity across art units. Avrg. Scope is the average number of characters in the first claim of the application at the initial time of application. IP Stars Ranked (2020) indicates whether the entity is listed in the 2020 version of the IP Stars Ranking. The other indicator variables determine the type of the entity.

B.2 Additional Details for Section 4

This appendix collects methodological detail and supplementary results for the strategic timing analysis in Section 4.2 and the revolving-door analysis in Section 4.3. We first describe the XGBoost hyperparameter tuning procedure used in the rolling prediction exercise. We then report supplementary prediction results for application success. Finally, we present timing-regression robustness using the raw split-window expected-leniency predictors instead of the fitted XGBoost leniency signal. The remaining appendix details the imputation estimator used for the revolving-door event study, the construction and validation of the LinkedIn-based revolving-door measure, the characteristics of the sample firms, and the timing of revolving-door transitions. Two further tables report alternative definitions of the counsel firm and a mediation analysis of the leniency-to-grant pathway.

B.2.1 XGBoost Hyperparameter Tuning

The prediction exercise uses rolling, art-unit-specific XGBoost models. For each art unit and each out-of-sample month, the model is trained on applications filed during the previous rolling three-year window (1,095 days). The resulting model is then applied to applications filed in the next month. All predictors are standardized using only the training sample within the corresponding rolling window. This ensures that both model estimation and predictor normalization use only information available before the out-of-sample prediction month.

Hyperparameters are selected separately within each rolling training window. For this purpose, we split the rolling training sample into an internal training sample and an internal validation sample, using the first 75% of observations for internal estimation and the remaining 25% for validation. For success prediction, the candidate model with the highest validation AUC is selected. For realized-leniency prediction, the candidate model with the lowest validation mean squared error is selected. The selected specification is then re-estimated on the full rolling training window before predicting the next out-of-sample month. If the internal validation split is too small or lacks variation in the dependent variable, we use the baseline XGBoost specification. The hyperparameter grid is reported in Table A2.

Table A2: XGBoost Hyperparameter Grid

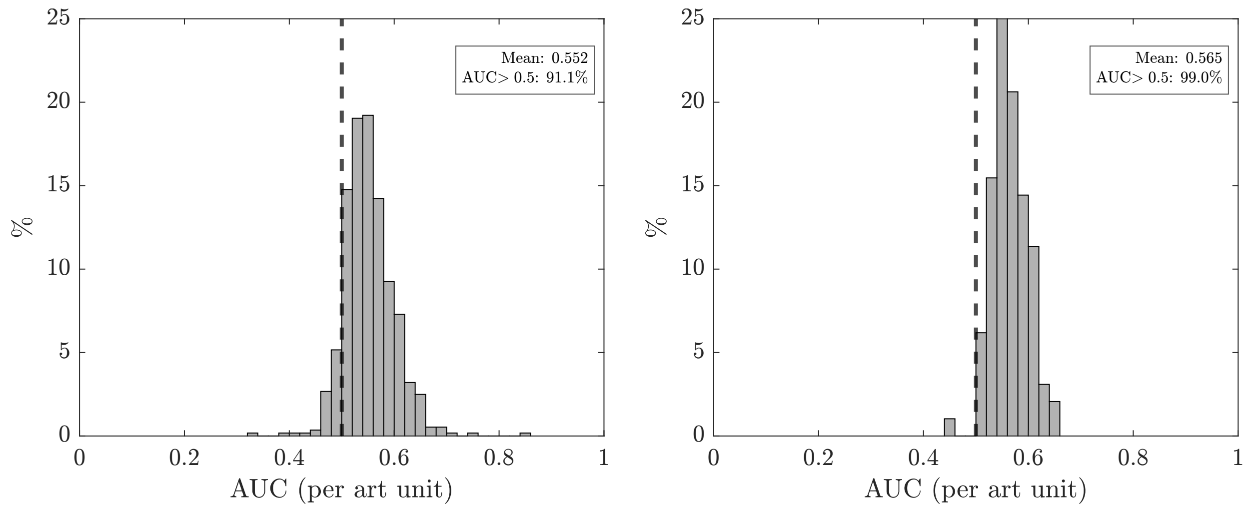
Specification	Trees	Depth	Learning rate	Min. child weight	λ
Depth 1	30	1	0.05	20	1
Baseline	30	2	0.05	20	1
More trees	60	2	0.03	20	1
Strong regularization	30	2	0.05	50	5

Note: Table reports the hyperparameter grid used in the rolling out-of-sample prediction exercise. The same grid is used for realized-leniency and success prediction. For realized leniency, XGBoost is estimated as a squared-error regression model and the validation criterion is mean squared error. For success, XGBoost is estimated as a classifier and the validation criterion is AUC. All specifications use subsample and column-subsample parameters of 0.85 and histogram-based tree construction.

B.2.2 Success Prediction Results

Figure A3: Distribution of Predictive Quality across Art Units

(a) All art units. (b) Art units with $\geq$ 10,000 applications.



Note: This figure shows the distribution of Art Unit-level out-of-sample AUC values from rolling XGBoost predictions of application success. The model uses nine latest-examiner predictors constructed from public information available before filing: mean expected leniency, log number of available examiners, and the share of available examiners in the 1–7, 8–14, and 15–21 day windows. The left panel includes all eligible Art Units; the right panel restricts the sample to the large-Art-Unit timing sample, defined as Art Units with at least 10,000 applications. The dashed vertical line marks an AUC of 0.5, so values to the right indicate predictive performance better than random ranking. The boxes report the mean AUC and the share of Art Units with AUC above 0.5.

B.2.3 Timing Robustness with Raw Predictors

Table A3: Timing Regressions Using Raw Split-Window Expected Leniency

	All AUs		Large AUs	
	Log sub.	Sub. 1/0	Log sub.	Sub. 1/0
	(1)	(2)	(3)	(4)
Exp. leniency 1–7	0.0001	0.0000	0.0003**	0.0004**
	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Exp. leniency 8–14	0.0001	0.0001	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Exp. leniency 15–21	0.0000	0.0000	−0.0000	−0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Entity Rank	0.0010	0.0011	−0.0007	−0.0010
	(0.0011)	(0.0014)	(0.0010)	(0.0013)
Exp. leniency 1–7 × Entity Rank	0.0004**	0.0006**	0.0007**	0.0009**
	(0.0002)	(0.0003)	(0.0003)	(0.0004)
Exp. leniency 8–14 × Entity Rank	0.0001	0.0002	0.0005***	0.0006***
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Exp. leniency 15–21 × Entity Rank	0.0002	0.0002	0.0004**	0.0004**
	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Clustering	Entity + Day	Entity + Day	Entity + Day	Entity + Day
Within R^2	0.00001	0.00001	0.00004	0.00003
Observations	88,145,642	88,145,642	34,087,861	34,087,861
Entity FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Art Unit FE	Yes	Yes	Yes	Yes

Note: Table reports timing regressions using the raw split-window expected-leniency predictors. The dependent variables are log submissions and an indicator for any submission by an entity on a given art-unit day. Columns (1)–(2) use all art units; columns (3)–(4) use large art units. Expected leniency variables are standardized within art unit. Entity ranks are based on backward-looking seven-year windows, adjusted for art-unit and filing-year fixed effects, and lagged by one year. All specifications include entity, day, and art unit fixed effects. Standard errors are two-way clustered by entity and day. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

B.2.4 Imputation Estimator for the Revolving-Door Event Study

The event study uses the imputation estimator of Borusyak et al. (2024); Gardner (2022) develops an equivalent two-stage formulation. The estimator is designed for staggered-adoption settings with heterogeneous treatment effects, where the standard two-way fixed-effects event study mixes already-treated observations into the comparison group. When treatment effects vary across cohorts or over event time, this contamination produces negatively-weighted comparisons that can flip the sign of the pooled estimand even when every cohort-specific effect is positive.

