Patents and Cumulative Innovation – Evidence from Post-Grant Patent Oppositions

Fabian Gaessler [a] Dietmar Harhoff [abc] Stefan Sorg [ad]

[a] *Max Planck Institute for Innovation and Competition, Munich*
[b] *Munich School of Management, Ludwig-Maximilians-University (LMU), Munich*
[c] *Centre for Economic Policy Research (CEPR), London*
[d] *Munich Graduate School of Economics, Ludwig-Maximilians-University (LMU), Munich*

June 13, 2017

ABSTRACT

Using large-scale data on opposition to patents at the European Patent Office (EPO), we investigate the causal effect of a patent's invalidation on follow-on inventions. We introduce a new instrumental variable exploiting the participation or absence of the patent examiner in the opposition proceeding. According to our baseline model, patent invalidation leads to a highly significant and sizeable increase of forward citations. While this is in line with previous studies, disentangling the effect leads us to results that stand in contrast to some of the literature. We find that the effects are most pronounced for patents in discrete technology areas, for areas where patent thickets are absent and for patents which are not protected by "patent fences". Moreover, the effect is particularly strong for small patent holders facing small follow-on innovators. We confirm these results using technology-specific samples of opposition cases, and citation measures based on either EPO or US Patent and Trademark Office citation data. (156 words)

KEYWORDS: cumulative innovation, patents, opposition.

JEL Classification: K41, L24, O31, O32, O33, O34

We thank Alberto Galasso, Klaus Gugler, Peter Haan, Yassine Lefouili, Yann Ménière, Maikel Pellens, Georg von Graevenitz, and Tetsuo Wada for their valuable comments. We also like to thank participants at the 16th REER Conference, the Annual Meeting of the Committee for Industrial Economics, the INNOPAT Conference, the SEI Faculty Workshop, and the Munich Summer Institute.

1 Introduction

Patents are considered a key policy instrument to spur innovation and technological progress. With a patent grant, inventors receive temporary exclusion rights in return for the creation and disclosure of their inventions. Inventions are rarely stand-alone achievements, but build to a large extent on previous inventions – "cumulative invention" has become a dominant characteristic of the modern industrial innovation apparatus. But the cumulative nature of technical progress may also cause major impediments for research and development. When inventions build on each other, exclusion rights on a preceding invention may limit the attractiveness of subsequent invention steps that build on the protected invention. Whether such distortions of research incentives exist, is ultimately an empirical question. This paper contributes to the literature on cumulative invention by providing an econometric analysis of patent invalidation at the European Patent Office. If patent invalidation is followed by "new" research and patenting activity, then this can be taken as evidence for the existence of such impediments. Our empirical results provide such evidence and allow us to identify situations in which the effects of patents on cumulative innovation are particularly pronounced.

Cumulative innovation and the knowledge spillovers that underpin it lie at the heart of the recent macroeconomic literature on innovation and endogenous growth, e.g., Grossman and Helpman (1991), Aghion and Howitt (1992), and Acemoglu and Akcigit (2012). However, a number of theoretical contributions have illustrated that the incentive created for one invention via a patent right may have a delaying or cost-increasing effect on follow-on inventions (see Hall and Harhoff (2012) for a literature survey). In general, the patent publication provides valuable information that allows follow-on inventors to build upon the inventions. The disclosure and the resulting knowledge spillovers are commonly seen to facilitate cumulative invention (Scotchmer and Green, 1990). But intellectual property rights on existing technologies require coordination between original and follow-on inventors (Scotchmer, 1991) which often takes the form of licensing agreements. In the absence of transaction costs, the parties involved in a cumulative invention process could reach a licensing agreement such that cumulative innovation is not blocked. However, due to transaction costs, such negotiations may not succeed or inefficiencies associated with licensing outcomes may limit the attractiveness of invention efforts for the parties involved. If that is the case, we would expect to see detrimental effects on technological progress and, ultimately, on economic growth.

Patent thickets and fragmented patent ownership have been identified as potential impediments to efficient licensing agreements and causes for bargaining failure (Heller and Eisenberg, 1998; Shapiro, 2001; Gallini and Scotchmer, 2002; Lemley and Shapiro, 2007; Bessen and Maskin, 2009). Follow-on innovation is particularly prone to patent encounter blockage in industries with complex and modular technologies and among small firms and market entrants that lack leverage for cross-licensing deals (Lanjouw and Schankerman, 2004; Cockburn et al., 2010). Hence, one may expect a positive effect on follow-on innovation by others after patent invalidation occurs in industries characterized by complex products or in cases where follow-on inventions would come from small players (Galasso and Schankerman, 2015).

However, this argument does not take into account the possibility of strategic patenting and differences in the effectiveness of patent rights across technologies (Teece, 1986; Ziedonis, 2004;

Harhoff et al., 2007). Moreover, as Cohen et al. (2000) have argued, cases in which one patent protects one product ("discrete" technologies) are rare. Inventions in "complex" technology areas are often protected by multiple patents, so that the reduction in protection from losing one patent could be relatively small compared to losing a patent in discrete technology areas. Furthermore, large patent portfolios with overlapping claims and dense patent thickets could marginalize the gain from the invalidation of a previous patent. Leaving aside the size of the patent portfolio, there are other complementary assets that may determine a firm's ability to exclude other parties. As the existence of complementary assets is likely correlated with the size of the patent holder, large patent holders should be more able to compensate for the loss of patent protection.

Hence, the gain in freedom to operate and to conduct R&D following patent invalidation could be larger in discrete technology areas and in cases where the focal patent holder cannot maintain protection with the help of overlapping patent claims or other complementary assets. We thus argue that the effect of invalidation on follow-on research by third parties should then be strongest where appropriability heavily relies on patent rights.

Several empirical studies as to if and where patents hinder follow-on innovation have been undertaken recently (see Table 1 for an overview). The identification strategies in these studies primarily exploit quasi-exogeneous variations in patent protection over time under the assumption that follow-on inventors would require a license from the upstream patent holder as long as the patent is enforceable. Since licensing agreements usually remain undisclosed, measures of follow-on innovation have had to rely on references to the focal invention in subsequent work. Commencing this stream of literature, Murray and Stern (2007), Huang and Murray (2009), and Williams (2013) focus on IP in biotechnology and analyze whether the protection of a particular genome sequence has any effect on follow-on activities, witnessed by either scientific studies, patents, or product development. Murray and Stern (2007) and Huang and Murray (2009) use differences-in-differences estimation models, exploiting granted patent protection as variation over time and gene sequences, with only a subset being the subject of a patent application in the first place. Both studies conclude that patent protection on genes impede subsequent research. Huang and Murray (2009) find this blocking effect to correlate with patent scope, patent thickets, and fragmented patent ownership. The results of Williams (2013) suggest that (non-patent) IP rights on a specific set of genes led to a 20-30 % decrease in subsequent scientific research and product development. Sampat and Williams (2015) further investigate the relationship of patent rights and follow-on innovation on human genes by comparing citations to successful and unsuccessful patent applications filed at the USPTO. To avoid issues arising from the presumable endogeneity of the patent grant event, they employ an instrumental variable based on the leniency of the respective patent examiner. The results of their analyses do not provide evidence for a blocking effect of human gene patents on follow-on innovation.

The exclusivity of patent protection is effectively limited in a compulsory licensing regime. Moser and Voena (2012) and Watzinger et al. (2016) each focus on cases where a set of patent rights became *de facto* ineffective in excluding others due to compulsory free licensing. Notably, Moser and Voena (2012) find an increase in innovation from compulsory free licensing in the chemical

Table 1: Prior empirical studies on patent rights and cumulative innovation

Study	Dependent variable	Identification	Technology	Sample
Patent grant				
Murray and Stern (2007)	Scientific citations	DiD estimation	Biotech	169 patent-paper pairs
Huang and Murray (2009)	Scientific citations	DiD estimation	Biotech	1,279 patent-paper pairs
Sampat and Williams (2015)	Scientific citations	IV (examiner fe)	Biotech	292,655 patent-gene pairs
Patent invalidation				
Galasso and Schankerman (2015)	Patent citations	IV (judge fe)	All	1,357 patents
Galasso and Schankerman (2016)	Patents	IV (judge fe)	All	1,469 patents
Compulsory licensing				
Moser and Voena (2012)	Patent citations	DiD estimation	Chemistry	130,000 patents
Watzinger et al. (2016)	Patent citations	DiD estimation	IT	4,509 patents

Notes: DiD = difference-in-differences; fe = fixed effects (or similar).

sector. Watzinger et al. (2016) look at patents of Bell Labs and find that compulsory free licensing favored in particular follow-on inventions by small and young firms. Both studies focus on particular technologies and use historical data. Given recent changes in patenting systems not all of the results may apply to the current context.

Most similar to our study, Galasso and Schankerman (2015) investigate the effect of patent invalidations by the U.S. Court of Appeals for the Federal Circuit (CAFC) on follow-on innovation. They address endogeneity of the patent invalidation event by exploiting the randomized allocation of judges at the CAFC to identify judge fixed effects. In a complementary study, Galasso and Schankerman (2016) use the same empirical setting to analyze the effect of patent invalidation on subsequent research activities of the focal patent holder.[1]

These studies are not fully comparable as the quasi-experimental settings differ in an important aspect. In the studies focusing on a compulsory licensing event, large numbers of patent rights lose the right to exclude simultaneously. Releasing a large set of patent rights into a compulsory licensing regime - and that at a price of zero - must have very different effects than the invalidation of a single patent right. One would expect that interactions between patent rights - as caused by thickets and fences - would not play a major role in the former scenario, but would limit the effect of invalidation of individual patents in the second case where the contextual restrictions from overlapping claims would be largely maintained.

With the present study, we contribute to this emerging stream of literature and investigate the causal effect of a patent's invalidation on follow-on innovation, using a relatively large dataset on opposition to patents granted by the European Patent Office (EPO). The EPO provides a harmonized

[1] The focal patent holder's activities are at the center of attention in several other studies (e.g., Baten et al., 2015; Farre-Mensa et al., 2016; Gaulé, 2016).

application procedure for patent protection in one or more member states of the European Patent Convention (EPC). By now, the EPO grants patents for 38 countries, covering a population of more than 450 million. Hence, patenting decisions by the EPO are economically at least as important to patent-owners and their rivals as corresponding decisions made by the USPTO. In the first nine months after grant, third parties can challenge the validity of a European patent at the EPO by filing an opposition against the granting decision.[2] The opposition procedure represents the last opportunity to centrally invalidate an EPO-granted patent before it is disassembled into national patent rights. With total costs between € 6,000 and € 50,000, the opposition procedure is relatively cheap compared to – sometimes inevitably duplicative – patent revocation proceedings at the national level (Mejer and van Pottelsberghe de la Potterie, 2012). As a result, opposition is a relatively frequent event with a historical opposition rate of about 6% which well exceeds litigation rates in Europe (Cremers et al., 2016) and the U.S. (Lanjouw and Schankerman, 2004; Bessen and Meurer, 2013). Oppositions should also be less prone to settlements given the short time horizon available for negotiations and given the possibility of the EPO pursuing an invalidation even after the parties have withdrawn the case. For these reasons, our data should be less selective than data for decisions at the CAFC as used by Galasso and Schankerman (2015).

