Screening Property Rights for Innovation

Online Appendices

William Matcham and Mark Schankerman

A Additional Tables and Figures

Table A.1. Summary Statistics

Variable	Observations	Mean	Median	Std. Dev.
Issued	4,846,053	0.70	1.00	0.46
Duration of Prosecution (years)	4,846,053	2.96	2.67	1.57
Number of Rounds	4,608,833	2.40	2.00	1.45
Independent Claims	3,838,553	2.99	3.00	2.94
Small Entity	4,781,012	0.24	0.00	0.43
Not Renewed at 4	410,667	0.13	0.00	0.33
Renewed at 4, not at 8	410,667	0.19	0.00	0.39
Renewed at 8, not at 12	410,667	0.23	0.00	0.42
Renewed at 12	410,667	0.46	0.00	0.50

Notes: Sample sizes are lower for rounds, claims, and examiner variables since the datasets containing these variables cover a subset of the years 2001-2017. On renewal variables, we restrict attention to patents granted before 2006 to ensure that we have full renewal data on all granted patents. Categorical variables may not sum to one due to rounding.

Table A.2. Estimated and Assigned Parameters

Estimated Parameters			
Variable	**Notation**	**Distribution**	**Parameters**
Examiner			
Intrinsic motivation	$\theta \sim G_{S,\theta}(\cdot)$	Log-normal	σ_θ, $\mu_{\theta,\text{junior}}$ or $\mu_{\theta,\text{senior}}$
Examiner Delay Cost	$\pi \sim G_\pi(\cdot)$	Log-normal	μ_π, σ_π
Error	$\varepsilon \sim G_{e,\varepsilon}(\cdot)$	Normal	σ_ε
Applicant			
Initial claim returns	$v_j^* \sim G_v(\cdot)$	Log-normal	μ_v, σ_v
Initial claim distances	$D_j^* \sim G_D(\cdot)$	Beta	α_D, β_D
Obsolescence	$\boldsymbol{\omega}$	Bernoulli	$P_{\omega,\text{pre}}$ or $P_{\omega,\text{post}}$
Application legal costs	f_{app}	Log-normal	$\mu_{f,\text{app}}, \sigma_{f,\text{app}}$
Issuance legal costs	f_{iss}	Log-normal	$\mu_{f,\text{iss}}, \sigma_{f,\text{iss}}$
Maintenance legal costs	f_{main}	Log-normal	$\mu_{f,\text{main}}, \sigma_{f,\text{main}}$
Amendment legal costs	f_{amend}	Log-normal	$\mu_{f,\text{amend}}, \sigma_{f,\text{amend}}$
Narrowing	η	-	-

Assigned Parameters		
Variable	**Notation**	**Values**
Discount rate	β	0.95
Depreciation	δ	$\frac{0.14-P_{\omega,\text{post}}}{1-P_{\omega,\text{post}}}$
Threshold by technology center	τ	Range from 0.48 to 0.52
Credits	$g^r(S,T)$	-
Finalizing fee	ϕ	\$2,268
RCE fees	$F^3_{\text{round}} = F^5_{\text{round}}$	\$1,034
	F_4	\$1,685
Renewal fees	F_8	\$3,791
	F_{12}	\$7,792

Table A.3. Application Fighting Costs by Technology Area

Parameter	**Symbol**	**Estimate**	**S.E.**
Chemical application fighting cost log-mean	$\mu_{f,\text{chem}}$	9.15	0.008
Chemical application fighting cost log-sigma	$\sigma_{f,\text{chem}}$	0.38	0.010
Electrical application fighting cost log-mean	$\mu_{f,\text{elec}}$	9.18	0.010
Electrical application fighting cost log-sigma	$\sigma_{f,\text{elec}}$	0.57	0.014
Mechanical application fighting cost log-mean	$\mu_{f,\text{mech}}$	9.02	0.008
Mechanical application fighting cost log-sigma	$\sigma_{f,\text{mech}}$	0.47	0.011

Notes: Standard errors are bootstrapped.

Table A.4. Applicant Fighting Costs by Technology Area

Parameter	**Symbol**	**Estimate**
Simple amendment fighting cost log-mean	$\mu_{f,\text{amend},\text{simp}}$	7.60
Simple amendment fighting cost log-sigma	$\sigma_{f,\text{amend},\text{simp}}$	0.37
Chemical amendment fighting cost log-mean	$\mu_{f,\text{amend},\text{chem}}$	8.13
Chemical amendment fighting cost log-sigma	$\sigma_{f,\text{amend},\text{chem}}$	0.45
Electrical amendment fighting cost log-mean	$\mu_{f,\text{amend},\text{elec}}$	8.07
Electrical amendment fighting cost log-sigma	$\sigma_{f,\text{amend},\text{elec}}$	0.38
Mechanical amendment fighting cost log-mean	$\mu_{f,\text{amend},\text{mech}}$	7.95
Mechanical amendment fighting cost log-sigma	$\sigma_{f,\text{amend},\text{mech}}$	0.43
Issuance cost log-mean	$\mu_{f,\text{iss}}$	6.54
Issuance cost log-sigma	$\sigma_{f,\text{iss}}$	0.62
Maintenance cost log-mean	$\mu_{f,\text{main}}$	5.67
Maintenance cost log-sigma	$\sigma_{f,\text{main}}$	0.46