The procedure has two stages. In the first stage, we fit a linear model for examiner leniency on the subset of untreated applications: those filed before the firm's first ex-post USPTO hire, together with applications from firms that never make such a hire. The first-stage specification absorbs firm and art-unit×year fixed effects and conditions on the application-level controls used throughout the paper (small-entity status, claim count, and scope). Restricting the first stage to untreated cells ensures that post-treatment leniency does not contaminate the estimated fixed effects, which would otherwise be the case if treated post-hire observations were used to identify the firm and art-unit-by-year means. Formally, for application i filed by firm f in art unit a and year t, the first stage estimates

$$Y_{ifat} = X_{ifat}'\beta + \alpha_f + \gamma_{at} + \varepsilon_{ifat} \quad \text{on} \quad \{(i, f, a, t) : t < E_f\},$$

where E_f is the calendar year of firm f's first ex-post USPTO hire (set to ∞ for never-treated firms), α_f is a firm fixed effect, and γ_{at} is an art-unit×year fixed effect.

In the second stage, we use the first-stage coefficients to predict counterfactual leniency $\hat{Y}_{ifat}$ for every observation in the panel and form residuals $\hat{\varepsilon}_{ifat} = Y_{ifat} - \hat{Y}_{ifat}$. For treated firms in post-hire periods these residuals identify the treatment effect, up to first-stage estimation error. We then regress the residuals on a saturated set of event-time indicators relative to each firm's first ex-post hire,

$$\hat{\varepsilon}_{ifat} = \sum_{k \neq -1} \theta_k \mathbf{1}\{t - E_f = k\} + u_{ifat},$$

with $t = -1$ omitted as the reference. No additional fixed effects enter the second stage: those have already been absorbed by the first stage, and adding them again would double-residualize the data and understate residual variance (Borusyak et al., 2024).

Inference accounts for the two-stage structure. We apply the analytical variance correction of Gardner (2022), which adjusts for first-stage estimation error in the imputed counterfactuals, and cluster standard errors at the firm level—the unit at which treatment is assigned.

B.2.5 Revolving-Door Data Construction

Table A4 reports the LinkedIn data processing pipeline. We include the full funnel for transparency: identifying revolving-door practitioners from raw LinkedIn profiles requires several layers of filtering, and each stage materially shrinks the sample. We start from 110.2 million US profiles, retain the 78.9 million with parsed employment histories, and extract 207.6 million individual experience entries. After matching organization names against USPTO name variants (Jaro-Winkler ≤ 0.08), filtering through a patent-office whitelist together with a blocklist of commonly confused agencies, restricting to patent-relevant job titles, and tightening firm-name thresholds, we are left with 2,585 unique revolving-door individuals connected to 1,363 canonical firms. Reporting the funnel allows readers to audit which false positives (e.g., postal-service employees, individuals with USPTO-adjacent but non-patent titles) are removed at each stage.

Table A4: LinkedIn Data Processing Funnel

Stage	**Rows/Entries**	**Unique Profiles**
LinkedIn profiles scanned	110,156,678	110,156,678
Profiles with experience data	78,939,953	78,939,953
Experience entries extracted	207,634,691	–
USPTO-matched profiles (JW $\leq$ 0.08)	120,514	120,514
Matched to canonical firms	58,731	31,021
After organization name cleaning	12,175	4,947
After job title filtering	8,614	3,421
After firm match validation	5,564	2,585
After deduplication (person-firm)	4,817	2,585

Note: The table traces the LinkedIn data from raw profiles through successive cleaning stages. The initial corpus covers 107 compressed JSONL files of US LinkedIn profiles. Organization matching uses Jaro-Winkler distance against USPTO name variants. Cleaning removes non-USPTO government agencies (e.g., Postal Service), filters to patent-relevant job titles, and tightens firm name matching thresholds.

Table A5 reports the direction of revolving-door movement in the LinkedIn sample. The revolving door operates in both directions: 43.9% of identified individuals move from private firms into the USPTO, 44.2% move from the USPTO into private firms, and 10.1% appear on both sides of the transition over their careers. Because the institutional-knowledge mechanism we test runs from the USPTO to private firms, the main analysis focuses on the ex-post (USPTO $\rightarrow$ firm) subgroup.

Figure A4 plots the timing of revolving-door transitions in the LinkedIn sample. Departures from the USPTO are concentrated in the 2010s, consistent with a wave of examiner-to-private-sector transitions during this period. The temporal spread of treatment timings is what makes

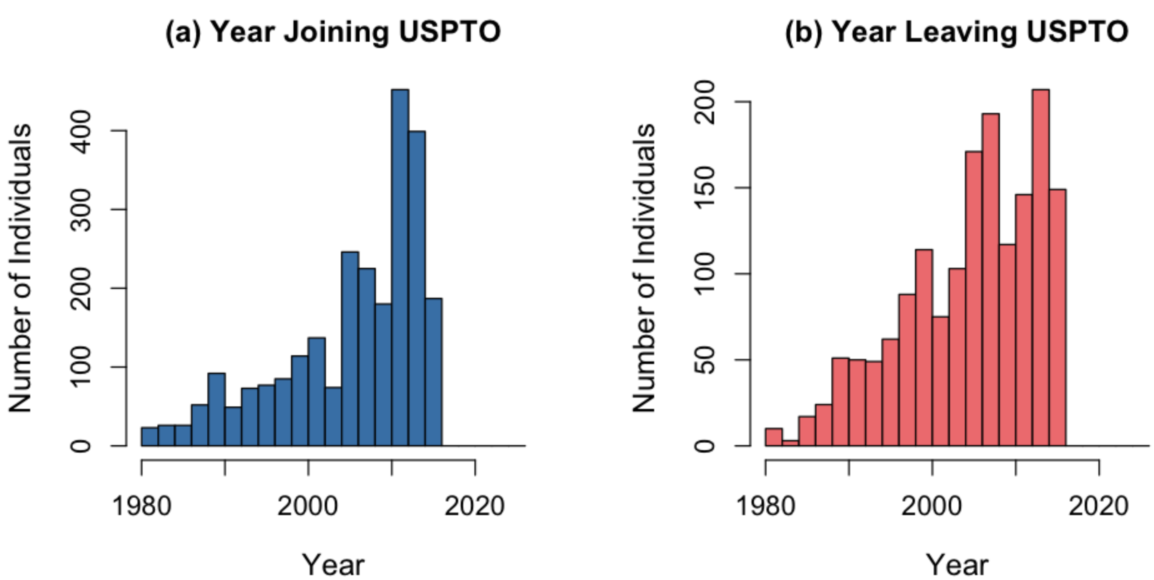
Table A5: Direction of Revolving Door Movement (LinkedIn)