In line with previous studies we use post-invalidation forward citations as a proxy for follow-on innovation. To address endogeneity issues concerning the outcome of opposition, we employ a new instrument exploiting variation in the participation of the patent examiner who granted the patent in the opposition division which decides on the grounds for opposition against the patent's validity. Although the rules and regulations of the EPO allow some personnel overlap in the examination and opposition procedure, they do not require the involvement of the examiner. In fact, the examiner participates in about 68% of all opposition proceedings covered by our dataset, with variation over time and technology field. This variation appears to primarily be a function of the non-availability of other examiners with expertise in the particular technology area.

In our baseline model, patent invalidation leads to a highly significant increase of other party and total forward citations, whereas the effect is insignificant when focusing on self citations. While this is in line with previous studies, disentangling the effect leads us to results that stand in stark contrast to the literature. We find that the effect is most pronounced for discrete technology areas such as chemistry, for areas where patent thickets are absent and for patents which are not protected by similar applications by the same applicant (patent fences). Besides, the effect is relevant mostly for small- and medium-sized patent holders facing follow-on inventors of similar size and, to a lesser extent, for large patent holders facing large follow-on patent holders. Taking the established factors for bargaining failure for granted, our results constitute a challenge for considering bargaining failure the dominant explanation for the effect of patent invalidation on follow-on innovation. In our robustness tests, we can show that the impact of relative size, thickets and fencing is visible even when we perform estimation for separate technologies. Hence, cumulative research is impacted through multiple channels in the aftermath of invalidation.

[2]The opposition procedure at the EPO can be compared to the *Post Grant Review* (PGR) at the U.S. Patent and Trademark Office (USPTO). PGR represents an option to challenge validity administratively at the USPTO during the first 9 months after grant without involvement of the judiciary.

Galasso and Schankerman (2015) focus on the causes for bargaining failure and their implications, but give less consideration to factors that determine the effectiveness of patent rights in excluding others. Furthermore, the findings for the highly selective sample of litigation cases at the appellate court for patent disputes cannot be extrapolated to the patent population in a straightforward fashion, because the selection mechanisms may substantially influence the composition with respect to a variety of both observable and unobservable characteristics beyond the commercial value of the patent. In fact, cases reaching a decision by the CAFC represent only a small share of litigation cases – the settlement rate of first instance patent litigation in the U.S. is in excess of 90% (Lanjouw and Schankerman, 2004) –, which in turn represent only a small share of all granted patents. Moreover, it is unclear to which extent first instance rulings impact expectations of market participants.

Our study contributes to the growing literature on the effect of intellectual property rights on cumulative innovation in several ways. First, compared to previous work, our study stands out in scope and scale of the underlying data. Variation in patent rights that can be used to study causal effects is scarce. So far, variation comes primarily from cases where patents are invalidated in court – as illustrated a highly selective and small sample of patents that can be very heterogeneous in age. Looking at post-grant opposition at the EPO, we exploit an institutional device to challenge validity that is more frequently used than patent litigation before ordinary courts. With more than 33,000 observations at the patent level, we capture a sample of patent invalidations that exceeds prior studies by more than an order of magnitude, although we restrict ourselves to a relatively short and recent time frame. Furthermore, with a narrow time window of 9 months right after grant, oppositions occur relatively early in patent life and are far less spread out across a patent's lifetime than in the case of data derived from patent litigation. As each patent can be subject to only one opposition proceeding, we have no observations where the same patent is litigated more than once. We also focus on the first decision on validity for the granted patent, for which, in contrast to cases heard by appeals courts, there is no prior decision that may blur the causal link between invalidation and follow-on innovation. Since the outcome of oppositions can be appealed, we perform related robustness tests, but given the low probability of reversals we expect the first outcome to give follow-on inventors trustworthy indication whether or not to re-engage in research on the subject matter. An additional aspect worth highlighting concerns technological scope. While large parts of the literature are limited to patents in discrete product technologies such as pharmaceuticals, biotechnology or chemicals (Moser and Voena, 2012; Williams, 2013; Sampat and Williams, 2015), our dataset comprises patents across all technologies – an advantage we share with the study by Galasso and Schankerman (2015).

Second, the fine-grained EP citation data used in this study alleviate a key point of criticism concerning the analysis of cumulative innovation proxied by forward citations. On the one hand, potential bias in citations may emerge if applicants can strategically disclose or withhold relevant prior art (cf. Alcacer et al., 2009; Sampat, 2010). In contrast to the US patent system, in Europe citations are made by EPO personnel during the search and examination phase and not by the applicant (Criscuolo and Verspagen, 2008). Furthermore, our dataset includes information on the origin of the citation, which allows us to exclude citations deriving from patents belonging to the applicant himself. On the other hand, it is unclear whether the subject-matter of the citing patent falls within the scope of the cited patent in the first place. The fact that a license may not be required to use the technology independent of the cited patent's invalidation puts the blocking effect of the focal patent right on follow-on innovation in question (Sampat and Williams, 2015). With no obligation of the applicant to disclose prior art, the average number of EPO patent citations is lower, whereas their technological relevance appears to be higher (Breschi and Lissoni, 2004).

Third, we employ an instrumental variable which is new in that it represents the first instrument for patent invalidation in the context of the European patent system, which lacks the randomized administrative processes that allow for a proper identification of fixed effects as used by Galasso and Schankerman (2015). Instead, we focus on the event of administrative personnel overlap in the examination and opposition procedure. Although well established, the literature on oppositions at the EPO focuses primarily on the determinants of opposition (Harhoff and Reitzig, 2004; Schneider, 2011; Harhoff et al., 2016) and opposition as error correction mechanism (Burke and Reitzig, 2007; Graham and Harhoff, 2014). We complement this literature, for the first time providing causal evidence of oppositions on subsequent innovative behavior.

The remainder of this study is structured as follows: Section 2 describes the institutional framework of patent opposition at the EPO. Section 3 provides details on the dataset as well as variables, and entails the descriptive statistics. Section 4 then presents the econometric analysis and a discussion of the results. Section 5 concludes.

2 Empirical Setting

The European Patent Office (EPO) provides a harmonized application procedure for patent protection in one or more member states of the European Patent Convention (EPC). As of now, a patent application granted by the EPO does not lead to a single "European patent." Instead, it is split into a bundle of national patent rights, each entering the patent system of the respective member states. As these rights exist independently of each other, the invalidation of a national patent in one country has no effect on its counterparts in other countries.

However, in the first nine months after grant third parties can challenge the validity of a European patent at the EPO by filing an opposition against the granting decision. As its outcome is binding for all designated states, the centralized opposition procedure represents the only option to invalidate a patent right with coverage of multiple European countries in a single, relatively inexpensive step.

2.1 Examination procedure

The majority of patent applications at the EPO is based on national first filings or international PCT filings (see Harhoff and Wagner (2009) for a detailed description). Only a small share of filings takes the EPO as its priority office. Publication of patent applications occurs at the EPO (as in many other patent authorities) exactly 18 months after the priority date; the publication of the patent document is acccompanied by the EPO Search Report. In the case of PCT filings, which are published by the

World Intellectual Property Organization (WIPO), an International Search Report is generated by an International Search Authority (ISA). Most International Search Reports are actually generated by the EPO. While the original patent application may contain many references to prior art inserted by the applicant, only the prior art listed in the Search Report is relevant for the examination process. The examiner has full control over the selection of prior art references already listed by the applicant for inclusion into the search report, while also generating references via own search efforts.

Within six months after the publication of the search report compiled by the patent office, the patent applicant has to request the examination of the patent application. If the applicant fails to do so, the application is deemed to be withdrawn. With the end of the search procedure, the responsibility for examining the application passes internally from the receiving section to an appointed examining division, which consists of a primary examiner, a secondary examiner, and the chairman. The primary examiner assesses whether the application and the invention meet the requirements of the European Patent Convention and whether the invention is patentable based on the search report. The primary examiner then either grants the patent directly, contingent on the approval by the other two members of the division, or requests a reply from the applicant within a certain time period that addresses the objections raised in the search report. If the objections are successfully overcome by the applicant, the primary examiner sends the version in which he intends to grant the patent, including his own amendments, to the applicant. After the applicant's approval and the completion of formalities, such as the payment of fees, the provision of translations, etc., the grant of the patent is published. The publication date of the EPO B1 document is the official grant date of the patent.

Currently, it takes on average more than four years from the filing of the application to the final decision on the grant of the patent (Harhoff and Wagner, 2009). Since the grant comes along with validation fees and costly translations into national languages, some applicants deliberately delay the examination process. However, in order to make complementary investment decisions or to claim injunctive relief before court, some applicants are interested in fast resolution of the patent examination and file a request for accelerated examination (Harhoff and Stoll, 2015).

2.2 Opposition procedure

The grant decision of the examination division is subject to a post-grant review mechanism, which is initiated by filing a notice of opposition within nine months after the publication of the mention of the patent grant. Oppositions can be filed by any party except the patent holder himself.[3] Receiving the notice of opposition, the primary examiner informs the patent holder and checks whether the grounds for opposition are admissible. Oppositions may be filed on the grounds that the subject-matter is not new or inventive, the invention is not sufficiently disclosed, or the granted patent extends beyond the content of the application as filed.

Consisting of three technically qualified examiners, the appointed opposition division has to decide whether the raised objections compromise the maintenance of the patent. If necessary, the opposition division invites patent holder and opponent to file observations on the other party's com-

[3] In case of multiple independently filed oppositions, all objections are dealt with in one combined proceeding.

munications. During this exchange of communications, the patent holder can amend the description, claims and drawings of the patent. An oral proceeding is summoned if requested by one of the parties, including the opposition division itself. Despite being optional, the oral proceeding before the opposition division is a rarely omitted part of the opposition procedure.

The opposition division usually states its decision verbally at the end of the oral proceeding. The conclusion of the oral proceedings is either the invalidation of the patent in its entirety, the maintenance of the patent as is, or the maintenance of the patent in amended form. A written decision, including the opposition division's reasoning, typically follows one to six months afterwards. If no oral proceeding was requested, the opposition division simply issues its decision in writing. Patent applicant and/or opponent may appeal against the decision of the opposition division. The involvement of the opposition division ends after the opposition phase. Appeal proceedings are heard by judges forming the Boards of Appeal, a separate and independent decision-making body within the EPO.

Withdrawals of oppositions may occur at any stage prior to the decision, but do not necessarily terminate the opposition proceedings. The opposition division has the option to continue the proceeding on its own motion (EPC Rule 84) and make a decision on the patent's validity based on the grounds of opposition previously stated. Since the opposed patent may still end up being invalidated, settlements between opponent and patent holder are relatively rare events. More than 85% of all oppositions conclude in a decision by the opposition division.[4]

2.3 Appointment of examination and opposition division

Technically qualified examiners are assigned to technical art units, so-called directorates. Patent applications are appointed to technical art units according to the application's underlying technology.[5] The examination division regularly consists of the previous search examiner as first member and two examiners appointed by the director as second member and chairman.[6]

The opposition division consists of a first examiner, a minute writer and a chairman. The director appoints the members of the opposition division under consideration of the technical qualifications relevant to the patent. The opposition division may be enlarged to a fourth member with a legal background, if there are complex legal questions to be resolved.