Table A.5. Robustness of Estimates

Parameter	**Symbol**	**Baseline**	1% $\boldsymbol{\tau}$	5% $\boldsymbol{\tau}$	$\boldsymbol{\beta = 0.99}$	**Definition of Seniority (GS13 + GS14)**
Junior intrinsic motivation log-mean	$\mu_{\theta,\mathrm{j}}$	3.92	3.96	3.96	3.90	4.16
Senior intrinsic motivation log-mean	$\mu_{\theta,\mathrm{s}}$	3.38	2.90	2.73	3.18	2.93
Intrinsic motivation log-sigma	σ_θ	0.77	0.82	0.79	0.90	0.99
Examiner delay cost log-mean	μ_π	0.19	0.16	0.18	0.49	0.12
Examiner delay cost log-sigma	σ_π	0.27	0.37	0.42	0.10	0.60
Error standard deviation	σ_ε	0.02	0.02	0.02	0.03	0.02
Initial returns log-mean	μ_v	10.55	10.59	10.88	10.07	10.28
Initial returns log-sigma	σ_v	1.32	1.13	1.61	2.94	0.57
Initial distance alpha	α_D	4.57	3.92	3.90	4.56	3.75
Initial distance beta	β_D	7.74	6.72	6.22	7.79	7.15
Narrowing probability	η	0.75	0.73	0.74	0.75	0.72
Application obsolescence probability	$P_{\omega,\mathrm{pre}}$	0.14	0.13	0.13	0.12	0.14
Renewal obsolescence probability	$P_{\omega,\mathrm{post}}$	0.04	0.04	0.04	0.04	0.04
Simple application fighting cost log-mean	$\mu_{f,\mathrm{simple}}$	8.53	8.43	8.56	8.60	8.53
Simple application fighting cost log-sigma	$\sigma_{f,\mathrm{simple}}$	0.87	0.97	0.79	0.74	0.95
SMM Objective		1.23	1.47	1.29	1.25	1.33

Notes: This table provides estimates of the model parameters across various model alternatives. The baseline model defines senior examiners as those at the GS14 level. The last column expands this to include GS13 and GS14.

Table A.6. Net Social Costs of Patent Prosecution: Robustness

	Patent Premium (ξ) = 0.10				Patent Premium (ξ) = 0.05					
Counterfactual	$\boldsymbol{T_1}$	$\boldsymbol{T_2}$ **(1.5)**	$\boldsymbol{T_3}$	**Total**	$\boldsymbol{T_1}$	$\boldsymbol{T_2}$ **(1.5)**	$\boldsymbol{T_2}$ **(2.0)**	$\boldsymbol{T_3}$	**Total (1.5)**	**Total (2.0)**
Baseline ($Bn)	6.4	0.7	17.6	24.7	6.6	0.0	0.2	20.6	27.2	27.4
25K Round Fee	5.9	1.8	16.4	24.1	6.3	0.7	1.4	19.1	26.1	26.8
50K Round Fee	6.1	3.1	15.1	24.2	5.5	1.7	3.5	17.1	24.7	26.1
Three Rounds	4.9	4.8	10.2	19.8	5.4	1.9	3.9	11.5	18.8	20.8
Two Rounds	2.9	7.4	4.7	14.9	2.9	3.2	6.6	5.2	11.4	14.8
One Round	0.0	6.3	0.7	7.0	0.0	1.6	3.3	0.8	2.4	4.1
15% IM	29.0	1.1	15.0	45.1	31.6	0.4	0.8	17.3	50.1	49.8
Credit$\searrow$	6.4	0.7	17.6	24.7	6.5	0.0	0.2	20.6	27.2	27.3
Credit$\searrow$ + 15% IM	24.3	1.9	15.8	42.0	23.7	0.7	1.5	18.2	47.8	43.3

Notes: This table provides the values of net social costs for alternative values of the patent premium and social multiplier. Columns denoted $\boldsymbol{T_2}$ **(1.5)** and $\boldsymbol{T_2}$ **(2.0)** provide values of type 2 net social costs when $\frac{\rho_{\text{soc}}}{\rho_{\text{priv}}}$ is equal to 1.5 and 2.0, respectively. Columns **Total (1.5)** and **Total (2.0)** provide the total net social costs when $\frac{\rho_{\text{soc}}}{\rho_{\text{priv}}}$ is equal to 1.5 and 2.0, respectively.