Direction	**N**	**%**
Firm → USPTO only	1,136	43.9
USPTO → Firm only	1,142	44.2
Both directions	262	10.1
Total	2,585	100

Note: Direction is classified by comparing the earliest and latest years of employment at the USPTO versus private firms in the cleaned LinkedIn sample. "Firm → USPTO" indicates the individual worked at a private firm before joining the USPTO. "Both" indicates movement in both directions across their career. The three categories sum to 97.9%; the remaining 45 individuals could not be classified due to incomplete dating.

the staggered-adoption event study feasible: with sufficient mass at different event times, we can identify post-hire dynamics that are not driven by any single cohort or calendar year.

Figure A4: Timing of Revolving Door Transitions (LinkedIn)



Note: Panel (a) shows the distribution of years in which revolving door individuals first appear at the USPTO. Panel (b) shows the distribution of years in which they last appear at the USPTO. Both panels are based on the cleaned LinkedIn sample of 2,585 unique revolving door profiles.

B.2.6 Additional Revolving-Door Results

Table A6 probes sensitivity to how we assign a counsel firm to each application. PatEx records up to three listed counsel firms per application, which leaves an analyst choice about which firm's revolving-door stock to attribute to the patent. Column (1) reproduces the baseline using the first-listed firm only. Column (2) instead uses the maximum stock across all listed firms, which is more inclusive but introduces ambiguity about which firm's behavior the coefficient reflects. Column (3) restricts to single-counsel applications, eliminating the ambiguity at the cost of dropping roughly 70% of the sample. The point estimate is stable at 0.05–0.09 percentage points across the three specifications and statistical significance is preserved throughout, indicating that the headline result in the main text is not an artifact of the first-listed convention.

Table A6: Alternative Revolving Door Measures and Examiner Leniency

		SampleLen	
	(1)	(2)	(3)
RD Stock, First-Listed (log)	0.0881***		0.0805*
	(0.0336)		(0.0483)
RD Stock, Any Firm (log)		0.0502***	
		(0.0187)	
Sample	All, 1st Listed	All, Any Firm	Single Firm Only
Clustering	Entity, Art Unit	Entity, Art Unit	Entity, Art Unit
R^2	0.49772	0.49772	0.50734
Observations	5,639,790	5,639,790	1,574,345
Entity FE	✓	✓	✓
Art Unit × Year FE	✓	✓	✓

Dependent variable is examiner leniency. Column (1) uses the first-listed counsel firm; (2) uses the maximum RD stock across up to three listed counsel firms; (3) restricts to applications with a single counsel firm. All specifications additionally control for small entity status, number of claims, and patent scope (coefficients suppressed for readability). Standard errors two-way clustered by entity and art unit in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A7 reports a mediation analysis of the leniency-to-grant pathway. Column (1) confirms that revolving-door stock predicts examiner leniency. Column (2) regresses the grant indicator directly on revolving-door stock; the coefficient is statistically and economically zero. Column (3) adds examiner leniency to the grant regression as a mediator: the leniency coefficient is large and highly significant (0.764), while the revolving-door coefficient remains zero. Together, these patterns are consistent with revolving-door connections shifting grant outcomes entirely through examiner assignment rather than through any direct channel such as differential application quality conditional on examiner. This pathway is also the assumption used to translate the event-study leniency effect into the dollar magnitude reported in Section **??**.

Table A7: Mediation Analysis: Revolving Door, Leniency, and Grant Probability

	Examiner Leniency (pp)	Granted (pp)	
	(1)	(2)	(3)
RD Stock (log)	0.0881***	-0.0999	-0.1671
	(0.0336)	(0.2146)	(0.2049)
Examiner Leniency (pp)			0.7635***
			(0.0060)
Dependent Variable	Leniency	Granted	Granted
Clustering	Entity, Art Unit	Entity, Art Unit	Entity, Art Unit
R^2	0.49772	0.14775	0.20915
Observations	5,639,790	5,639,790	5,639,790
Entity FE	✓	✓	✓
Art Unit × Year FE	✓	✓	✓

Column (1) confirms that RD stock predicts examiner leniency. Columns (2)–(3) regress the grant indicator on RD stock, with (3) adding examiner leniency as a mediator. The RD stock coefficient is insignificant in both grant regressions, indicating the effect operates through the leniency channel. All specifications additionally control for small entity status, number of claims, and patent scope (coefficients suppressed for readability). Standard errors two-way clustered by entity and art unit in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

B.3 Additional Details for Section 5

Summary statistics for the dependent and independent variables are provided in Table A8. We analyze data on over 5.7 million patent applications. The applications show an average leniency of 71.4% and about one fifth of them are made by applicants qualifying for a discounted fee. The median application has 20 claims and while the average is slightly below the median, the range of this variable is between one claim and over 8,000 claims. The scope is more right-skewed with an average number of 774 characters in the first claim of the patent.

Table A8: Summary Statistics on the Application-level

	n	Mean	St.Dev.	1. Quartile	Median	3. Quartile
Examiner Leniency (in 100%)	5,708,617	71.41	20.91	59.52	76.71	87.69
Discounted Application	5,708,617	0.216	0.412	-	-	-
No. Claims	5,708,617	19.46	26.28	12	20	21
Scope	5,708,617	774.60	831.55	413	658	979

Note: Examiner leniency is defined as the average probability of an application's examiner to grant a patent in the applications' year and art unit. The application is discounted if it comes from an individual, a nonprofit organization or a company with fewer than 500 employees. No. Claims is the number of claims listed at the time of the application. Scope is the number of characters in the first claim of the application at the initial time of application.

Table A9 shows results of an analysis similar to the one displayed in Table 6 but including Art unit $\times$ Year $\times$ USPC subclass fixed effects. Results are consistent in sign and significance with the exception of the inverse measure of application scope which now becomes negative and significant. If larger scope of the application is a proxy of a higher value of π_i then this is in line with our theoretical predictions. Note, however, that such a claim is not unambiguously supported in the literature (e.g., Righi and Simcoe, 2019). The magnitude of the small entity indicator coefficients are reduced. This could be because including more granular fixed effects might control for the strategic behavior by applicants and their legal counsel which is what drives coefficient magnitude in our theoretical model.