As substantive examiners with the necessary technical qualification, the members of the examination division are natural candidates for the opposition division. Concerning the participation of the grant examiners in the opposition proceeding, Article 19(2) of the European Patent Convention states the following:

> "An Opposition Division shall consist of three technically qualified examiners, at least two of whom shall not have taken part in the proceedings for grant of the patent to

[4]According to our data (see Figure B-1 in the Appendix), the patent holder surrenders the opposed patent in about 5.1% of all oppositions, whereas opponents withdraw their notice without continuation in about 7.7% of all oppositions.

[5]The technical art units are based in Berlin, Den Haag and Munich.

[6]The primary examiner used to be different from the search examiner. This has changed due to the "BEST" ("Bringing Search and Examination Together") initiative, with the goal to have search report as well as examination decision made by the same examiner.

which the opposition relates. An examiner who has taken part in the proceedings for the grant of the European patent may not be the Chairman."

Statements of interviewed EPO officials and our empirical findings show that the primary examiner of the examination division frequently participates in the opposition proceeding of the same patent. Case law has established that the patent holder or opponent cannot object the director's decision regarding the appointment of a particular examiner in the opposition division. The opposition division's decision can in principle be appealed on the ground of suspected lack of impartiality among the division members. However, there are only very few cases where this has occurred; the precedent cases that we are aware of refer to different allegations than the involvement in the previous grant decision.[7]

3 Data and Descriptive Analysis

We use data on opposed patents granted at the EPO between 1993 and 2011 to empirically analyze the causal effect of patent invalidation on follow-on invention. 1993 is taken as the starting point of our data collection as this is the year when the members of the opposition division were – for the first time – explicitly listed in the rulings of the opposition divisions. In order to allow for a sufficiently large time span of 5 years for citations to occur, 2011 marks the last opposition decision year of our data set. This section provides detailed information on our data sources, a discussion of the variables we derive, and a selection of descriptive statistics.

3.1 Data sources

We construct a sample of all patents granted between 1993 and 2011 that became subject to an opposition by drawing on several distinct patent data sources. For each granted patent at the EPO we first observe in the EPO PATSTAT Register whether an opposition was filed within the statutory period of nine months after the grant date.[8] Via the patent application number, we gather all relevant document files concerning the examination and opposition procedure from the online file inspection system of the European Patent Register.[9] We read out documents on the grant decision, the oral proceedings and the opposition decisions in order to extract the names of the examination division and opposition division members, since this information is not available from patent data providers.[10] We elaborate on our read-out and parsing efforts in Appendix D.

[7]For instance in the case G 0005/91 with a decision from May 5, 1992, a patent holder's objection originated from a former employment relationship between examiner and opponent.

[8]Unless otherwise noted below, we use the EPO PATSTAT Statistical Database – 2016 Spring Edition for the selection of patent filings and for extracting citation information.

[9]See https://register.epo.org/regviewer?lng=en. The European Patent Register provides access to digital documents in the public part of a patent file (also known as online file inspection or "file wrapper"). The documents are grouped by procedural stage and include the full written correspondence between the EPO, the applicant, and the opponent. Outgoing communications become available online on the day after the date of dispatch. Incoming communications become available once the filed document has been coded by the EPO.

[10]For PCT patent applications with a filing date from 2011 onwards, the WIPO patent database contains information on the examiner.

We rely on the procedural steps data in the EPO PATSTAT Register data to determine the result and date of the first instance as well as the final decision of the opposition proceeding.[11] Furthermore, the EPO PATSTAT Register provides us with information on the name and address of the opponents. For bibliographic data on the opposed patents, the patent holders, and forward citations, we again use the EPO Worldwide Patent Statistical Database. A few important aspects of the examination process, such as the assigned technical art unit and the examination location, are not covered by any of the above patent databases. We obtain those details from the EPO's administrative database EPASYS (April 2015).

3.2 Dependent variable

A common way to capture a technology's dependence on a past technology is to use citation data. This approach assumes that a cited patent represents the exclusion right that is important when determining the scope of patent protection of the latter patent application. To measure follow-on invention to a focal patent, we therefore look at its number of forward citations in a fixed time window after the opposition outcome. We discuss potential weaknesses of this approach below. As we are most interested in analyzing the effect of the patent's invalidation on follow-on invention, we distinguish citing patents by their filing date relative to the date of invalidation. In order to link the effect to inventive activity and not to application behavior, we use the earliest application date within the DOCDB family of the citing patent. This is also the priority date of subsequent filings, and thus closest to the actual date of invention of the presumed follow-on invention.

We further categorize forward citations by the citing party. Comparing names of the citing applicant with the focal patent holder and the opponent, we distinguish between citations from patents by the patent holder itself ("self citations"), and citations by third parties ("other citations"). In contrast to the US patent system, most citations of European patent applications are generated by the examiners during the search and examination phase and not by the applicant (Criscuolo and Verspagen, 2008). We restrict the citations used to those included in the EPO Search Report or the International Search Report generated by the EPO as International Search Authority. These citations are fully under the control of the examiner. Thus, by design of our dependent variable, we avoid the use of measures impacted by (strategic) citation patterns which may occur when using US citation data (cf. Alcacer et al., 2009; Sampat, 2010).[12]

While we maintain that EPO citations should be more suited to our analysis, it would be comforting to obtain qualitatively similar results when using USPTO data. Therefore, we replicate our

[11] An alternative data source represents the EPO Worldwide Patent Statistical Database. However, this source entails only final opposition outcomes with limited means to reconstruct the result of reversed first instance decisions.

[12] A prominently raised limitation of citation analyses is the lack of distinction between citations where the citing patent is within the scope of protection of the cited patent, and citations where the citing patent is beyond the scope of protection (cf. Sampat and Williams, 2015). In the latter case, a license to use the technology is not required independent of the cited patent's invalidation – blurring the causal effect of patent rights on follow-on invention. With no obligation of the applicant to disclose prior art relevant for the examination at the EPO, the average number of patent citations is lower in comparison with US patent citations, while the technological relevance appears to be higher (Breschi and Lissoni, 2004). To further address the concern of irrelevant citations, we will exploit in future versions of the paper the fact that the EPO provides a detailed categorization of all citations listed in the search report that aims primarily at differentiating prior art by its relevance to the patentability of an invention.

empirical analysis on the basis of USPTO citations and present the results in the appendix. As information on the origin of citations is only available for citations made from 2001 onwards, we include examiner as well as applicant citations published by the USPTO. Moreover, even the distinction available after 2001 may not be fully satisfactory, as the US examiner will add missing references, but not mark applicant-generated references as relevant or not. The European-type search report provides that information.

3.3 Independent variables

The independent variables used in the main empirical analysis capture characteristics of the opposition proceeding, the involved parties, and the focal patent.

Opposition variables

The decision of the opposition division may have three mutually exclusive results for the opposed patent: "valid", "valid in amended form", and "invalid". We operationalize the result in line with Galasso and Schankerman (2015), with the outcomes "invalid' and "valid in amended form" equaling 1, while the outcome "valid" equals 0. The decision of the opposition division can be subject to appeal. In fact, almost half of all decisions in our sample are appealed. However, the reversal rate of the Board of Appeals is very low and skewed; that is, pro-patent holder outcomes are more likely to be overruled in favor of the opponent than vice versa.[13] As appeals considerably delay the final outcomes of opposition proceedings to the effect of substantial truncation in our sample, we focus on the first decision of the opposition division with the expectation that potential bias from disregarding appeals – if at all – understates the effect of invalidation.[14]

Patent holder, opponent and third party variables

Prior literature has found that the risk of bargaining failure between patent holder and potential licensees varies by the vertical position and size of the parties. Furthermore, the country of residence may influence patenting and appropriation strategies. Hence, the selection of patents into opposition, as well as the effect of opposition outcome on follow-on invention, is likely a function of patent holder, opponent and third party characteristics.[15] In line with previous work (Harhoff and Reitzig, 2004), we include the sector (corporate entity or not), the country of residence, and the patent portfolio size of each entity as independent dummy variables. See the explanations below for details on coding.

[13] Which is in line with the established view that arguing against already identified novelty destroying prior art is considerably more challenging than presenting new subject matter.

[14] A further reason to focus on the opposition outcome is the fact that our instrumental variable has a direct effect on first instance outcomes, but merely an indirect effect on appeal outcomes.

[15] Harhoff et al. (2016) argue that non-corporate applicants hold on average patents of lower commercial value and higher novelty, with implications on the selection into opposition.

Patent and procedural variables

We include patent value indicators and technology controls to reduce asymptotic variances and to mitigate bias. To preempt endogeneity issues, we focus on patent value indicators that are set at a very early stage of the patent application and are thus independent of the examination and opposition proceeding. These are a dummy variable for international patent applications (PCT) and count variables for DOCDB patent family size, IPC subclasses, claims, applicants, inventors and patent as well as non-patent backward citations. We include pre-opposition self and other citations restricted to the first three years after filing as further proxy variables for patent value.

We assign each patent to a technology area by mapping the IPC classes in line with the concordance table developed by the Fraunhofer ISI and the Observatoire des Sciences et des Technologies in cooperation with the French patent office (cf. Schmoch, 2008). The IPC codes are clustered into 34 technology areas, each belonging to one of five main technological areas: (a) electrical engineering, (b) instruments, (c) chemistry, (d) mechanical engineering, and (e) other fields.

In our empirical analysis, we further aim to account for contextual factors of the focal patent. We employ a time-variant variable that measures the density of patent thickets in the focal patent's technology area (cf. Von Graevenitz et al., 2011). The focal patent may also be part of a "patent fence" consisting of several similar patents held by the patent holder. With the help of a novel approach that calculates a measure of similarity using a sophisticated semantic comparison of patents' full texts (abstract, description, claims and title) (cf. Harhoff, 2014), we count the number of patents that are highly similar to the focal patent and belong to the patent holder's portfolio.

In order to show randomness of our instrumental variable, we test correlations with a set of further variables specific to the patent examination process. These variables include the duration of examination, the language of the proceeding, and the granted request for accelerated examination.