Figure A.1. Match of internal data and model moments

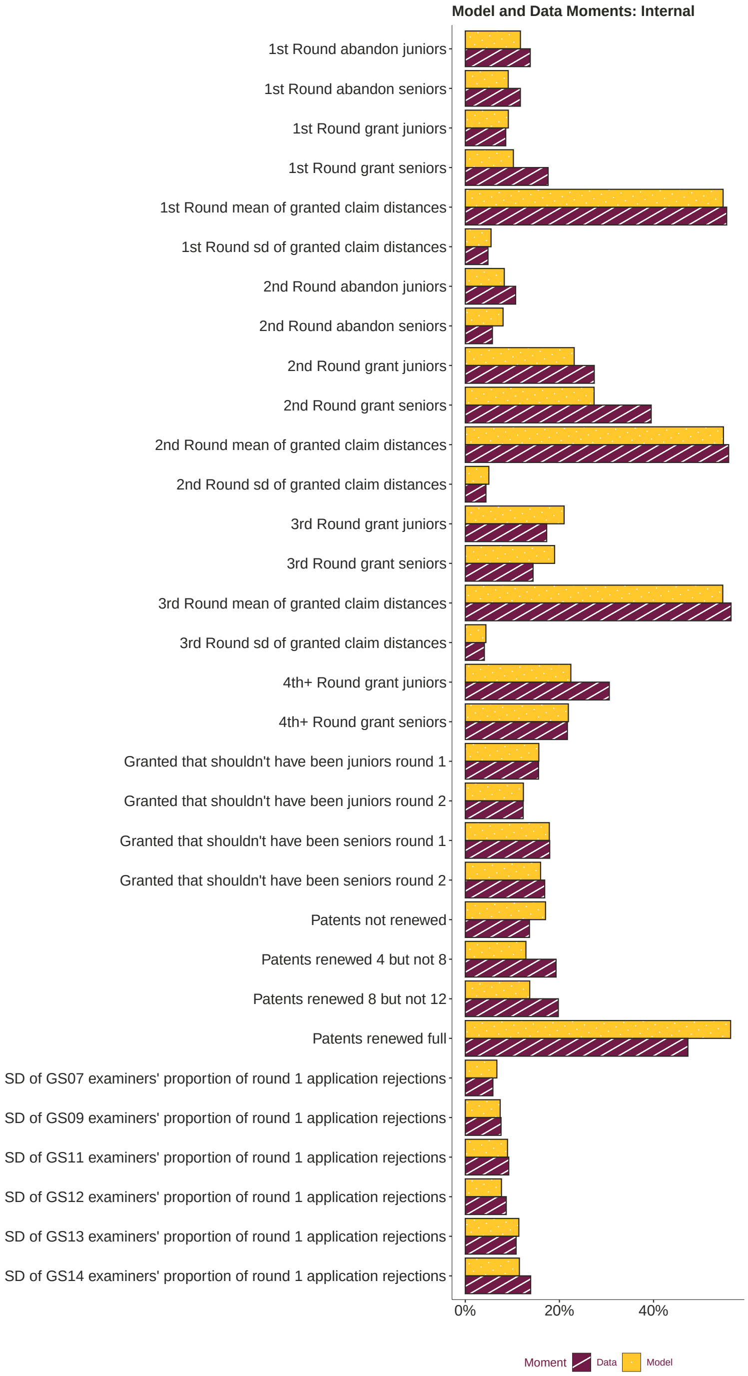
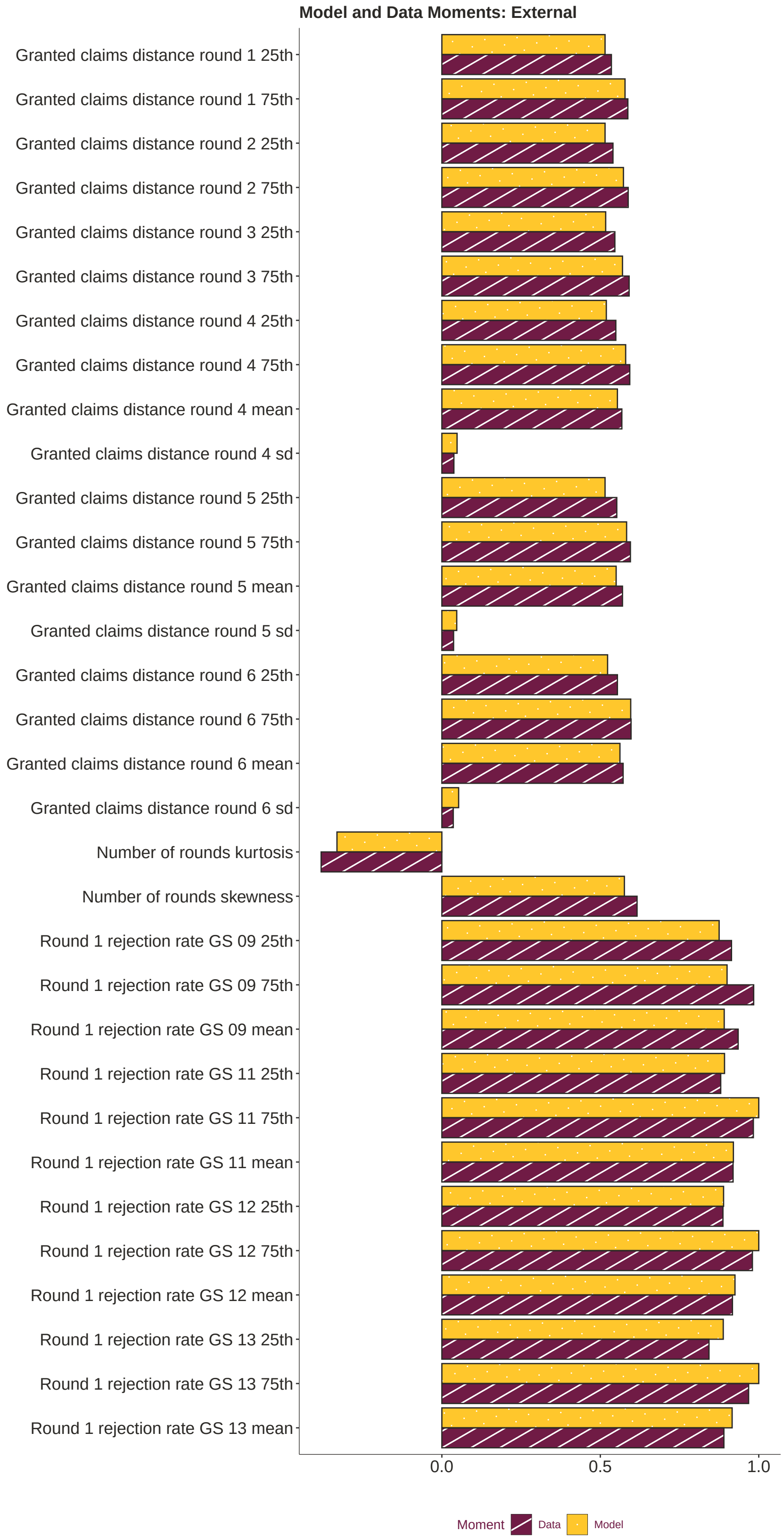


Figure A.2. Match of external data and model moments



B Data Sources

If the links are broken, the documents are available upon request.

B.1 Publicly Available Datasets

1. *U.S.PTO Patent Application Claims Full Text Dataset* and *U.S. PTO Patent Claims Full Text Dataset*: https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-claims-research-dataset

2. *Patent Examination Research Dataset*: https://www.uspto.gov/ip-policy/economic-research/research-datasets/patent-examination-research-dataset-public-pair

3. *U.S.PTO Maintenance Fee Events Dataset*: https://developer.uspto.gov/product/patent-maintenance-fee-events-and-description-files

4. *U.S.PTO Office Action Research Dataset*: https://www.uspto.gov/ip-policy/economic-research/research-datasets/office-action-research-dataset-patents

5. *Frakes and Wasserman (2019)*: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ABE7VS

B.2 Data from Public Documents

6. GDP Deflator: https://fred.stlouisfed.org/series/GDPDEF.

7. *AIPLA Report of the Economic Survey*: See https://www.aipla.org/detail/journal-issue/economic-survey-2017 for 2017.