Table A9: Application-level Evidence with USPC Subclass Fixed Effects

	Dependent Variable: Examiner Leniency					
	(1)	(2)	(3)	(4)	(5)	(6)
Discounted Appl.	-0.507^{***}				-0.508^{***}	
	(0.037)				(0.037)	
Small Entity		-0.502^{***}				-0.503^{***}
		(0.037)				(0.037)
Micro Entity		-0.617^{***}				-0.602^{***}
		(0.116)				(0.116)
No. Claims			0.006^{***}		0.006^{***}	0.006^{***}
			(0.002)		(0.002)	(0.002)
Application Scope				-0.0001^{***}	-0.0001^{***}	-0.0001^{***}
				(0.00002)	(0.00002)	(0.00002)
Fixed effects	Art unit $\times$ Year $\times$ Subcl.	Art unit $\times$ Year $\times$ Subcl.	Art unit $\times$ Year $\times$ Subcl.	Art unit $\times$ Year $\times$ Subcl.	Art unit $\times$ Year $\times$ Subcl.	Art unit $\times$ Year $\times$ Subcl.
Clustered st. err.	Art unit	Art unit	Art unit	Art unit	Art unit	Art unit
Observations	5,708,617	5,708,617	5,708,617	5,708,617	5,708,617	5,708,617

Note: Table shows the results of estimating Equation (9) either in full (columns (5) and (6)) or with each explanatory variable in isolation (columns (1) through (4)). Examiner leniency is defined as the average probability of an application's examiner to grant a patent in the applications' year and art unit. The application is discounted if it comes from an individual, a nonprofit organization or a company with fewer than 500 employees. Micro entities get a larger discount than small entities. No. Claims is the number of claims listed at the time of the application. Scope is the number of characters in the first claim of the application at the initial time of application. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

B.4 Additional Details for Section 6

This subsection provides additional material for the trading-strategy analysis in Section 6. It includes summary statistics for firm characteristics across residualized examiner-leniency terciles in Table A10, the transition matrix for these tercile assignments in Figure A5, and multi-horizon trading-strategy results in Table A11.

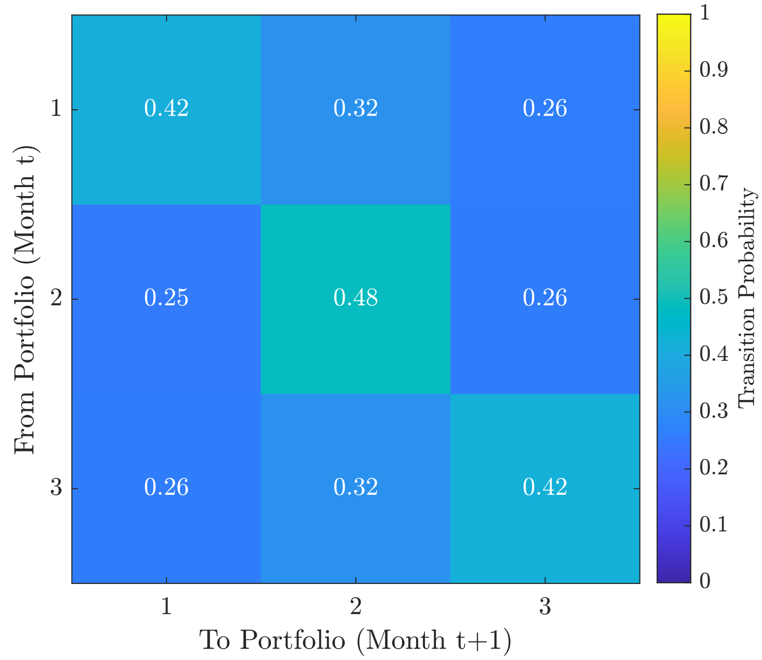
Table A10: Firm Characteristics by Residualized Examiner Leniency Tercile

	T1	T2	T3	T3–T1	p-value (T3–T1)
Beta	1.211	1.151	1.202	-0.010	0.528
Size	16.271	44.522	15.354	-0.916	0.341
Book to Market	0.280	0.275	0.305	0.025	0.000
Momentum (6m)	0.066	0.071	0.067	0.000	0.929
Illiquidity	0.000	0.000	0.000	0.000	0.378
Operating Profitability	0.136	0.280	0.156	0.019	0.812
Gross Profits to Assets	0.257	0.287	0.294	0.037	0.000
Net Payout Yield	-0.009	0.015	0.002	0.011	0.000
Asset Growth	0.164	0.139	0.165	0.001	0.929
Max. Return	0.061	0.054	0.059	-0.002	0.064
ReturnSkew	0.105	0.070	0.106	0.001	0.877
Idiosyn. Volatility	0.020	0.018	0.020	-0.001	0.001
Num. of Analysts	11.488	14.888	11.153	-0.335	0.139
Patents to RD	0.015	0.014	0.014	-0.001	0.264
Cash to Assets	0.302	0.264	0.265	-0.037	0.000
Investment to Revenue	0.980	0.980	0.994	0.014	0.029

Note: This table reports average firm characteristics across terciles formed on residualized examiner leniency. In each month, firms are assigned to terciles based on the firm-month residualized leniency measure used in the portfolio-sort analysis. T1 denotes the lowest-leniency tercile and T3 the highest-leniency tercile. For each characteristic, the table reports the time-series average of the cross-sectional mean within each tercile, the difference between the highest and lowest terciles (T3–T1), and the corresponding p-value based on Newey-West standard errors. Size is measured as lagged market capitalization and is reported in billions of U.S. dollars. Firm characteristics are obtained from the Open Asset Pricing signal library (openassetpricing.com); detailed variable definitions are available on that website. Firms are first sorted into terciles of residualized examiner leniency and, for each characteristic, averages are computed using all nonmissing observations within a tercile-month. The sample is restricted to months in which each tercile contains at least 50 firms.

Table A11 complements the baseline one-month trading results by extending the holding period to multiple months. The purpose of this exercise is to assess whether the return differential persists beyond the month immediately following disclosure. Following the overlapping portfolio methodology of Jegadeesh and Titman (1993), the returns for each horizon are averaged across H active portfolios formed in months $t-1$ through $t-H$. We find that the predictive content of examiner leniency is highly concentrated in the period immediately following disclosure. The

Figure A5: Transition Matrix for Residualized Examiner Leniency Terciles



Note: This figure reports the month-to-month transition matrix for firms sorted into terciles based on the firm-month residualized examiner leniency measure used in the portfolio-sort analysis. In each month, firms are assigned to terciles using the information available at month-end. The entries report the probability that a firm in tercile i in month t is in tercile j in month $t+1$, so each row sums to one. T1 denotes the lowest-leniency tercile and T3 the highest-leniency tercile. The sample is restricted to months in which each tercile contains at least 100 firms.

risk-adjusted returns exhibit a sharp attenuation as the horizon lengthens. For instance, the CAPM alpha of 43.05 basis points observed at $H = 1$ declines to economically smaller and statistically insignificant magnitudes for all horizons of $H \geq 2$. A similar decay is observed under the FF3 and FF3+MOM specifications. While mean excess returns remain positive at longer horizons, the rapid decay in risk-adjusted alphas implies that the abnormal performance associated with the identity signal is largely realized within the first month after the information becomes public. Such a pattern is consistent with a slow information diffusion process where market participants gradually process the complex implications of examiner assignment. As the "identity signal" is gradually incorporated into valuations, the initial mispricing is corrected and abnormal returns dissipate.