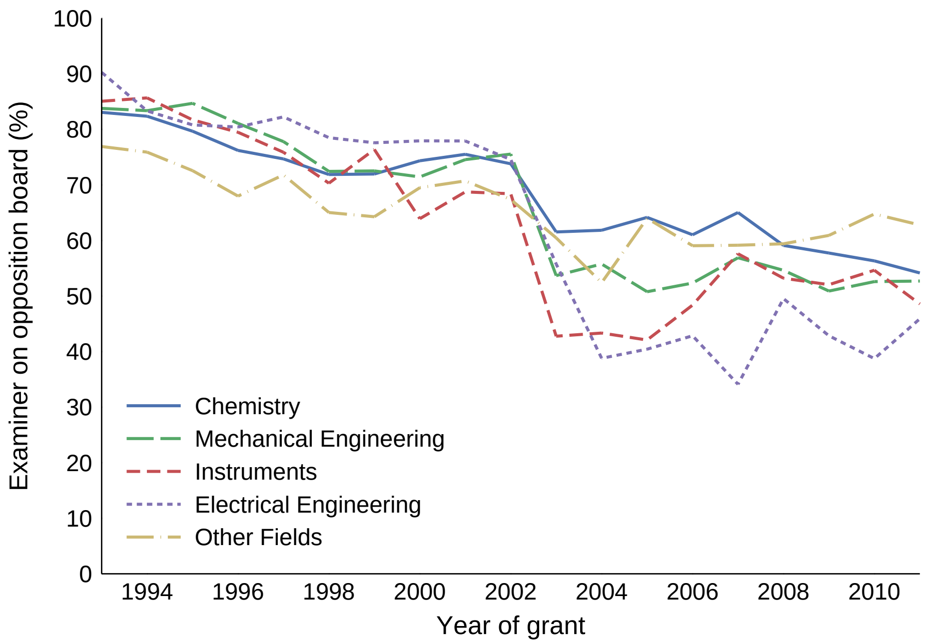
3.4 Instrumental variable

The opposition division consists of three technically qualified substantive examiners, of which at least two must not have taken part in the examination of the opposed patent. Opposition cases are decided by a vote of all three persons on the opposition board.[16] It seems like a natural assumption that the examiner who granted the patent is generally more inclined to be in favor of the patent holder than of the opponent, who dissents with the examiner's prior decision. Given that this pro-patent holder effect exists, two requirements must be fulfilled so that we can exploit the participation of the examiner in the opposition proceeding as an instrumental variable. First, we need continuous variation in examiner participation across time and within cohorts. Second, we must be able to exclude any endogeneity in the determination whether the patent examiner participates in the opposition division or not.

We conducted interviews with EPO officials to explore the process by which opposition divisions are formed. These discussions revealed that the reasons for the participation of the examiner are primarily found in the non-availability of other examiners with expertise in the particular technology

[16]Voting follows a simple majority. In case of parity, the vote of the chairman is decisive.

Figure 1: Annual rate of examiner participation in opposition proceeding



Notes: This graph shows the annual rate of examiner participation in opposition proceedings by technology main area. The sample includes oppositions with first outcome after 2011.

area. If the number of substantive examiners relative to oppositions is large, the granting examiner is less likely to take part in the opposition proceeding as the third member of the division. The supply of substantive examiners depends *inter alia* on the labor market – staff shortage induces the granting examiner to become indispensable for the opposition proceeding. Figure 1 shows that the average participation rate is well above 60% before 2003, but then declines to an average rate of about 55% with increasing variations between technology main areas. This drop is caused by a sharp increase in the number of substantive examiners eligible to participate in opposition proceedings in the course of the "BEST" initiative.[17] We conclude that the event "examiner participation in opposition proceeding" is exogenous and frequent, yet by far not always the case – with continuous variation within cohorts and technology areas.

To further argue against potential endogeneity, we discuss the instrument's randomness and its virtue as exclusion restriction. In Table R-1, we show that common patent value indicators as well as characteristics of the patent holder and opponent do not show any significant effect on the likelihood of the examiner's participation in the opposition proceeding. This supports the view of EPO officials and patent attorneys that the participation or absence of the examiner is independent of the opposed patent and beyond the influence of the patent holder or the opponent. However, one legitimate concern is that the duration of examination may affect the likelihood of examiner

[17]The "BEST" ("Bringing Search and Examination Together") initiative had the goal to have the search report and examination decision made by the same examiner. For this purpose, search examiners were - on a large scale - trained and promoted to substantive examiners.

participation as well as follow-on citations. An applicant with a considerable pipeline of follow-on inventions may be interested in having the patent granted as quickly as possible. As prior empirical analyses (e.g., Harhoff and Wagner, 2009) have shown, the duration of examination is not perfectly exogenous, i.e., the applicant has considerable influence in speeding up or delaying the examination process. This may present a problem to the instrumental variable if the duration of the proceeding affects the examiner's availability to participate in the opposition proceeding. For instance, the granting examiner may become unavailable due to retirement, promotion or transfer in a different technical art unit. However, our sample does not show any effect of examination length on the likelihood of the examiner's participation in the opposition proceeding. Accelerated examination constitutes an additional issue. Even when controlling for length of examination, the request of accelerated examination positively affects the participation dummy. We assume this is due to the fact that the accelerated examination request releases the examiner from further duties and provides him with a free schedule to participate in the opposition proceeding. To underline the robustness of our instrument, we remove cases with accelerated examination (about 11% of the sample) in a robustness test, yet we find no significant changes throughout our results.

A random instrument could still violate the exclusion restriction if the outcome is affected through different ways than just the first stage.[18] This would be the case if the applicant foresaw whether the examiner is part of the opposition proceeding earlier than at the decision on the patent's validity, providing him with enough time to adjust his behavior accordingly. However, this seems very unlikely. While the composition of the division is set at the beginning of the opposition proceeding, all correspondence between the applicant/opponent and the EPO is channeled through the formalities officer. Only at the time of the oral proceeding, which usually ends in a decision on the case, the opposition division members become known to the parties.[19] The applicant may also be able to foresee whether the examiner is part of the opposition proceeding if examiner-specific participation rates are concentrated at zero and one. As can be seen from Figure A-1, this concern is unfounded.

3.5 Descriptive statistics

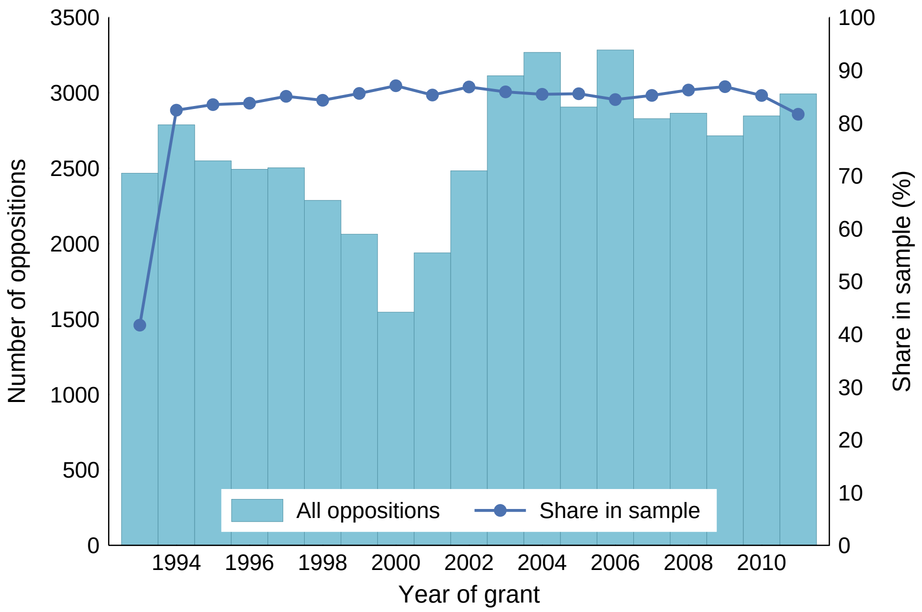
We count 49,938 patents granted between 1993 and 2011 with opposition at the EPO. Since the composition of the examination and opposition board is essential to construct our instrumental variable, our sample is limited to those patents where we are able to gather the names of the examiners involved in the grant and opposition decision. For several reasons outlined in Table B-1, we are forced to exclude about 17% of patents, leading to a sample size of 41,358 patents. We assume that this selection has little relevance for our subsequent analysis. The fact that the excluded patents are equally distributed over time (cf. Figure 2) supports this view.

A second sample restriction comes into play when constructing the follow-on citation variables. To mitigate truncation effects for more recently invalidated patents, we exclude patents with a first

[18]This concern follows a similar argument raised and discussed in Farre-Mensa et al. (2016).

[19]In those cases where applicant and opponent waive the oral proceeding, the parties learn about the identity of the opposition division members only through the published decision.

Figure 2: Annual number of opposed patents and sample rate



Notes: This graph includes all opposition proceedings (at the patent level) with grant date between 1993 and 2011. The low sample rate in the first year is due to the fact that the EPO introduced the grant document type that contains examiner names only in mid of 1993. The used sample includes oppositions with first outcome after 2011.

instance opposition decision after 2011. This reduces our main sample of analysis to 33,075 observations at the patent level.

Opposition proceedings usually result in one of three distinct outcomes for the opposed patent: valid, amended, invalid. In line with prior analyses of oppositions at the EPO, we find fairly equal shares across the three outcomes. Yet, time trends appear to exist in our sample (see Figure 3a): invalidations have seen a considerable increase over the last 20 years, whereas fewer and fewer patents survive opposition perfectly unscathed.

The opposition rates across technology fields differ substantially (Harhoff et al., 2016). These rates hardly correspond to recent technology-specific estimates of weak patents granted by the EPO (de Rassenfosse et al., 2016). Since patent invalidation is ex ante uncertain and its benefits often difficult to internalize, potential opponents may be reluctant to invest in a post-grant validity challenge. This public good problem weakening the error correction mechanism is most prevalent in complex technology areas with a low concentration of patent ownership and a high density of patent thickets (Harhoff et al., 2016). In line with this, Figure 3b shows that with negligible variation over time the predominant share of oppositions in our samples are filed against patents in the technology areas "Chemistry" and "Mechanical Engineering."

We present the summary statistics of patent and procedural characteristics in Table 2. Among the patent characteristics, we distinguish between self/other forward citations within three years after filing and self/other forward citations within five years after the opposition decision. While the latter represent our dependent variables of interest, we consider the former, which are indepen-

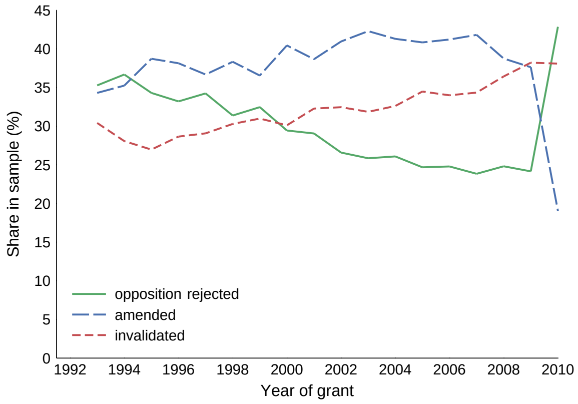
Table 2: Patent and procedural characteristics

Variable	Mean	SD	min	max
Patent characteristics				
Self forward citations (3 years after filing)	0.38	0.97	0	20
Other forward citations (3 years after filing)	0.85	1.83	0	83
Self forward citations (5 years after decision)	0.13	0.49	0	10
Other forward citations (5 years after decision)	0.77	1.42	0	32
Age of patent (yr)	8.84	2.47	3	26
DOCDB family size	10.75	10.56	1	263
No of patent holders	1.07	0.32	1	13
No of inventors	2.61	1.76	1	21
No of claims	13.12	10.05	0	329
No of IPC subclasses	2.74	2.45	1	56
No of patent backward references	6.31	4.82	0	128
No of non-patent backward references	1.15	3.39	0	110
PCT application (d)	0.44	0.50	0	1
Year of application filing	1996.22	4.71	1981	2008
Year of grant decision	2001.01	4.62	1993	2010
Patent technology main area				
Electrical Engineering (d)	0.10	0.31	0	1
Chemistry (d)	0.39	0.49	0	1
Instruments (d)	0.13	0.33	0	1
Mechanical Engineering (d)	0.31	0.46	0	1
Other Fields (d)	0.06	0.24	0	1
Examination proceeding				
Duration filing to examination (yr)	1.72	1.22	0	18
Duration of examination (yr)	3.98	1.80	0	16
Accelerated examination (d)	0.11	0.31	0	1
Opposition proceeding				
Examiner participation (d)	0.68	0.47	0	1
Outcome: valid (d)	0.29	0.45	0	1
Outcome: invalid (d)	0.71	0.45	0	1
Appeal (d)	0.46	0.50	0	1
Outcome reversal (d)	0.07	0.26	0	1
Observations		33,075		

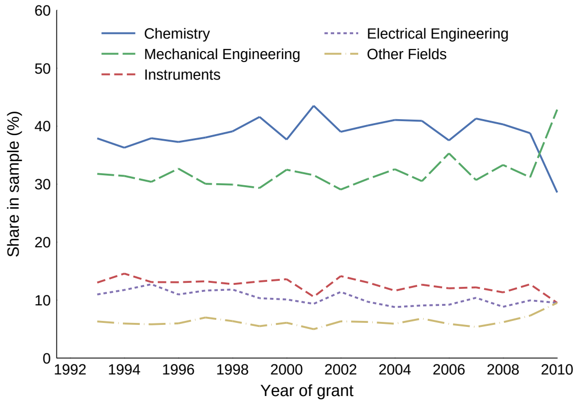
Notes: This table presents characteristics of the patent and examination as well as opposition proceeding at the level of opposition cases.