8. Industry concentration: https://www.census.gov/content/dam/Census/programs-surveys/economic-census/data/archived_tables/2007/sector31/2007_31-33_Con_Ratios_US.zip.

9. Patent Office fees: https://www.govinfo.gov/content/pkg/CFR-2011-title37-vol1/pdf/CFR-2011-title37-vol1.pdf or from https://www.uspto.gov/sites/default/files/aia_implementation/AC54_Final_Table_of_Patent_Fee_Changes.pdf.

10. Patent operations costs:

    2005: https://www.uspto.gov/sites/default/files/about/stratplan/ar/USPTOFY2005PAR.pdf
    2010: https://www.uspto.gov/sites/default/files/about/stratplan/ar/USPTOFY2010PAR.pdf

2015: https://www.uspto.gov/sites/default/files/documents/USPTOFY15PAR.pdf

11. Patent applications: https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm.

12. R&D expenditures: https://www.nsf.gov/statistics/infbrief/nsf14307/.

C Distance Measure

This section provides details on how we construct our patent distance metric. We describe our preferred choice, the paragraph vector approach.[1] The method consists of four steps: (1) standardizing the independent claim text, (2) turning the text into a numerical vector, (3) calculating the distances between a focal patent claim on an application to all existing granted patent claims and (4) calculating the distance to the closest existing independent claim.

The first step before converting text into a numerical vector is text standardization. We perform basic changes to the content of the text and remove words that carry no informational content. Once we standardize the text, we drop any claims with fewer than two words or illegible text.

We use the paragraph vector approach to represent the text of a patent claim as a numerical vector. The paragraph vector approach is an improvement of the word vector approach. We implement the Paragraph Vector approach using Gensim's `Doc2Vec` Python model (Řehůřek and Sojka, 2010).

The step above converts all patent claims, including those on applications and those granted, into a numerical vector. The next step involves taking every focal application patent claim vector and calculating its distance to every *existing* granted claim at the point of application. After representing a patent claim's text as a numerical vector, we use cosine similarity and angular distance, both of which are standard in the text matching and the NLP literature. We compute the cosine similarity (CS) between claim text vectors x and y as

$$cs(x, y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_j x_j^2 \sum_j y_j^2}}.$$

Then, we calculate the angular distance (AD) metric, $AD(x, y) = \arccos(cs(x, y))/\pi$ and then double AD to obtain a normalized distance in the interval $[0, 1]$.

[1] At the time of writing this paper, we used the state-of-the-art approach, but there is a fast-moving frontier. The most recent approaches use GPT-4 or BERT word embeddings integrated directly into Neural Networks. See Elliot and Hansen (2023) for details on text algorithms.

With all distances computed, it is a simple step to find the closest 50 claims to each application. We experiment with different choices on which percentile of the closest 50 distances to use. We also experimented with taking an average of the five closest distances for example, and the resulting distances were similar.

D Descriptive Results

We show how patent application outcomes vary with technology center and examiner seniority. First, we regress a binary variable equal to one if the application process lasts more than one round against fixed effects for examiner seniority grade, technology center, year of application, and a small entity indicator (applying firm having fewer than 500 employees). The results in Column (1) of Table D.1 reveal substantial variation across technology centers; e.g., Computer Networks (TC-24) has a 12 percentage point higher likelihood of multi-round negotiation than the reference category, Biotechnology (TC-16). Further, the likelihood of any negotiation decreases with the seniority of the examiner, with senior (GS-14) examiners nine percentage points less likely to require negotiation relative to the most junior, holding technology center and application year fixed. Further, small entities are 12 percentage points less likely to negotiate (all else fixed).

In Column (2), we do the same analysis for the dependent variable equal to one if the examiner grants a patent. We match the findings of Frakes and Wasserman (2017) – senior examiners are more likely to grant and grant rates vary substantially across technology centers. In our model, we explain this variation by letting the distribution of intrinsic motivation vary with seniority level, by incorporating differences in the credit structure for examiners that vary across seniority and technology centers, and by allowing fighting costs to differ for applicants, with technology category-specific distributions. Our parameter estimates enable us to disentangle the effects of these factors in explaining the variation in outcomes, as we discuss in the text.

Table D.1. Regression Results

Variable	(1) Negotiation	(2) Grant
Intercept	0.7433 (0.006)	0.542 (0.005)
GS-7	-0.002 (0.004)	0.003 (0.005)
GS-9	-0.016 (0.004)	0.035 (0.004)
GS-11	-0.020 (0.004)	0.066 (0.004)
GS-12	-0.034 (0.004)	0.092 (0.004)
GS-13	-0.045 (0.004)	0.126 (0.004)
GS-14	-0.091 (0.004)	0.178 (0.004)
Chemicals (17)	0.064 (0.002)	0.067 (0.002)
Comp. Software (21)	0.105 (0.002)	0.196 (0.002)
Comp. Networks (24)	0.123 (0.002)	0.192 (0.002)
Communications (26)	0.047 (0.002)	0.198 (0.002)
Electronics (28)	-0.010 (0.001)	0.244 (0.001)
Other (36)	0.065 (0.002)	0.136 (0.002)
Mech Engineering (37)	0.042 (0.002)	0.139 (0.001)
Small Entity	-0.120 (0.001)	-0.170 (0.001)
Year FE	Yes	Yes
N	1,641,333	1,759,313

Notes: Omitted grade is GS-5 and omitted technology center is Biotechnology and Organic Fields (16). Technology center "Other" refers to Center 3600, which is "Transportation, Electronic Commerce, Construction, Agriculture, Licensing and Review." Following Frakes and Wasserman (2017), we omit GS-15 grade examiners. We report heteroskedasticity robust (HC1) standard errors in parentheses.