Table A11: Value-weighted CAPM alphas of multi-horizon trading strategy

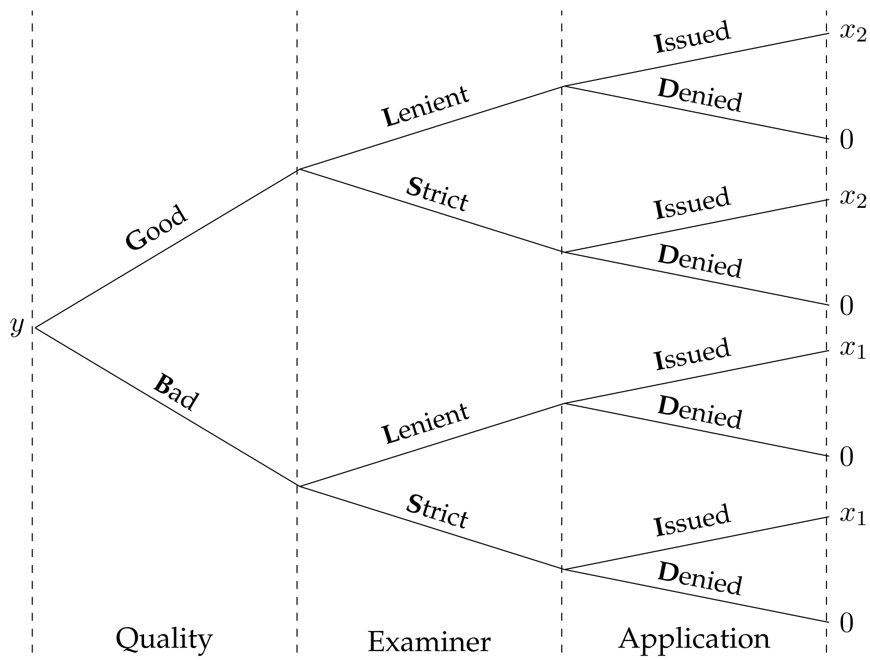
	$H = 1$	$H = 2$	$H = 3$	$H = 4$
Mean Returns (bps)	43.05***	18.85	15.65	13.26
	(14.91)	(12.50)	(12.36)	(11.99)
CAPM Alpha (bps)	35.10**	10.89	7.51	5.53
	(15.15)	(12.74)	(12.55)	(12.34)
FF3 Alpha (bps)	35.28**	11.48	8.22	6.16
	(15.20)	(13.16)	(13.07)	(12.81)
FF3 + MOM Alpha (bps)	35.80**	10.86	7.50	4.87
	(15.27)	(13.45)	(13.08)	(12.91)

Note: This table reports value-weighted average monthly excess returns and risk-adjusted alphas (in bps) for the long-short strategy (T3–T1) across holding periods H. At the end of each month $t-1$, firms are assigned to terciles based on the mean residual leniency of their assigned patent examiners. Residual leniency is defined as the examiner's grant rate demeaned by Art Unit over the previous calendar year. For $H > 1$, monthly returns follow the overlapping portfolio methodology of Jegadeesh and Titman (1993), averaging the returns of H vintages formed in months $t-1, \ldots, t-H$. Stocks are value-weighted by market capitalization at the time of formation. Risk adjustments include the CAPM, Fama-French three-factor (FF3), and FF3 plus momentum (FF3+MOM) models. Newey and West (1986) standard errors with a bandwidth of H are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix C Coexistence of result with Shu et al. (2022)

In this section, we make a short formal argument why the result of Shu et al. (2022) can coexist with ours without any contradiction. For this demonstration, we assume that all random variables are binary. A patent can be of good (G) or bad (B) quality. The examiner can be lenient (L) or strict (S). And the application can be issued (I) or denied (D). The entire situation, in order of the timing of the patent application process, is shown in Figure A6.

Figure A6: Event tree of a patent application process



Note: Stock outcome dependent on patent quality (first branch), examiner leniency (second branch) and issuance (third branch). All three variables are assumed to be binary.

We are interested in the return to the patent, denoted y. The return of a granted patent of good quality is x_2, that of a granted patent of bad quality is x_1. If a patent application is denied, there is no return. We assume, in line with virtually all research on the matter, that $x_2 > x_1 > 0$. We further make two assumptions on the patent application process, the first is that, conditional on the quality Q of the patent, the probability of it being issued is higher for lenient examiners than for strict ones ($P(I|L \cap Q) > P(I|S \cap Q)$ for $Q = G, B$). This is an innocuous assumption, as it is consistent with the definition of examiner leniency. The second assumption is that, conditional on the examiner E, good quality patents are more likely to be issued than bad quality ones ($P(I|G \cap$

$E) > P(I|B \cap E)$ for $E = L, S$). This assumption is already implied in the study by Shu et al. (2022) and also relatively uncontroversial.

Based on the model in Figure A6 and the assumptions, we can now formalize the two hypotheses that Shu et al. (2022) and we test. We begin with our assumption. We state that an assignment of a patent application to a lenient examiner increases the return. This can be formally written as

$$\mathbb{E}[y|L] > \mathbb{E}[y|S]. \tag{A18}$$

Shu et al. (2022), on the other hand, state that the return to an issued patent will be higher when assigned to a strict examiner than when assigned to a lenient examiner. They thus hypothesize that

$$\mathbb{E}[y|L \cap I] < \mathbb{E}[y|S \cap I]. \tag{A19}$$

Neither of the two hypotheses or empirical results requires that firms manipulate the examiner assignment process. We thus assume here that $P(S|G) = P(S|B)$ and $P(L|G) = P(L|B)$. Coexistence of the two results can also be shown under a different dependence structure that takes the results of Section 5.2 into account, but we opt for the simple setting because it makes the exposition clearer and is sufficient for our purposes. We can write (A18) as

$$x_2 P(G \cap I|L) + x_1 P(B \cap I|L) > x_2 P(G \cap I|S) + x_1 P(B \cap I|S). \tag{A20}$$

By the definition of conditional probability,

$$x_2 P(G|L)P(I|G \cap L) + x_1 P(B|L)P(I|B \cap L) > x_2 P(G|S)P(I|G \cap S) + x_1 P(B|S)P(I|B \cap S). \tag{A21}$$

By virtue of the no manipulation assumption, we have $P(E \cap G) = P(E)P(G)$ and thus $P(G|L) = P(G|S) = P(G)$ and $P(B|L) = P(B|S) = P(B)$. (A18) is thus equivalent to

$$x_2 P(G)[P(I|G \cap L) - P(I|G \cap S)] + x_1 P(B)[P(I|B \cap L) - P(I|B \cap S)] > 0. \tag{A22}$$

The inequality follows because lenient examiners are more likely to issue patents than strict examiners. This shows that our hypothesis will always be true, given the assumptions above.