Figure 3: Time trends in oppositions

(a) Opposition outcomes over time



(b) Oppositions over time by technology area



Notes: Both graphs include all opposition proceedings (at the patent level) which are part of our main sample of analysis. Grant year 2010 includes only 21 opposition proceedings.

dent of the subsequent opposition proceeding, as control variables. As further exogenous patent value indicators we draw on the DOCDB family size and counts of applicants, inventors, claims, IPC subclasses and backward references. With application filing years between 1981 and 2008, the average patent has spent about 4 years in examination and is close to 9 years old when the opposition division decides on its validity. That is, opposition outcomes occur relatively early in patent life and are far less spread across a patent's lifespan than the outcome in patent litigation (see Figure A-2).

Concerning the opposition proceeding, the overall participation rate of an examiner on the opposition board is about 68%, with considerable variation over time and technology areas as already elaborated in Section 3.4. Almost half of all opposition decisions are appealed before the EPO's board of appeals. However, the reversal rate (computed as the share of all cases where the appeal outcome is different from the opposition outcome) stands at mere 7%. Moreover, appeals initiated by the patent holder, that is, where the decision in first instance was rather in favor of the opponent, are even less common to induce a reversal than vice versa (see Figure B-2).

Oppositions are mostly filed by corporations and directed at corporate patent holders. Table 4 shows that 94% of patent holders and 98% of opponents are companies with practically no involvement of parties from the academic or non-profit sector.[20] The opposition proceeding may consolidate multiple notices of opposition that were filed during the nine months window after grant. On average, about 1.3 parties represent the validity challenging side. We account for cases with more than one opponent in our subsequent empirical analysis.

The distribution of the patent holders' countries of residence is very similar to the overall distribution among all granted patents. Naturally, as the grant of EP patents affects primarily companies active in EPC countries, the share of opponents with residence in one these countries is considerably higher in comparison. To capture effects varying with the patent holder's size, we classify the patent

[20] EPO caselaw has rendered the use of a "straw man" representing the real party interested in the opposition eligible. In those rare occasions, our data reference a law firm or a single patent attorney as opponent.

Table 3: Characteristics of patent holder and opponent

	Patent holder				Opponent			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Number of parties	1.07	0.31	1	11	1.28	0.76	1	19
Sector								
Company (d)	0.94	0.25	0	1	0.98	0.15	0	1
Country of residence								
EPC (excl. GB) (d)	0.58	0.49	0	1	0.83	0.37	0	1
GB (d)	0.04	0.20	0	1	0.04	0.20	0	1
US (d)	0.23	0.42	0	1	0.10	0.29	0	1
JP (d)	0.12	0.32	0	1	0.02	0.14	0	1
Other (d)	0.03	0.16	0	1	0.01	0.10	0	1
Size								
Large (d)	0.38	0.49	0	1		–		
Medium (d)	0.28	0.45	0	1		–		
Small (d)	0.34	0.47	0	1		–		
Observations		33,075				33,075		

Notes: This table presents characteristics of the patent holder(s) and the opponent(s) at the level of opposition cases. In case of multiple patent holders / opponents, we give preference according to the ordering of sector, country of residence, and size. Size categories are proxied by the number of patents (incl. applications) filed during the last five years prior to the opposition decision (large: 200 and more patents, medium: 20 and more patents, small: fewer than 20 patents).

holder as either small, medium or large according to his patent portfolio. This measure seems less appropriate to proxy the opponent's size. For instance, oppositions against pharmaceutical patents are frequently filed by generic drug companies that hold few if any patents. As we are more interested in the size of firms with innovative follow-on activities, we disregard this aspect of the opponent.

We capture follow-on inventions by the number of forward citations the focal patent receives within the first five years after the opposition outcome. In line with prior empirical analyses, we distinguish between "self citations", where the citing applicant and the focal patent holder are the same entity, and "other citations", where the citing applicant and the focal patent holder are different entities. We focus on forward citations linking two patent families on the basis of patent applications published by the EPO or WIPO. The EPO/WIPO citation data are unusually rich, letting us distinguish between citations originating from the citing applicant or from a patent office examiner and providing information on the technological relevance of the cited patent. As can be seen from Table 4, citation characteristics differ between self citation and other citations. If the citing applicant is also the holder of the cited patent, the citation is more likely to originate from himself than from an examiner.[21]

[21]This suggests that citation data based on applicant information only may be prone to substantial bias.

Table 4: Characteristics of EP/WO forward citations by relationship to cited patent

	Self citations				Other citations			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Publication authority								
EPO	0.45	0.50	0	1	0.45	0.50	0	1
WIPO	0.55	0.50	0	1	0.55	0.50	0	1
Citation characteristics								
Citation lag (yr)	10.14	2.69	1	22	10.43	2.88	1	25
DOCDB family size	6.77	5.45	1	85	5.93	5.55	1	254
Sector (citing applicant)								
Company (d)	0.98	0.15	0	1	0.92	0.28	0	1
Country of residence (citing applicant)								
EPC (excl. GB) (d)	0.63	0.48	0	1	0.57	0.50	0	1
GB (d)	0.02	0.15	0	1	0.04	0.19	0	1
US (d)	0.24	0.42	0	1	0.24	0.43	0	1
JP (d)	0.10	0.30	0	1	0.10	0.30	0	1
Other (d)	0.01	0.11	0	1	0.06	0.24	0	1
Size (citing applicant)								
Large (d)	0.52	0.50	0	1	0.32	0.47	0	1
Medium (d)	0.28	0.45	0	1	0.25	0.43	0	1
Small (d)	0.20	0.40	0	1	0.43	0.50	0	1
Observations	4,139				25,413			

Notes: This table includes examiner forward citations for patents subject to opposition proceedings in our main sample of analysis. The unit of observation is the citation. We only consider citation links established in search reports issued by the EPO. In case of multiple citations coming from the same patent family, we keep the earliest citation. In case of multiple citing applicants, we give preference according to the ordering of sector, country of residence, and size. Size categories are proxied by the number of patents (incl. applications) filed during the last five years prior to the opposition decision (large: 200 and more patents, medium: 20 and more patents, small: fewer than 20 patents).

4 Empirical Analysis

4.1 Baseline specification and identification strategy

Our data on oppositions is a cross section where the unit of observation is the opposition proceeding involving the unique patent p. Our main empirical specification is

$$\log(\textit{Forward citations}_p) = \beta_1 \textit{ Invalidated}_p + \beta_2 \textit{ Patent}_p + \beta_3 \textit{ Patent holder}_p + \\ + \beta_4 \textit{ Opponent}_p + \beta_5 \textit{ Age}_p + \beta_6 \textit{ Year}_p + \beta_7 \textit{ Tech}_p + \epsilon_p .$$

The coefficient β_1 captures the effect of invalidation on subsequent forward citations the opposed patent receives. If patent rights have a positive or no impact on follow-on innovation, we would expect $\beta_1 \leq 0$. Vice versa, a finding of $\beta_1 > 0$ would suggest that patents block follow-on innovation.

Our dependent variable captures the number of forward citations within the first five years after the opposition outcome. We distinguish between forward citations in total, those from patents held by the focal patent holder himself ("self citations") and those from patents held by others ("other citations"). To control for heterogeneity in the value that the patent has for the patent holder and follow-on inventors, we include patent value indicators, such as the number of claims and the number of self citations and other citations received within the first three years as covariates in the regression. We also include age, grant as well as decision year, and technology field dummies to control for additional heterogeneity that may correlate with the court decision and subsequent citations.

As previous studies have amply illustrated, our main empirical challenge is the endogeneity of the opposition division's decision to invalidate the patent. More valuable inventions may lead to more forward citations, but may also induce the patent holder to heavily defend the patent. This negative correlation, biasing the OLS estimate of β_1, renders this specification inappropriate to estimate causal effects. To address this endogeneity, we need an instrument that affects the likelihood of patent invalidation, but does not belong directly in the citations equation, hence creating exogenous variation in patent invalidation.

We construct our instrument around the participation of the primary examiner in the opposition proceeding – an approach new to the literature, which has focused on the use of decision maker fixed effects (Sampat and Williams, 2015; Galasso and Schankerman, 2015). Following the basic intuition that the primary examiner is more likely to come to the same conclusion concerning the validity of the patent as in the examination proceeding than an arbitrary examiner, namely a confirmation of the patentability of the subject matter, we expect his participation to negatively affect the probability of invalidation. To verify this, we use probit estimation models to regress the binary opposition outcome variable "Invalidated" on the "Examiner participation" dummy and all other exogenous variables $\boldsymbol{x}$,

$$\text{Prob}(\textit{Invalidated}_p) = \Phi(\gamma_1 \textit{Examiner participation}_p + \boldsymbol{\gamma} \boldsymbol{x}_p) \\ \rightarrow \textit{Predicted probability of invalidation}_p . \qquad (4.1)$$

We find strong evidence that examiner participation indeed has an effect on the opposition outcome (p-value < 0.001). More importantly, we use the probit regression to obtain a fitted probability (propensity score) of invalidation for each observation, which we use as our instrument throughout the paper. We then apply standard Two-Stage Least Squares (2SLS) regression analysis, instrumenting the dummy of the opposition outcome with the predicted probability, [22]

$$\begin{aligned} \textit{Invalidated}_p &= \alpha_1 \; \textit{Predicted probability}_p + \boldsymbol{\alpha x}_p \; + u_p \\ \log(\textit{Forward citations}_p) &= \beta_1 \; \widehat{\textit{Invalidated}}_p \; + \boldsymbol{\beta x}_p \; + \; \epsilon_p \,. \end{aligned} \tag{4.2}$$

In Table 5, columns (1) and (2), we report detailed results of the probit regression models of the invalidation dummy on the examiner participation dummy. The estimated effect in column (1) indicates that examiner participation is associated with a decrease of about 6.6 percentage points in the likelihood of invalidation. The results are similar when we add the full set of control variables (cf. column (2)) – examiner participation is associated with a highly significant decrease of about 4 percentage points in the probability of invalidation. We also find that patents with a larger number of claims are more likely to be invalidated, whereas variables concerning the time until grant have no significant effect.