These results show stark differences in average grant rates and likelihood of negotiation across technology centers and examiner seniority grades. Next, we investigate the variation in *examiner-specific* decisions within and between seniority grades and technology center pairs. To do this, we calculate examiner-specific outcomes (average grant rates, number of rounds, length of examination period, probability of negotiation, etc.) within each seniority grade examiners are in at the time. We decompose the variation in these examiner averages into within and between seniority grade-technology center pairs by introducing dummies for each seniority-grade-technology-center dyad in Table D.2. The proportion of within-group variation in examiner grant rates is 80%, im-

Table D.2. ANOVA Results

Variable	**Grade × TC Fixed Effects**
Grant rate	79.84
Duration of examination (years)	75.79
Number of rounds	80.89
No negotiation (one round)	89.53
Independent claims granted	74.93

Notes: For each variable y, and an examiner e when they are in seniority grade S and technology center T, we calculate $\bar{y}_{eST}$. Then we regress $\bar{y}_{eST}$ on a set of interactive dummies for seniority grade and technology center. We report $1 - R^2$ (as a percentage) for these regressions, thereby providing the proportion of within group variation.

plying substantial variation in examiner grant rates not explained by seniority and technology centers. Our model explains this variation in examiner-specific grant rates within the technology center and seniority groups by incorporating group-specific *distributions* of examiner intrinsic motivation and costs of delay.

E Examiner Credit Structure

Here we provide expressions for $g^r_{GR}(S,T)$, $g^r_{ABN}(S,T)$, $g^r_{RCE}(S,T)$ and $g^r_{REJ}(S,T)$. For $y \in \{GR,\ ABN,\ REJ,\ RCE\}$, we write $g^r_y(S,T) = \nu^r_y \cdot c(S,T)$, and give expressions for ν^r_y and $c_(S,T)$ separately.

E.1 Credits

Granting in the first round gives the examiner a payoff of $\nu^1_{GR} = 2$ credits. Rejecting in the first round gives $\nu^1_{REJ} = 1.25$. If the applicant abandons in round one, the examiner obtains $\nu^1_{ABN} = 0.75$. Granting in the second round gives $\nu^2_{GR} = 0.75$ credits. Rejecting in the second round gives $\nu^2_{REJ} = 0.25$ credits, with an extra $\nu^2_{ABN} = \nu^2_{RCE} = 0.5$ credits whether the applicant abandons or continues to an RCE. Ultimately, the examiner obtains two credits irrespective of what happens in the first two rounds. The only difference is whether they obtain the credits immediately (say, from an immediate grant) or spread out over two rounds.

The structure of the payoffs in the first RCE are the same, except $\nu^3_{REJ} = 1$ and $\nu^3_{GR} = 1.75$. In this case, irrespective of what happens in the RCE, the examiner will obtain 1.75 credits. The

Table E.1. Seniority Corrections

Seniority Grade	**Signatory Authority**	$c_{SEN}(S)$
GS-5	None	0.55
GS-7	None	0.7
GS-9	None	0.8
GS-11	None	0.9
GS-12	None	1.0
GS-13	None	1.15
GS-13	Partial	1.25
GS-14	Partial	1.25
GS-14	Full (primary examiner)	1.35

Notes: This table provides the seniority factors for credit adjustment. In the empirical work, we use 1.15 for GS-13 and 1.25 for GS-14.

difference comes from whether they receive all 1.75 credits at once by granting, or 1 credit from their non-final rejection and $\nu^4_{REJ} = 0.25$ plus $\nu^4_{ABN} = \nu^4_{RCE} = 0.5$ credits from the applicant's response.

In the second and any subsequent RCEs, the structure of the payoffs is still the same, except $\nu^{2r+1}_{REJ} = 0.75$ and $\nu^{2r+1}_{GR} = 1.5$ ($r > 1$). As before, the examiner will receive 1.5 credits from second and subsequent RCEs. The difference comes from whether they receive all 1.5 credits at once from granting, or 0.75 credits from their non-final rejection and $\nu^{2r+2}_{REJ} = 0.25$ plus $\nu^{2r+2}_{ABN} = \nu^{2r+2}_{RCE} = 0.5$ credits from the applicant's response.

E.2 Seniority and Technology Complexity Adjustments

The seniority and technology complexity adjustment term is

$$c(S,T) = \frac{c_{TECH}(T)}{c_{SEN}(S)}.$$

Table E.1 gives the values of $c_{SEN}(S)$ across the GS categories. Higher seniority factors imply larger values of c_{SEN}, and therefore lower values of credits. Table E.2 gives the values of $c_{TECH}(T)$ we created for the different technology centers and use in the estimation of the model. The Patent Office does not have adjustments at the technology center level, but rather at the more detailed U.S. Patent Class (USPC) level. We obtained the adjustments at the USPC level from the Patent Office and constructed a patent-application weighted average for each technology center.

Table E.2. Technology Center Adjustments

Technology Center T	**U.S.PTO Number**	**Correction** $(c_{TECH}(T))$
Chemical and Materials Engineering	17	22.2
Computer Architecture Software and Information Security	21	31
Computer Networks, Multiplex, Cable and Cryptography/Security	24	29
Communications	26	26.5
Semiconductors, Electrical and Optical Systems and Components	28	21.4
Transportation, Electronic Commerce, Construction, Agriculture...	36	22.4
Mechanical Engineering, Manufacturing and Products	37	19.9

F Moment Selection and Identification Intuition

First, we provide further details on the possible moments we could use to estimate our model. Then, we provide some information on our methods to prune moments from the full set. Finally, we provide some intuition on how the moments identify the model parameters.

F.1 Available Moments

We have seven sets of moments available, which we describe in turn.

Our first group of moments corresponds to examiners' issuance and applicants' abandonment decisions. For each round in the model and each seniority level, we calculate the proportion of applications examiners grant and the proportion that applicants abandon. Since there are nine seniority grade-signatory authority pairs, and we observe at least six rounds, this implies at least 108 moments on grants and abandonments.