What remains to be shown is that (A19) can also be true under the given assumption. We can write it as

$$x_2 P(G|I \cap L) + x_1 P(B|I \cap L) < x_2 P(G|I \cap S) + x_1 P(B|I \cap S) \tag{A23}$$

$$x_1 + (x_2 - x_1)P(G|I \cap L) < x_1 + (x_2 - x_1)P(G|I \cap S) \tag{A24}$$

$$P(G|I \cap L) < P(G|I \cap S). \tag{A25}$$

From Bayes' theorem $P(G|I \cap E) = \frac{P(I|G \cap E)P(G|E)}{P(I|E)}$ and thus by the no manipulation condition

$$\frac{P(I|G \cap L)}{P(I|G \cap S)} < \frac{P(I|L)}{P(I|S)}. \tag{A26}$$

This implies that under our assumptions the hypothesis of Shu et al. (2022) can be true as long as the relative increase in grant probability from strict to lenient examiners is larger for bad quality applications than for good quality applications. This seems in line with the rubber stamping hypothesis the authors formulate in their study.

Appendix D Details on Examiner Leniency IVs

D.1 Derivation of the IV Estimand

This appendix derives the population 2SLS estimand under the data-generating process described in the main text. Recall

$$Z = \gamma_Z Q + \varepsilon_Z, \tag{A27}$$

$$X = \gamma_X Q + \beta_Z Z + \varepsilon_X, \tag{A28}$$

$$Y = \gamma_Y Q + \beta_X X + \varepsilon_Y, \tag{A29}$$

where $Q, \varepsilon_Z, \varepsilon_X$, and ε_Y are mutually independent, mean zero random variables with finite variances.

The population IV estimand for the just-identified model is

$$\beta^{IV} = \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, X)}. \tag{A30}$$

We first derive the denominator. Using the expression for X,

$$\text{Cov}(Z, X) = \text{Cov}\big(Z, \gamma_X Q + \beta_Z Z + \varepsilon_X\big) \tag{A31}$$

$$= \gamma_X \text{Cov}(Z, Q) + \beta_Z \text{Var}(Z), \tag{A32}$$

where $\text{Cov}(Z, \varepsilon_X) = 0$ by independence. Since $Z = \gamma_Z Q + \varepsilon_Z$,

$$\text{Cov}(Z, Q) = \gamma_Z \text{Var}(Q), \tag{A33}$$

$$\text{Var}(Z) = \gamma_Z^2 \text{Var}(Q) + \text{Var}(\varepsilon_Z). \tag{A34}$$

Hence,

$$\text{Cov}(Z, X) = \gamma_X \gamma_Z \text{Var}(Q) + \beta_Z \big(\gamma_Z^2 \text{Var}(Q) + \text{Var}(\varepsilon_Z)\big). \tag{A35}$$

Next consider the numerator. Using the expression for Y,

$$\text{Cov}(Z, Y) = \text{Cov}\big(Z, \gamma_Y Q + \beta_X X + \varepsilon_Y\big) \tag{A36}$$

$$= \gamma_Y \text{Cov}(Z, Q) + \beta_X \text{Cov}(Z, X), \tag{A37}$$

since $\text{Cov}(Z, \varepsilon_Y) = 0$. Substituting $\text{Cov}(Z, Q) = \gamma_Z \text{Var}(Q)$ yields

$$\text{Cov}(Z, Y) = \gamma_Y \gamma_Z \text{Var}(Q) + \beta_X \text{Cov}(Z, X). \tag{A38}$$

Combining numerator and denominator,

$$\beta^{IV} = \frac{\gamma_Y \gamma_Z \mathrm{Var}(Q) + \beta_X \mathrm{Cov}(Z, X)}{\mathrm{Cov}(Z, X)} = \beta_X + \frac{\gamma_Y \gamma_Z \mathrm{Var}(Q)}{\mathrm{Cov}(Z, X)}. \tag{A39}$$

Substituting for $\mathrm{Cov}(Z, X)$ gives

$$\beta^{IV} = \beta_X + \frac{\gamma_Y \gamma_Z \mathrm{Var}(Q)}{\gamma_X \gamma_Z \mathrm{Var}(Q) + \beta_Z \big(\gamma_Z^2 \mathrm{Var}(Q) + \mathrm{Var}(\varepsilon_Z)\big)}. \tag{A40}$$

This expression characterizes the asymptotic bias of the IV estimator arising from the violation of the exclusion restriction through the common factor Q.

D.2 Reassessment of Sampat and Williams (2019)

In their paper, Sampat and Williams (2019, henceforth SW) consider the influence of patents on follow-up innovation. They use the patent examiner leniency as an instrument for patent grant. In light of our findings, this seems questionable. In Section IV.D of their paper, SW consider the possibility that a latent application-quality dimension affects both patentability and examiner assignment. In the notation of Figure 7, this latent variable is denoted by Q.

They provide a test of the quality dimension and argue that based on this test, its importance can be discarded – a conclusion at odds with our results. Below, we reconsider their empirical set-up, their test of the quality dimension, and their empirical identification strategy. We argue that given the values provided in the paper, one can construct data in which there is no influence of a patent grant on follow-up innovation but the identification strategy nevertheless results in a statistically significant result. Their empirical exercise is informative, but it does not by itself rule out omitted-variable bias in the IV design.

We start by summarizing the key values provided by SW in the notation of Figure 7. In a regression $X = \tilde{\beta}_z Z + \eta$, they obtain $\tilde{\beta}_Z = 0.876$ with a standard error of 0.037 (p. 222). In a regression of X on the quality dimension Q, they obtain a statistically significant influence, not specified further than the p-value of the F statistic (p. 225). However, from this regression, they take the predicted values of X, denote them $\hat{X}^Q$, and regressing $\hat{X}^Q$ on Z, they obtain a coefficient (that we denote $\tilde{\pi}_Q$) of 0.013 with a standard error of 0.003 (p. 225). Lastly, in a 2SLS regression of Y on X, instrumented with Z, they obtain $\tilde{\beta}_X^{2SLS} = -0.023$ with a standard error of 0.010 (Table 3, Panel A, Column (1), p. 226). Together, they take these numbers as evidence for a statistically significant, negative and causal influence of a granted patent (X) on follow-up innovation (Y). Their argument relies on the idea that the estimated relationship between examiner leniency and the quality-predicted grant probability is sufficiently small that omitted quality is unlikely to materially bias the IV estimate.