Column (3) explores the interrelation of the observable control variables with examiner participation to provide some additional perspective concerning the exogeneity assumption. Variables with the potential to raise concerns have statistically insignificant coefficients close to zero. For a more detailed overview, especially concerning patent characteristics, see Table R-1 in the appendix.

Note that weak identification is never an issue in the 2SLS regressions in the following, with heteroskedasticity-robust first-stage F-statistics ranging from >70 for one of the considered subsamples to 700 for the full sample.

4.2 Results and discussion

In Table 6 we examine how patent invalidation or partial invalidation in an opposition proceeding affects the number of subsequent EP/WO forward examiner citations. Column (1) shows the baseline OLS regression of the logarithmized number of forward citations of parties other than the focal patent holder within five years after the opposition decision on the invalidity dummy and an extensive set of control variables. There is a weakly significant negative correlation between patent invalidation and future citations. In contrast, turning to the 2SLS instrumental variables regression in column (2), we find a highly significant positive coefficient. The obvious discrepancy from the OLS estimate is in line with the expected endogeneity of invalidation, a suspicion confirmed on the 5% level by a test of endogeneity. The estimated coefficient implies that patent invalidation causes a significant increase in citations by other parties in the five years following the opposition outcome. Note that the instrument explains a sizable part of the variation in patent invalidation, which is underlined by the first stage heteroskedasticity-robust F-statistic of 700 – a value that easily

[22]The resulting estimator is asymptotically efficient in the class of estimators where the instrumental variables are functions of all exogenous variables (Wooldridge, 2010, p. 939, Procedure 21.1).

Table 5: Examiner participation and opposition outcome (EP/WO citations)

	(1)	(2)	(3)
Estimation method	Probit	Probit	Probit
Dependent variable	Invalidated (d)	Invalidated (d)	Examiner participation (d)
Exam. participation (d)	-0.066^{***}	-0.040^{***}	
	(0.005)	(0.005)	
log(No of claims)		0.031^{***}	−0.005
		(0.004)	(0.004)
log(CitEPExaPre3Other)		0.000	0.001
		(0.005)	(0.005)
log(CitEPExaPre3Self)		−0.010	0.006
		(0.007)	(0.007)
Duration of examination (yr)		−0.007	0.004
		(0.006)	(0.007)
Duration of wait (yr)		0.009	0.007
		(0.007)	(0.007)
Year effects	No	Yes^{***}	Yes^{***}
Age effects	No	Yes^{*}	Yes^{*}
Technology effects	No	Yes^{***}	Yes^{***}
Patent characteristics	No	Yes^{***}	Yes^{**}
Patent holder characteristics	No	Yes^{*}	$\text{Yes}^{\dagger}$
Opponent characteristics	No	Yes^{***}	Yes
Model degrees of freedom	1	113	112
χ^2-statistic	154.3	2,409.5	2,777.4
Pseudo-R^2	0.004	0.083	0.073
Observations	33,075	33,075	33,075

Marginal effects; Robust standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The probit regressions in columns (1) and (2) illuminate the relevance of the "Examiner participation" dummy for the outcome of the opposition proceeding. The invalidation predictions of the probit regression in column (2)—or equivalent predictions for subsamples and other citation measures—are used as the instrument in the 2SLS instrumental variables regressions throughout the paper. Column (3) shows the probit regression of the "Examiner participation" dummy on the other exogenous variables. One is added to all citation variables before taking the logarithm to include patents with no forward citations. A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

exceeds the Stock and Yogo (2005) (i.i.d. error) critical values for weak identification tests. Column (3) presents the results of the same baseline specification, however, with the dependent variable restricted to citations from patents held by the focal patent holder himself. We find no significant effect of invalidation on the focal patent holder's follow-on inventive activity. Column (4) presents the results of the baseline specification on the total number of citations. While these positive average effects over the whole sample for 'other' as well as total citations are in line with the findings of Galasso and Schankerman (2015), the results investigating the origin of the effect stand in stark contrast.

The following four tables disentangle the average effect on other citations by technology area, complexity of the technology and size of both the focal and the citing patent holder.

First, Table 7 lists the estimation results on subsamples defined by technology main area. While the coefficients for "Electrical Engineering", "Instruments" and "Chemistry" are all positive, the latter is the only one with statistical significance. It appears the effect of invalidation on citations by others is most coherent in "Chemistry" – an area which is commonly associated with discrete technologies, while "Electrical Engineering" and "Instruments" encompass predominantly complex technologies.

Second, given that the fairly large standard errors for "Electrical Engineering" and "Instruments" hint at potential heterogeneity in the effect of invalidation on citations, in Table 8 we split the sample based on the nature of the underlying technology and based on the size of the focal patent holder. In column (1) we restrict our sample to complex technology areas, resulting in no significant effect of invalidation on forward citations by others. In contrast, the subsample of patents in "discrete" technologies in column (2) shows a highly significant positive invalidation effect. These results go hand in hand with the common perception of the difference between complex and discrete technologies. While the protection of an invention in discrete technologies is concentrated in a single patent, resulting in profound consequences for the IP landscape in the case of an invalidation, inventions in complex technologies are typically spread across two or more patents, rendering the implications of an invalidation less severe and more heterogeneous. We further explore this channel in Table 10. Column (3) and column (4) concern the size of the focal patent holder. We find a much stronger and highly significant effect of invalidation on citations by others if the focal patent holder is small or medium-sized. This result is difficult to align with the findings of Galasso and Schankerman (2015), who find that the effect of invalidation on forward citations is larger if the holder of the invalidated patent is large. According to our results, bargaining failure, which presumably blocks follow-on innovation as long as the patent remains in force, is considerably less likely in negotiations with large patent holders. We investigate this channel in more detail in Table 9.

We include both aspects, complexity and size, in the subsample definitions used in column (5) and column (6). While there is some positive effect for the subsample that includes all complex patents and/or large focal patent holders, we find the coefficient estimates almost twice as large in magnitude and highly significant for the subsample based on patents in discrete technology areas which are held by non-large patent holders.

Table 6: Impact of invalidation as opposition outcome on EP/WO citations

	(1)	(2)	(3)	(4)
Estimation method	OLS	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Self	Total
Invalidated (d)	-0.014^{*}	0.223^{***}	0.044	0.245^{***}
	(0.007)	(0.063)	(0.027)	(0.066)
log(No of claims)	0.061^{***}	0.054^{***}	0.014^{***}	0.062^{***}
	(0.005)	(0.005)	(0.002)	(0.005)
log(CitEPExaPre3Other)	0.129^{***}	0.129^{***}	$0.005^{\dagger}$	0.129^{***}
	(0.006)	(0.006)	(0.003)	(0.006)
log(CitEPExaPre3Self)	0.019^{*}	0.021^{*}	0.047^{***}	0.050^{***}
	(0.008)	(0.008)	(0.005)	(0.009)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes**	Yes*	Yes	Yes**
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes***	Yes**	Yes	Yes**
Patent holder characteristics	Yes**	Yes***	Yes***	Yes
Opponent characteristics	Yes***	Yes***	Yes	Yes***
Underidentification test		344.6	344.6	344.6
Weak identification test		700.0	700.0	700.0
Observations	33,075	33,075	33,075	33,075

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Columns (1) and (2) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on EP/WO examiner citations to patents held by other parties than the focal patent owner, as measured by EP/WO examiner forward citations in a 5-year window following the decision of the opposition proceeding. Columns (2)–(4) show 2SLS regressions for the impact of invalidation on the number of follow-on patents held by other parties than the focal patent owner, on the number of follow-on patents held by the focal patent owner himself and on the total number of follow-on patents, respectively. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table 7: Impact of invalidation as opposition outcome on EP/WO citations – technology main areas

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExaPost5…)	Other	Other	Other	Other
Technology area	Electr Eng	Instruments	Chemistry	Mech Eng
Invalidated (d)	0.216	0.394	0.245**	−0.193
	(0.176)	(0.252)	(0.086)	(0.151)
log(No of claims)	0.060***	0.062***	0.043***	0.075***
	(0.015)	(0.015)	(0.008)	(0.009)
log(CitEPExaPre3Other)	0.140***	0.169***	0.099***	0.143***
	(0.017)	(0.017)	(0.009)	(0.012)
log(CitEPExaPre3Self)	0.089**	0.025	0.006	0.032*
	(0.031)	(0.024)	(0.011)	(0.016)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes	Yes***	Yes	Yes*
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes***
Patent holder characteristics	Yes***	Yes	Yes**	Yes
Opponent characteristics	Yes	Yes*	Yes	Yes**
Underidentification test	51.9	54.9	164.6	74.7
Weak identification test	90.9	75.8	324.9	114.5
Observations	3,432	4,220	13,011	10,384

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Columns (1)–(4) show the impact of invalidation on EP/WO examiner forward citations to patents held by parties other than the focal patent holder for the technology main area subsamples Electrical Engineering, Instruments, Chemistry and Mechanical Engineering, respectively. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table 8: Impact of invalidation as opposition outcome on EP/WO citations – technology and size

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	IV	IV	IV	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Other	Other	Other	Other
Subsample	Complex	Discrete	Large	Non-large	Complex or large	Discrete, non-large
Invalidated (d)	0.137	0.256^{***}	0.054	0.298^{***}	0.123	0.323^{***}
	(0.112)	(0.074)	(0.116)	(0.073)	(0.090)	(0.086)
log(No of claims)	0.066^{***}	0.044^{***}	0.046^{***}	0.058^{***}	0.061^{***}	0.043^{***}
	(0.007)	(0.007)	(0.009)	(0.006)	(0.006)	(0.009)
log(CitEPExaPre3Other)	0.154^{***}	0.107^{***}	0.113^{***}	0.136^{***}	0.134^{***}	0.115^{***}
	(0.009)	(0.008)	(0.010)	(0.008)	(0.008)	(0.010)
log(CitEPExaPre3Self)	0.032^{*}	$0.017^{\dagger}$	$0.020^{\dagger}$	0.032^{**}	$0.019^{\dagger}$	0.030^{*}
	(0.014)	(0.010)	(0.011)	(0.011)	(0.010)	(0.014)
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes	Yes	Yes**	Yes†	Yes†
Technology effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes*	Yes*	Yes	Yes**	Yes	Yes**
Patent holder characteristics	Yes**	Yes*	Yes	Yes	Yes†	Yes
Opponent characteristics	Yes*	Yes*	Yes*	Yes*	Yes**	Yes
Underidentification test	140.6	199.9	102.3	263.0	180.7	167.8
Weak identification test	276.6	426.9	225.8	477.9	397.9	318.1
Observations	14,946	18,129	11,014	22,061	20,911	12,164