Second, we observe the distribution of the proportion of claims rejected, both by round (six) and by seniority grade-signatory authority pair (nine). These observations generate another 54 moments. Third, we observe the proportion of granted patents that renew at four, eight, and twelve years after issuance. These observations generate four moments on patent renewals (don't renew at four, renew at four but not eight, renew at eight but not twelve and renew at twelve).

Fourth, we calculate the distribution of claim distances by round. We calculate the mean and standard deviation of the distance distribution by round for at least six rounds, implying at least 12 moments on distance. Another moment comes from the within-application distance correlation. Fifth, at each of the nine seniority grades, we calculate each examiner's *leniency*, which is their average rejection rate across all the applications they examine. Hence for each seniority grade-signatory authority pair, we obtain a distribution of examiner rejection rates, for which we can calculate the mean and standard deviation of the distribution of examiner fixed effects. From this we obtain another 18 moments.

Next, given that we can identify the distance threshold externally, we calculate the proportion of granted patents containing at least one invalid claim (that is, a claim whose distance is below the distance threshold). Hence, for each round and each seniority level, we calculate the proportion of patents granted containing an invalid claim, implying another 54 moments.

Finally, we observe the distribution of application fighting costs. We have six moments on the distribution of legal application fees for four technology categories (simple, chemical, electrical and mechanical), which we match to the technology centers on which we estimate the model. This implies another 24 moments.

F.2 Choosing Moments

We have more than two hundred data moments that we can calculate from endogenous variables in the model. Since we have 21 model parameters to estimate with simulated method of moments, in principle, we are over-identified. However, not all moments will aid the estimation procedure in identifying the parameters, so we begin by pruning the set of moments for estimation.

We follow a rigorous, data-driven methodology to create a subset of the moments that best estimate the parameters. To do this, we calculate the sensitivity matrix described in Andrews, Gentzkow, and Shapiro (2017). As the authors explain, "sensitivity gives a formal, quantitative language in which to describe the relative importance of different moments for determining the value of specific parameters." If a moment had a small value in the sensitivity matrix for all parameters, we considered it as not useful in estimating our model. Further, as described in Jalali, Rahmandad, and Ghoddusi (2015), for each parameter and moment, we plot the value of the moment for different values of the parameter, fixing the other parameters at their estimates. If this curve is flat, this parameter does not influence on the value of the moment. For a given moment, if the curve is flat across *all* parameters, it suggests that the moment offers no useful variation to identify the parameters.

For each parameter, we also plot the value of the SMM objective across all values of the parameter, fixing other parameters at their estimates. Ideally, the SMM will be U-shaped in each parameter to ensure a well-defined global minimum exists. By doing this, we learn how well we pin down parameters based on the set of moments we have available.

By combining the sensitivity matrix with moment and SMM plots, we pruned the set of moments down to those that offer some assistance in estimating the parameters. Since we split many parameters into two seniority groups (junior and senior), we split some of our moments into the same seniority categories.

F.3 Full Set of Moments

The full set of moments we use for estimation is as follows. The selected moments corresponding to outcomes for examiners are:

(i) The proportion of applications granted in each round for juniors and seniors, for rounds one, two, three, and all rounds after four combined [eight moments]

(ii) The standard deviation of the distribution of examiner rejection rates for the six seniority categories used by the Patent Office (GS levels 7, 9, 11, 12, 13, and 14) [six moments]

(iii) The proportion of patents granted containing an invalid claim (for juniors and seniors) for rounds one and two [four moments]

The moments corresponding to outcomes for applicants are:

(i) The proportion of abandonments in each round, when the assigned examiner is junior and senior, for rounds one and two [four moments]

(ii) The proportion of granted patents not renewed, renewed at year four but not eight, renewed at year eight but not twelve, and renewed at year twelve [four moments]

(iii) The mean and standard deviation of the distribution of granted claim distances for rounds one, two, and three [six moments]

(iv) Mean and median of legal application fees for simple applications and complex applications in electrical, mechanical, and chemical technologies [eight moments]

F.4 Identification

A model is either point identified or not, and technical conditions on the required variation in exogenous variables determine whether a model is identified (Andrews, Gentzkow, and Shapiro, 2017). Due to our model's complicated and nonlinear nature, we cannot calculate these conditions. Identification with simulated method of moments is based on how different moments are affected by specific parameters. While we cannot identify this link exactly, we provide some intuition of how moments aid in pinning down specific parameters of the model.

We start with the parameters relating to the applicant. The renewal rates, together with first-round abandonment decisions, aid in identifying the parameters of the distribution of flow returns, i.e., μ_v and σ_v. This is because, all else equal, an applicant with higher returns is less likely to abandon after learning their examiner and more likely to renew their patent, conditional on being granted. The renewal moments also aid in identifying the post-grant obsolescence probability $P_{\omega,\text{post}}$. Similarly, the ex post claim distribution of padded distances, as calculated using the distance between text vectors, aids in identifying the parameters of the distribution of ex ante unpadded distance, i.e., α_D and β_D. Moments on application fighting costs directly pin down the distribution of application fighting costs, $\mu_{f_{app}}$, and $\sigma_{f_{app}}$.

Regarding pre-grant obsolescence $P_{\omega,\text{pre}}$, the only case in which an applicant abandons in interim rounds two to four is when they become obsolete. If an applicant, upon learning their examiner calculates that they will want to abandon in any round after the first, they will abandon immediately in round one. Therefore, interim round abandonments offer substantial assistance in identifying the obsolescence probability in the application process.

Intuition for examiner parameters is more complicated. Observing that examiners grant several invalid patents could result from low intrinsic motivation, high examiner error, or high examiner delay costs. Three factors make this challenge less formidable. First, since we assume that only intrinsic motivation varies by seniority, differences in grant rates and examiner errors by seniority pick up the value of intrinsic motivation, μ_{im} by seniority, and differences in the variation in examiner-specific grant rates by seniority capture the variation in intrinsic motivation, σ_θ by seniority.