Restatement. We restate the problem in the notation of Figure 7. Consider the linear system (which is very similar to the DGP in (13))

$$Y = \beta_X X + \gamma_Y Q + \varepsilon_Y, \tag{A41}$$

$$X = \beta_Z Z + \gamma_X Q + \varepsilon_X + \mu, \tag{A42}$$

$$Z = \gamma_Z Q + \varepsilon_Z, \tag{A43}$$

where $Q, \varepsilon_Y, \varepsilon_X, \varepsilon_Z$ are mutually independent, mean-zero random variables with finite variances, and μ is a constant used to match the sample mean of X.

For the defense of their empirical identification strategy, SW form the predicted probability of patent grant using quality proxies. In our notation, the population counterpart of this fitted value can be written as

$$\hat{X}^Q = \alpha_Q + \delta Q. \tag{A44}$$

From substituting (A43) into (A42) we can see that

$$X = \beta_Z(\gamma_Z Q + \varepsilon_Z) + \gamma_X Q + \varepsilon_X + \mu = (\gamma_X + \beta_Z \gamma_Z)Q + \beta_Z \varepsilon_Z + \varepsilon_X + \mu. \tag{A45}$$

Hence the population projection of X onto Q has slope $\delta = \gamma_X + \beta_Z \gamma_Z$. SW then proceed to regress $\hat{X}^Q$ on Z. The corresponding population coefficient is

$$\tilde{\pi}_Q = \frac{\text{Cov}(\hat{X}^Q, Z)}{\text{Var}(Z)}. \tag{A46}$$

Using $\hat{X}^Q = \alpha_Q + \delta Q$ and $Z = \gamma_Z Q + \varepsilon_Z$,

$$\text{Cov}(\hat{X}^Q, Z) = \delta \gamma_Z \text{Var}(Q). \tag{A47}$$

Substituting this and $\delta = \gamma_X + \beta_Z \gamma_Z$ renders

$$\tilde{\pi}_Q = \frac{(\gamma_X + \beta_Z \gamma_Z)\gamma_Z \text{Var}(Q)}{\text{Var}(Z)} = 0.013. \tag{A48}$$

The value of $\tilde{\pi}_Q$ thus provides us with a relationship for γ_X, β_Z, γ_Z, Var(Q), and Var(Z). However, it does not by itself imply that the resulting IV bias is small. We can see this by considering $\tilde{\beta}_X^{2SLS}$ directly. From the main text and Appendix D.1, we know that the just-identified IV estimand $\tilde{\beta}_X^{2SLS} = \frac{\text{Cov}(Z,Y)}{\text{Cov}(Z,X)}$ can be written as

$$\tilde{\beta}_X^{2SLS} = \beta_X + \frac{\gamma_Y \gamma_Z \text{Var}(Q)}{\beta_Z \text{Var}(Z) + \gamma_X \gamma_Z \text{Var}(Q)}. \tag{A49}$$

This makes the key identification point transparent. Even if the value of $\tilde{\pi}_Q$ is small, the IV estimand can still differ materially from β_X whenever $\gamma_Y \neq 0$. The SW auxiliary regression therefore does not bound the magnitude of omitted-variable bias unless one also imposes restrictions on γ_Y. Suppose that the true causal effect of patent grant on follow-on innovation is zero:

$$\beta_X = 0. \tag{A50}$$

Then (A49) reduces to

$$\tilde{\beta}_X^{2SLS} = \frac{\gamma_Y \gamma_Z \mathrm{Var}(Q)}{\beta_Z \mathrm{Var}(Z) + \gamma_X \gamma_Z \mathrm{Var}(Q)}. \tag{A51}$$

For fixed values of β_Z, γ_Z, and $\mathrm{Var}(Q)$, Equation (A51) shows that an appropriate value of γ_Y can reproduce the reported IV coefficient even when $\beta_X = 0$.

Calibration. To see that the reported effect can appear even if $\beta_X = 0$, we now show a calibration exercise. The first stage coefficient $\tilde{\beta}_Z = 0.876$ results from $\tilde{\beta}_Z = \frac{\mathrm{Cov}(X,Z)}{\mathrm{Var}(Z)}$. Using the DGP, we can see that

$$\begin{aligned} \mathrm{Cov}(X,Z) &= \mathrm{Cov}(\beta_Z Z + \gamma_X Q + \varepsilon_X + \mu, Z) && \text{(A52)} \\ &= \beta_Z \mathrm{Var}(Z) + \gamma_X \mathrm{Cov}(Q,Z) && \text{(A53)} \\ &= \beta_Z \mathrm{Var}(Z) + \gamma_X \gamma_Z \mathrm{Var}(Q) && \text{(A54)} \end{aligned}$$

and thus

$$\tilde{\beta}_Z = \beta_Z + \frac{\gamma_X \gamma_Z \mathrm{Var}(Q)}{\mathrm{Var}(Z)}. \tag{A55}$$

From Table 1 (p. 212), we obtain $E[x] = 0.3043$ which implies $\mathrm{Var}(X) = 0.2117$. From own analysis of the USPTO data, we know that Art-unit-by-year demeaned examiner leniency has $\mathrm{Var}(Z) = 0.0225$. Since no coefficient for the quality dimension is reported[25], we assume $\gamma_X = 0.2$. We further assume that Z and Q explain exactly $\varphi_X = 50\%$ of the variance of X. $\mathrm{Var}(Q)$ can be defined from

$$\varphi_X \mathrm{Var}(X) = \beta_Z^2 \mathrm{Var}(Z) + \gamma_X^2 \mathrm{Var}(Q) + 2\beta_Z \gamma_Z Cov(Z,Q). \tag{A56}$$

By the DGP, $\mathrm{Cov}(Z,Q) = \gamma_Z \mathrm{Var}(Q)$. Substituting and rearranging renders

$$\mathrm{Var}(Q) = \frac{\varphi_X \mathrm{Var}(X) - \beta_Z^2 \mathrm{Var}(Z)}{\gamma_X^2 + 2\beta_Z \gamma_X \gamma_Z}. \tag{A57}$$

[25] And, in fact, their quality scale has two variables, so reporting a single coefficient is not possible.

Under the assumption that $\beta_X = 0$, we thus have the three equations

$$-0.023 = \frac{\gamma_Y \gamma_Z \text{Var}(Q)}{\beta_Z \text{Var}(Z) + \gamma_X \gamma_Z \text{Var}(Q)}, \tag{A58}$$

$$0.013 = \frac{(\gamma_X + \beta_Z \gamma_Z)\gamma_Z \text{Var}(Q)}{\text{Var}(Z)}, \tag{A59}$$

$$0.876 = \beta_Z + \frac{\gamma_X \gamma_Z \text{Var}(Q)}{\text{Var}(Z)}. \tag{A60}$$

Together with Equation (A57), these are four equations with four unknowns: Var(Q), γ_Y, γ_Z, and β_Z. They imply the values provided in Table A12 below. We can see that $\beta_X = 0$ could be rationalized by a relatively large effect of quality on the follow-up dimension. Interestingly, the marginal effect of the quality dimension on examiner leniency would not have to be large. A one standard deviation shift in Q only implies a 0.1 percentage point shift in leniency. This is about on par with our results from Table 6 where a one standard deviation shift in the number of claims implies an 0.079 percentage point shift in leniency.