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the impact of invalidation on EP/WO examiner forward citations to patents held by parties other than the focal patent holder for different sample splits. Columns (1) and (2) compare the effect in complex technologies to that in discrete technologies, columns (3) and (4) compare the effect for patents held by large patent holders to that for patents held by non-large patent holders and columns (5) and (6) compare the effect for patents which are in complex technologies or held by a large patent holder to that for patents which are in discrete technologies and held by a non-large patent holder. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Third, motivated by our findings on patent holder size, in Table 9 we further explore the heterogeneity of the invalidation effect with respect to the differences in size between the owner of the citing (dependent variable) and the owner of the focal patent (subsample). Columns (1) and (2) show the effects of invalidation of a large-holder patent on forward inventive activity by large and non-large patent holders, respectively. Both coefficients are insignificant on a 10% level. In contrast, for columns (3) and (4), which display the corresponding effects for the invalidation of patent held by a non-large owner, we find highly significant coefficients. However, the effects on non-large other parties appears stronger than the one for large other parties. These results imply an ordering with regard to bargaining failure in the presence of a patent right. Frictions are most pronounced for non-large focal patent holders and non-large follow-on innovators (4), significant for non-large original applicants and large subsequent innovators (3), insignificant for large focal patent holders and large follow-on innovators (1) and close to zero for large original applicants and non-large subsequent applicants (2). This is consistent with intuition: While small firms struggle to efficiently negotiate a path for follow-on innovation building on a second small firm's patented invention, they are free to operate after an invalidation (4). However, small firms are unable to profit from the invalidation of a patent held by a large company which is able to retain protection of its invention by further patents or by other means (2). Although this logic may apply to large focal patent holders facing small follow-on innovators, it seems to be inappropriate for those faced with a large competitor, where the invalidation does not have a perceptible effect (1). Finally, although non-large original applicants enjoy some protection against large follow-on innovators as long as the patent right is in place (3), it is by no means as effective as against small subsequent innovators. It seems that large follow-on innovators can to a considerably greater extent rely on being capable of building on a small applicant's patented invention than that of a large original inventor.

Fourth, to further inquire into the findings for complex technologies in Table 8, we discuss the invalidation effect in the presence of patent thickets and patent fences in Table 10. In columns (1) and (2) the sample is split into technology areas with and without patent thickets, respectively. Consistent with intuition we do not find a significant effect of invalidation in areas with thickets, but a positive and significant effect for those without. Similarly, there is no significant effect for patents protected by a fence, i.e., protected by the presence of one or more similar patents filed by the focal patent holder before the opposition proceeding of the focal patent, while there is a strong and highly significant effect in the case of the absence of a protecting fence.

Fifth, in order to examine potential differences in the invalidation effect with respect to patent age and value, Table 11 shows the results for sample splits at the age median of 8 years and the DOCDB family size median of 8. The effect seems to be primarily driven by younger and more valuable patents.

Table 9: Impact of invalidation as opposition outcome on EP/WO citations – by sizes of focal and citing patent holders

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExa...Post5Other)	Large	Non-large	Large	Non-large
Patent holder subsample	Large	Large	Non-large	Non-large
Invalidated (d)	0.089	−0.031	0.119**	0.217***
	(0.070)	(0.106)	(0.042)	(0.065)
log(No of claims)	0.016**	0.036***	0.013***	0.051***
	(0.006)	(0.007)	(0.004)	(0.006)
log(CitEPExaPre3Other)	0.060***	0.073***	0.060***	0.101***
	(0.008)	(0.008)	(0.005)	(0.007)
log(CitEPExaPre3Self)	0.023**	0.002	0.023**	0.018†
	(0.009)	(0.009)	(0.007)	(0.010)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes	Yes	Yes*	Yes
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes**
Patent holder characteristics	Yes	Yes	Yes**	Yes***
Opponent characteristics	Yes	Yes*	Yes†	Yes**
Underidentification test	102.3	102.3	263.0	263.0
Weak identification test	225.8	225.8	477.9	477.9
Observations	11,014	11,014	22,061	22,061

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the impact of invalidation on EP/WO examiner citations with respect to the differences in size between the holder of the citing patent (dependent variable) and the holder of the focal patent (subsample). Columns (1) and (2) show the effect of invalidation on citations to patents held by large and non-large patent owners, respectively, for the subsample of patents held by large patent owners, columns (3) and (4) analogously for the subsample of patents held by non-large patent owners. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table 10: Impact of invalidation as opposition outcome on EP/WO citations – patent thickets and patent fences

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Other	Other
Subsample	Thicket	No thicket	Fence	No fence
Invalidated (d)	−0.069	0.202**	0.113	0.289***
	(0.124)	(0.073)	(0.116)	(0.074)
log(No of claims)	0.055***	0.057***	0.046***	0.055***
	(0.015)	(0.006)	(0.010)	(0.006)
log(CitEPExaPre3Other)	0.128***	0.131***	0.111***	0.135***
	(0.016)	(0.007)	(0.011)	(0.007)
log(CitEPExaPre3Self)	0.031	0.011	0.014	0.039***
	(0.022)	(0.009)	(0.012)	(0.011)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes	Yes*	Yes	Yes*
Technology effects	Yes**	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes**	Yes†	Yes*
Patent holder characteristics	Yes***	Yes**	Yes*	Yes†
Opponent characteristics	Yes	Yes***	Yes	Yes***
Underidentification test	101.6	245.3	87.6	272.6
Weak identification test	125.6	555.5	156.1	531.0
Observations	3,239	28,494	8,826	24,233

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the different effects of invalidation on EP/WO examiner citations in the presence or absence of patent thickets and patent fences. Columns (1) and (2) represent a sample split with respect to the presence of a patent thicket in the focal patent's technology area. We consider a thicket to be present if the area triples variable derived by Von Graevenitz et al. (2011) lies at or above the 90th percentile in the full sample. Columns (3) and (4) show the effect of invalidation for a sample split with respect to the presence of a patent fence erected by the holder of the focal patent. We consider a fence to be present if we find at least one similar patent by the focal patent owner prior to opposition. The similarity measure we use is sensitive to the title, the claims, the technology area and the full text of the patent. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table 11: Impact of invalidation as opposition outcome on EP/WO citations – Patent age and value

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Other	Other
Subsample	Younger	Older	Smaller family	Larger family
Invalidated (d)	0.240*	0.073	0.101	0.261**
	(0.102)	(0.092)	(0.084)	(0.091)
log(No of claims)	0.071***	0.041***	0.048***	0.059***
	(0.007)	(0.007)	(0.007)	(0.008)
log(CitEPExaPre3Other)	0.173***	0.098***	0.149***	0.111***
	(0.010)	(0.008)	(0.009)	(0.008)
log(CitEPExaPre3Self)	0.034**	0.005	0.030*	0.016
	(0.012)	(0.010)	(0.013)	(0.010)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes	Yes	Yes*
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes**	Yes	Yes*
Patent holder characteristics	Yes	Yes***	Yes	Yes***
Opponent characteristics	Yes	Yes**	Yes**	Yes†
Underidentification test	128.2	195.8	80.5	222.6
Weak identification test	231.4	339.8	311.7	346.4
Observations	16,981	16,094	17,188	15,880

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: In this table we explore the differences of the invalidation effect with respect to the age of the focal patent at the time of the opposition division's decision and with respect to the size of its DOCDB family, a common patent value indicator. In columns (1) and (2) we split the sample at the age median (8 years), where "Younger" refers to patents of age ≤ 8 years and "Older" refers to patents of age > 8 years. In columns (3) and (4) the sample is split at the median DOCDB family size (8 members), "Smaller family" referring to patents with a family size ≤ 8, "Larger family" referring to patents with a family size > 8. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

4.3 Robustness tests

Robustness across main technology area subsamples

To verify that the results reported in Tables 9 and 10 are not exclusively driven by a single technology area, we report analogous regressions for the chemistry and the electrical engineering / instruments subsamples in Tables R-2 to R-5, finding qualitatively very similar coefficients.

Exclusion of particular cases

Table R-6 shows that our results are not merely artifacts of very particular patents or final outcomes. In column (1) we exclude "dead" patents, i.e., patents solidified in the opposition proceeding which lapse prior to the end of the citation window 5 years after the opposition decision. Column (2) presents the results with patents with accelerated examinations excluded, to rule out the possibility that the effect is solely driven by patents of special interest to the applicant. To mitigate concerns addressing the use of the opposition decision instead of the final outcome of a potential appeal, in columns (3) and (4) we exclude all cases in which an appeal leads to a reversal of the opposition decision and in which any appeal is filed, respectively.

Focus on the extensive margin

Additionally, we limit our count of forward citations to the first of each unique follow-on innovator within the respective time frame. This operationalization allows us to estimate the effect of invalidation on the extensive margin of follow-on innovation. The results are very similar to the ones in our main section.

Dummy citation variables

The regressions of Table R-7 follow our baseline specification with all citation variables replaced with the corresponding dummy variables indicating that at least one citation has been made. The results closely reproduce the findings of Table 6.

Alternative definition of opposition outcome

We further test whether the results are robust to an alternative operationalization of our key dependent variable "invalidation". Instead of treating all patents subject to an amendment as invalidated, we choose a demarcation based on the relative loss of patent scope due to opposition. Patents that lose a smaller number of claims relative to the median of all amendment cases ($N = 5,415$) are treated as remaining valid. The coefficients are quite similar to the ones when using the standard operationalization but less precisely estimated (see Table R-8 and R-9).

Exclusion of citations by focal patent's examiner

To rule out potential concerns that the involvement of the focal examiner in the opposition proceeding may modify his powers of recall, we include only those citations, for which we can exclude that they were made by the focal patent's examiner (Table R-10). Due to resulting data restrictions we have to limit the sample to patents with an application filing year ≥ 2001. Despite a substantial reduction in the number of observations and in the citation count, the results closely resemble those of Table 7. We can hereby rule out potentially modified powers of recall when a focal examiner involved in the opposition proceeding is compiling subsequent search reports as a main driver of the observed effect.

US citations

Tables US-5 to US-10 demonstrate that all findings are qualitatively similar when using US citations. This alternative measure of follow-on innovation results in a dependent variable with much higher variation and more non-zero observations (see Figure A-3 in the Appendix). Besides, we are able to rule out the citation behavior of EP/WO examiners as the key driver of the effect.