Second, we assume that each examiner has the same delay cost across all applications and rounds but faces varying intrinsic motivation costs at each round of every application (because $\mathcal{R}^r$, the proportion of invalid independent claims varies across rounds and applications). This implies that the proportion of invalid patents granted in rounds one and two offer the best assistance in identifying the mean examiner intrinsic motivation and mean examiner delay costs. Third, examiner error is two-sided and symmetric. This feature creates cases where examiners do not grant valid patents, whereas intrinsic motivation and delay costs only incentivize examiners to grant when they should not. Otherwise, we know that an examiner, making no mistake, and facing a fully valid patent, will always issue it. Together, this implies that we can use the residual variation in grant rates (valid and invalid) by round and seniority to learn about the distribution of examiner error.

F.5 Details on Model Fit

As shown in Figure A.1, we match most of the internal moments well, though there are two exceptions. The first is the proportion of fully renewed patents, which we overestimate. The other exception is the second-round grant rate. This moment is difficult to match with our model because examiners have incentives to wait until the third round and obtain RCE credits if they do not choose to grant in the first round. Since examiners have incentives and targets across applications on their desks (docket management), they are more likely to grant in the second round than our baseline model predicts.

G Quantification of Social Costs

G.1 Implementing Type 1 Social Cost Calculation

As indicated in the text, a key challenge in implementing our calculation of type 1 social costs comes from the fact that the estimates of the value of patent rights for invalid patents include potential litigation costs. To impute the "value at stake" in litigation for these patents, we need to adjust our methodology to exclude these costs.

To do this, we make two assumptions:

A1: Valid patents are not litigated. This assumption holds in a model with perfect courts, where a competitor knows (or can pay a fee to discover) whether a patent is valid or not, and then choose whether to litigate based on the result.[2] This assumption allows us to calculate the value of patent rights for valid patents, $\tilde{V}$, as equal to the observed value since there are no litigation costs to net out.

A2: The *distribution* of the value at stake, $G_{\tilde{V}}(\cdot)$, is the same for valid patents as invalid patents. The basis for this assumption is that initial distances and values are uncorrelated in the model. This assumption allows us to draw values from the observed distribution of $\tilde{V} = V$ for valid patents and use them as draws from the distribution of $\tilde{V}$ for invalid patents.

Given A1 and A2, the procedure for calculating type 1 social costs is as follows:

1. Estimate the parameters of a log-normal distribution for the value at stake for *valid* patents.[3] Let the estimated distribution be denoted as $\hat{G}_{\tilde{V}}(\cdot)$.

[2]This assumption is *not* at odds with Schankerman and Schuett (2022), where *high types* are litigated with some probability even though they will not be invalidated. The important point is that high types in their model (patents that would not be developed without patent rights) are not the same as valid patents in our model, which are defined as those with distance larger than the threshold.

[3]The sum of log-normal distributions is approximately log-normal (Dufresne, 2004), which our simulation here exhibits.

2. Let $\bar{P}$ be the total number of *invalid* patent grants for the given period we simulate. Then, for each $p = 1, \ldots, \bar{P}$:

   (a) Take a draw from the estimated distribution of *valid* patents' value at stake (ex post value), $\hat{G}_{\tilde{V}}(\cdot)$, to represent the value at stake for the invalid patent p

   (b) Using the draw, calculate S_{1p} from Equation (11).

3. Calculate the total social cost of type 1 error as $\sum_{p=1}^{\bar{P}} S_{1p}$.

Finally, note that we calculate the threshold for exposure to litigation from the *empirical* distribution of the value at stake for valid patents, $\hat{G}_{\tilde{V}}(\cdot)$.

G.2 Implementing Type 2 Social Cost Calculation

The primary challenge in implementing our calculation of type 2 social costs comes from calibrating the value of the invention without patent rights (π), particularly for inventions with $\Gamma^* \leq 0$, where we cannot use the patent premium. In a similar vein to our approach to type 1 social costs, we assume that the distributions of π for those with positive and negative Γ^* are the same and then draw values of π from this distribution for those inventions.

To be precise, our specific implementation is as follows:

1. Draw a pilot set of potential inventions, used to calculate a distribution of π. Run these set of potential inventions through the model and calculate Γ^*. For those with positive Γ^*, create a distribution of π using the relationship $\Gamma = \xi\pi$.

2. Now start the simulation for type 2 social costs by drawing a new set of potential inventions (returns, distances, number of claims, fighting costs, examiner etc.). For each potential invention i, calculate Γ_i^*. If $\Gamma_i^* > 0$, calculate $\pi_i = \frac{\Gamma_i^*}{\xi}$. If $\Gamma_i^* \leq 0$, draw a value of π_i from the distribution calculated in 1. Also, draw a development cost κ_i.

3. For each of the potential inventions i, work out the set $i = 1, \ldots, \mathcal{I}_{\text{no dev}}$ that do not develop as those with $\max\{\Gamma_i^*, 0\} + \pi_i < \kappa_i$

4. For $i = 1 \ldots, \mathcal{I}_{\text{no dev}}$, run the potential invention through a model where, at the point of abandonment, the inventor obtains all valid claims they have, and so obtains the patent value of their valid claims, instead of a payoff of 0. By definition, this scenario has the property that all abandoned claims are invalid, so that there is no type 2 error. Let Γ_i' denote the expected value of patent rights in this new scenario.