Table A12: Calibrated parameters for the SW IV design

	Calibrated Value	Marg. Effect 1 s.d. increase in Z	Marg. Effect 1 s.d. increase in Q
Var(Q)	2.215		
β_Z	0.862	0.129	
γ_Y	-0.299		0.445
γ_Z	0.0007		0.001

Note: Reported the values resulting from numerically solving Equations (A57) through (A60). From SW, we take the values $\tilde{\beta}_Z = 0.876$, $\tilde{\pi}_Q = 0.013$, $\tilde{\beta}_X^{2SLS} = -0.023$ and Var$(X) = 0.2117$. From own data analysis, we take Var$(Z) = 0.0225$. We assume $\gamma_X = 0.2$ and $\varphi_X = 0.5$.

Interpretation. The argument provided by SW against the importance of the quality dimension is not only quantitative but also graphical. Quantitatively, they argue that the small coefficient size for $\tilde{\pi}_Q$ implies a low importance of the quality dimension in their data. We show above that this argument is misleading. The second argument put forward by SW is a graphical one. In their Figure 3, they argue that the statistical relationship of Z and $\hat{X}^Q$ is difficult to determine when compared to the statistical relationship of Z and X. However, the graphical comparison in SW is not very informative for identification. By construction, the relationship between Z and X will be much stronger than the relationship between Z and the component of X predicted by observed quality proxies. But this says little about the magnitude of bias in $\tilde{\beta}_X^{2SLS}$, which also depends on the effect of latent quality on the outcome, γ_Y. Thus, even if the graph suggests only a weak relationship between Z and $\hat{X}^Q$, omitted-variable bias in the IV estimate can still be substantial.

We do not claim that SW's substantive conclusion is necessarily false. Our point is narrower. The moments they report are jointly consistent with a data-generating process in which patent grant has no causal effect on follow-up innovation. Therefore, their Section IV.D exercise does not suffice to rule out omitted-quality bias in the examiner-leniency IV design. Even a small effect of latent quality on examiner leniency can materially bias causal inference if latent quality also affects the outcome.

Appendix E Attorney Name Cleaning and Law Firm Disambiguation

The data prep links each application to up to three legal entities of record in two stages.

Stage 1: Practitioners to firms via the USPTO roster. From the PatEx application-level practitioner records we extract the registered patent attorneys and agents associated with each application, identified by USPTO registration number. We join these to the historical USPTO practitioner rosters—monthly snapshots maintained by the USPTO that list each registered practitioner's firm of record, mailing address, and registration status. The monthly snapshots are aggregated into a single longitudinal roster. Because practitioner affiliations change over time, we assign each practitioner to the firm observed in the roster snapshot closest to but not after the application's filing date; if no earlier snapshot exists, we fall back to the closest subsequent one. We then aggregate from the practitioner level to the application level by counting how many linked practitioners map to the same firm and retaining the top three firms per application, with ties broken by source order. This three-firm cap is intended to capture the entities most relevant to prosecution, since fourth- and lower-ranked firms typically make only a marginal contribution to the prosecution effort.

Stage 2: Firm-name cleaning. The raw firm-name strings produced by Stage 1 contain substantial variation—a single law firm may appear under dozens of representations due to punctuation differences, abbreviations, typos, OCR errors, and individual-attorney listings. We extract and deduplicate the firm-name strings across all applications (yielding roughly 19,000 unique raw names) and then proceed in seven ordered steps.

1. **Normalization.** Each name is lowercased, stripped of diacritics, has common abbreviations expanded (e.g., "P.C." $\rightarrow$ "professional corporation"), and its punctuation regularized. Several canonical keys (sorted-token order, hash, phonetic) are precomputed for use as blocking keys in later stages.

2. **Initial fuzzy matching.** Names sharing an identical canonical key are linked with full confidence. The remaining pairs are scored with inverse-document-frequency–weighted Levenshtein and token-overlap similarity inside length-bucketed candidate blocks. Candidate edges are recorded in two tiers: high-confidence "seed" edges at similarity ≥ 0.80, and weaker "attach" edges in $[0.58, 0.70)$ that additionally require a shared rare anchor token (a token whose corpus-wide inverse-document-frequency exceeds a fixed threshold). A manual exclusion filter discards candidate matches whose only shared tokens are generic legal terms ("law," "group," "LLP," and the like).

3. **Language-model–assisted cleanup.** GPT-5-mini (temperature 0, with content-addressed caching for reproducibility) is used in three roles: (i) correcting typos and OCR errors in raw names, (ii) semantically clustering near-duplicates above a text-similarity threshold of 0.85, and (iii) disambiguating ambiguous cases using shared mailing addresses and filing-year overlap as side information. A second fuzzy pass is then run on the typo-corrected names to recover matches the first pass missed.

4. **Entity-type classification.** Each unique name is labeled as a law firm, in-house counsel, individual, government entity, or unknown. Deterministic rules apply first: suffix-based (LLP/PC/PA $\to$ law firm; Inc/Corp/LLC absent legal terms $\to$ in-house counsel), keyword-based ("Department of $\cdot$," military branches, "University of $\cdot$" $\to$ government entity), personal-name lookup against a list of common given names, and manual overrides for known entities (e.g., IBM, Fish & Richardson, Morrison & Foerster). Names not resolved with deterministic-rule confidence above 0.92 are sent to GPT-5-mini with a structured response schema returning a classification, a confidence score, and a small set of anchor tokens. A post-processing layer then catches known language-model failure modes—for example, "Legal Department" overrides any law-firm label to in-house counsel, and a university-name pattern forces a government label.

5. **Manual overrides.** A curated list of force-merge and force-split pairs corrects specific entities that prior stages misclassified, addressing both known false positives (firms that should not have been merged) and known false negatives (firms that should have been but were not).

6. **Clustering.** A union-find pass consolidates edges in priority order: typo-correction and force-merge edges first, then seed edges, then attach edges. Attach edges may only join a component that already contains a seed edge, which prevents weak similarity scores from initiating new clusters. Force-split pairs are blocked from co-clustering. An address gate provides a final safeguard: it vetoes merges between names that share a mailing address but disagree substantially on their non-address core tokens, on the logic that different firms at a shared office should not be collapsed. The shortest name in each resulting cluster is selected as the canonical display name.

7. **Audit and apply.** A second language-model pass re-verifies every merged cluster, flagging any cluster whose members the model judges inconsistent. The final canonical mapping is then joined back to the application-level data on each of the three firm-name fields. Raw names never observed in the cleaning corpus receive deterministic singleton identifiers, leaving them trackable but unmerged.

Each application thus exits the pipeline with up to three canonical legal entities and their assigned types, which we use throughout the analysis to identify the law firms representing each application.