5. For $i = 1\ldots, \mathcal{I}_{\text{no dev}}$, calculate the set $i = 1, \ldots, \mathcal{I}_{\text{now dev}}$ who have $\max\{0, \Gamma_i'\} + \pi_i \geq \kappa_i$. This is the set who do not develop because of type 2 error but do develop in the absence of type 2 error.
6. For $i = 1, \ldots, \mathcal{I}_{\text{now dev}}$, calculate $S_{2i} = \frac{\rho_{\text{soc}}}{\rho_{\text{priv}}}\Big(\max\{0, \Gamma_i\} + \pi_i\Big) - \kappa_i$ and calculate the total type 2 social cost as

$$T_2 = \sum_{i=1}^{\mathcal{I}_{\text{now dev}}} S_{2i}.$$

G.3 Calibrating Deadweight Loss

In the derivation of deadweight loss, note that

$$DWL = \frac{1}{2}\Delta\wp\Delta q = \frac{1}{2}\frac{\Delta q}{q}q\Delta\wp = \frac{\lambda}{2}\frac{\Delta\wp}{\wp}\tilde{V},$$

by the definitions of $\tilde{V}$ and λ. Further, note that

$$\frac{\Delta\wp}{\wp} = \frac{q\Delta\wp}{q\wp} = \frac{\text{lic. rev}}{\text{sales}} = \frac{\text{lic. rev}}{\text{R\&D}} \cdot \frac{\text{R\&D}}{\text{sales}}$$

As described in the text, we use Schankerman and Schuett (2022) for the ratio of licensing revenue to R&D, and data from the Bureau of Economic Analysis for the ratio of R&D to sales.

G.4 Deadweight Loss Under Cournot Competition

In the main text, we compute deadweight loss from a patented invention assuming symmetric licensees operate in a perfectly competitive industry. Suppose instead that the licensees compete in a Cournot setting. By standard calculations, the equilibrium price-cost margin is $\frac{\wp - c}{\wp} = \frac{m^*}{\lambda}$ where $m^* = \frac{1}{N}$ is the average market share and λ is the demand elasticity. We write this as $\frac{\wp - c}{\wp} = \frac{H^e}{\eta}$ where H^e is the symmetric-equivalent Herfindahl index of concentration. Thus for $H^e < 1$

$$\wp = \frac{c}{1 - \frac{H^e}{\lambda}}.$$

With imperfect competition, the change in equilibrium price is larger than the Arrow royalty due to double marginalization: $\Delta\wp = \frac{\Delta c}{1-\frac{H^e}{\lambda}} > \Delta c$. The associated deadweight loss with Cournot competition is

$$DWL_{\text{cournot}} = \frac{1}{2}\Delta\wp\Delta q = \frac{1}{2}\frac{\Delta c}{1 - \frac{H^e}{\lambda}}\Delta q = DWL_{\text{pc}} \cdot \frac{1}{1 - \frac{H^e}{\lambda}},$$

where it should be noted that in this case $\tilde{V} = q\Delta c$ denotes total *royalty payments*. Since $H^e \in (0, 1)$ and we require that $|\lambda| > 1$, deadweight loss in this imperfect competition setting is larger than in perfect competition case.

Using U.S. Census data for 2007, the value added weighted-average Herfindahl index for manufacturing industries (based on the 50 largest firms), H, for manufacturing sectors is 0.05. As is well-known,the Herfindahl index can be decomposed as $H = \frac{1}{N} + N \cdot \text{Var}(m) = H^e + N \cdot \text{Var}(m)$, where m is the market share of each firm. Thus, the observed H overstates the unobserved H^e, so the computed deadweight loss will be an upper bound to the true value of DWL. Despite this, the upper bound for the Cournot setting is not materially different from the competitive case in the text.

The value of H varies widely across industries. We do not compute deadweight loss using industry-specific values because it is difficult to assign patents in different patent classes to industries, and the existing Patent Office concordance is problematic (e.g., the mapping is not unique).

G.5 Calibrating Litigation Costs

To calibrate litigation costs, $\mathcal{C}(\tilde{V})$, we use data from the American Intellectual Property Law Association (AIPLA) surveys on litigation costs as a function of (intervals) of the value at stake, which we assume is the same for the patentee and challenger. We use the linear specification

$$\mathcal{C}(\tilde{V}) = \ell_0 + \ell_1 \tilde{V}$$

Using this same specification, Schankerman and Schuett (2022) estimate $\ell_0 = \$624,000$ and $\ell_1 = 0.162$ (2018 USD). Note that this calibration of legal costs is at the patent, not claim, level.

G.6 Calibrating Development Costs

We apply the estimates from Schankerman and Schuett (2022) to our context. They assume that development costs κ are exponential, with mean equal to $k_0 + k_1 s$, where s is the size reduction of the invention and k_0 and k_1 are estimated as 254.6×10^3 and 2.33×10^{10}, respectively. Regarding the size reduction, they assume that s is log-logistic distributed with parameters $\beta_0 = 1.02$ and $\beta_1 = 1.14 \times 10^{-6}$. We use the mean value of s in our calibration.

In the baseline quantification, we draw values of κ from the distribution described above, which assumes that development costs are independent of Γ^* and π. In this model, inventors know their development costs prior to their decision to develop their idea. We also experiment with another model, which makes the opposite assumption that inventors do not know their development costs and thus use the mean value, $\bar{\kappa} = k_0 + k_1 \bar{s}$, to make their development decision. Both models produce similar conclusions; results are available upon request.

G.7 Calibrating the Number of Ideas

To compute the number of ideas, we start with the average annual number of utility patent applications in the period 2011–2013. We convert this number into the number of ideas in two steps. First, we use the estimates from Schankerman and Schuett (2022) that about two-thirds of applications are "low type" inventions (defined by them as those that would have been developed even without patent protection), and second, that one-third of ideas become a low type patent application. Together, this implies about one million ideas for potential inventions for each cohort of applications.

References

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

Dufresne, D. (2004): "The Log-Normal Approximation in Financial and Other Computations," *Advances in Applied Probability*, 36, 747–773.

Elliot, A. and S. Hansen (2023): "Text Algorithms in Economics," *Unpublished Working Paper*.

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

Jalali, M., H. Rahmandad, and H. Ghoddusi (2015): *Using the method of simulated moments for system identification*, MIT Press.

Řehůřek, R. and P. Sojka (2010): "Software Framework for Topic Modelling with Large Corpora," *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50.

Schankerman, M. and F. Schuett (2022): "Patent Screening, Innovation, and Welfare," *The Review of Economic Studies*, 89, 2101–2148.