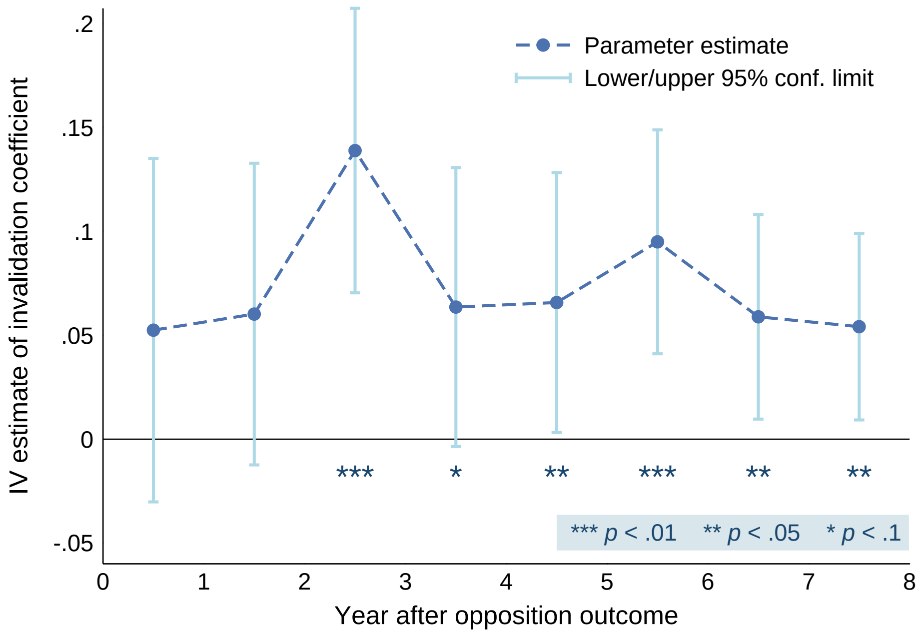
Timing of the invalidation effect

Figure 4 provides some insights into the timing of the invalidation effect. For each year after the opposition outcome, we run IV regressions with a dummy dependent citation variable indicating whether or not a patent has been cited in the respective time span. Significant coefficients of invalidation are only found starting from the third year after opposition, with the third and the sixth year showing particularly large effects. This supports the interpretation of a true change in inventive behavior underlying the increase in the probability of being cited and attenuates the potential concern that the effect is mainly driven by the examiners' increased attention and memory for invalidated patents when searching prior art for subsequent inventions. Figures A-4 and A-5 show the corresponding results for the chemistry and the electrical engineering / instruments subsamples.

5 Conclusion

In this study we investigate the causal effect of a patent's invalidation on follow-on inventions. As our empirical setting we use the opposition procedure at the European Patent Office which allows rivals and other third parties to invalidate patents centrally, i.e., before they are converted to a bundle of national patent grants in up to 38 signatory states of the European Patent Convention. In order to take the presumably endogenous nature of opposition outcomes into account, we introduce a new instrument that exploits the participation or absence of the patent examiner who had granted the patent in the first place in the opposition proceeding. Participation of the examiner in the opposition division is associated with a significant reduction of invalidation outcomes. Given that opposition is relatively frequent with about 6% of patents being attacked, we can test the impact of invalidation on post-opposition citations referencing the opposed patent. In line with the results presented by Galasso and Schankerman (2015), our baseline model shows that overall patent invalidation leads to a highly significant increase of other party and total forward citations. Forward citations by others increase by about 23% following an invalidation. At the same time we do not find evidence that follow-on inventive activity by the holder of the focal patent is affected.

Figure 4: Timing of the invalidation effect



Notes: Blue points depict the coefficients of invalidation resulting from IV regressions for each year after opposition outcome, where as the dependent variable we use a dummy citation variable indicating whether or not a patent has been cited in the respective time span. The usual independent citation control variables (Pre3Self and Pre3Other) are also replaced by dummies. Error bars show the corresponding lower and upper 95% confidence limits. The significance levels are indicated by stars below each parameter estimate.

However, concerning the origin of the invalidation effect, our results strongly contradict previous findings. First, the positive effect of invalidation on subsequent citations is confined to discrete technology industries; second, to small and medium-sized patent holders facing same-size follow-on innovators; third, to a lesser extent, to large patent holders facing large follow-on innovators, and finally, to areas where patent thickets are absent and to patents which are not protected by "patent fences", i.e., similar applications of the same applicant. We probe the robustness of our results in various ways. In the relatively large subsamples of chemistry and electrical engineering / instruments patents, we confirm that invalidation does not appear to lead to an increase in citations when thickets or fences are present. In a second round of robustness tests we use USPTO citations and confirm for most cases the results obtained with EP/WO citations.

Our results for chemistry are in line with those of Murray and Stern (2007), Huang and Murray (2009), and Williams (2013) who focus on IP protection in genome analysis. Moreover, Moser and Voena (2012) find an increase in innovation from compulsory free licensing in the chemical sector. We do not obtain a strong result for electrical engineering and instruments patents per se, but identify a positive effect of invalidation on subsequent citations for small and medium-sized patent-holders and same-size follow-on inventors. These results are broadly in line with findings by Watzinger et al. (2016).

In the study most similar to ours, Galasso and Schankerman (2015) obtain very different results.

In their regressions, the invalidation of chemistry patents by the CAFC does not result in an increase of forward citations, while they do find such an effect for computers, electronics, and medical instruments (pooled with biotechnology). They also find that the strongest effect emerges when patents of large firms are invalidated, while in our EPO oppositions sample we find that the invalidation of patents of small and medium-sized firms leads to a particularly strong increase in subsequent forward citations, in particular for citations coming from other small and medium-sized applicants. Our results are definitely in line with a view that bargaining problems are particularly pronounced between small players, while having a large player on either or both sides of the negotiation table helps to alleviate the bargaining problems. We find that view more appealing than one in which the presence of large players causes bargaining problems.

Nonetheless, the differences are intriguing. They may suggest new avenues for exploring the nature of cumulative effects in invention processes. We speculate that the samples used here differ in more ways than we initially anticipated. Opposition is far less selective than selection into cases heard at the CAFC. That allows us to work with a much larger number of observations, but the nature of the cases may change as well. Presumably the CAFC cases are particularly valuable, but even more, they are particularly uncertain which presumably prevented the parties to achieve some form of settlement. In the case of oppositions, settlements are prevented by institutional means, reducing the ex ante selection effects, but increasing the heterogeneity of cases in the sample. The comparatively low cost of opposition work in the same direction. Further insights illuminating the context-specific selection mechanisms are presumably needed to explain the differences. One way of resolving the issue might be to include in a study of US litigation data cases that reached a final conclusion at an earlier litigation stage and ended with an invalidation (or rejection thereof) and were not brought to the CAFC. A second approach would be to refine the Europen opposition sample we use here further to select cases in a way that mimicks the US selection. We intend to pursue this approach. Leaving the differences of recent results aside, our estimates clearly support the view that the invalidation of patents does - in the aggregate - lead to a rekindling of inventive processes.

References

Acemoglu, D. and U. Akcigit (2012). Intellectual Property Rights Policy, Competition and Innovation. *Journal of the European Economic Association 10*(1), 1–42.

Aghion, P. and P. Howitt (1992). A Model of Growth Through Creative Destruction. *Econometrica 60*(2), 323–351.

Alcacer, J., M. Gittelmann, and B. Sampat (2009). Applicant and Examiner Citations in US Patents: An Overview and Analysis. *Research Policy 38*(2), 415–427.

Baten, J., N. Bianchi, and P. Moser (2015). Does compulsory licensing discourage invention? evidence from german patents after wwi.

Baum, C., M. Schaffer, and S. . Stillman (2010). ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression.

Bessen, J. and E. Maskin (2009). Sequential Innovation, Patents, and Imitation. *RAND Journal of Economics 40*(4), 611–635.

Bessen, J. and M. J. Meurer (2013). The Patent Litigation Explosion. *Loyola University Chicago Law Journal 45*(2), 401–440.

Breschi, S. and F. Lissoni (2004). *Handbook of Quantitative Science and Technology Research*, Chapter Knowledge Networks from Patent Data, pp. 613–643. Springer Netherlands.

Burke, P. F. and M. Reitzig (2007). Measuring Patent Assessment Quality – Analyzing the Degree and Kind of (In)consistency in Patent Offices' Decision Making. *Research Policy 36*(9), 1404–1430.

Cockburn, I. M., M. J. MacGarvie, and E. Mueller (2010). Patent Thickets, Licensing and Innovative Performance. *Industrial and Corporate Change 19*(3), 899–925.

Cohen, W. M., R. R. Nelson, and J. P. Walsh (2000). Protecting Their Intellectual Assets Appropriability Conditions and Why US Manufacturing Firms Patent (or not). National Bureau of Economic Research.

Cremers, K., M. Ernicke, F. Gaessler, D. Harhoff, C. Helmers, L. McDonagh, P. Schliessler, and N. van Zeebroeck (2016). Patent Litigation in Europe. *European Journal of Law and Economics, forthcoming*.

Criscuolo, P. and B. Verspagen (2008). Does it Matter Where Patent Citations Come from? Inventor vs. Examiner Citations in European Patents. *Research Policy 37*(10), 1892–1908.

de Rassenfosse, G., A. B. Jaffe, and E. Webster (2016). Low-quality Patents in the Eye of the Beholder: Evidence from Multiple Examiners.

Farre-Mensa, J., D. Hedge, and A. Ljungqvist (2016). The Bright Side of Patents.

Galasso, A. and M. Schankerman (2015). Patents and Cumulative Innovation: Causal Evidence from the Courts. *Quarterly Journal of Economics 130*(1), 317–369.

Galasso, A. and M. Schankerman (2016). Patents Rights, Innovation and Firm Exit.

Gallini, N. and S. Scotchmer (2002). *Innovation Policy and the Economy, Volume 2*, Chapter Intellectual Property: When Is It the Best Incentive System?, pp. 51–77. MIT Press.

Gaulé, P. (2016, July). Patents and the Success of Venture Capital Backed Startups.

Graham, S. J. and D. Harhoff (2014). Separating Patent Wheat from Chaff: Would the US Benefit from Adopting a Patent Post-Grant Review? *Research Policy 43*(9), 1649–1659.

Grossman, G. M. and E. Helpman (1991). Quality ladders in the theory of growth. *The Review of Economic Studies 58*(1), 43–61.

Hall, B. and D. Harhoff (2012). Recent Research on the Economics of Patents. *Annual Review of Economics 4*, 541–565.

Harhoff, D. (2014). Patent Similarity - An Analysis Tool and Applications. http://www.oecd.org/site/stipatents/1_2_Harhoff.pdf. [accessed: 22 October 2016].

Harhoff, D., B. Hall, G. von Graevenitz, K. Hoisl, and S. Wagner (2007). The Strategic Use of Patents and Its Implications for Enterprise and Competition Policies. Technical report, Tender for No ENTR/05/82.

Harhoff, D. and M. Reitzig (2004). Determinants of Opposition Against EPO Patent Grants. *International Journal of Industrial Organization 22*(4), 443–480.

Harhoff, D. and S. Stoll (2015). Exploring the Opaqueness of the Patent System.

Harhoff, D., G. von Graevenitz, and S. Wagner (2016). Conflict Resolution, Public Goods and Patent Thickets. *Management Science 62*(3), 704–721.

Harhoff, D. and S. Wagner (2009). The Duration of Patent Examination at the European Patent Office. *Management Science 55*(12), 1969–1984.

Heller, M. A. and R. S. Eisenberg (1998, May). Can Patents Deter Innovation? The Anticommons in Biomedical Research. *Science 280*(5364), 698–701.

Huang, K. G. and F. E. Murray (2009). Does Patent Strategy Shape the Long-Run Supply of Public? Knowledge Evidence from Human Genetics. *The Academy of Management Journal 52*(6), 1193–1221.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97 – 126.

Lanjouw, J. O. and M. Schankerman (2004). Protecting Intellectual Property Rights: Are Small Firms Handicapped? *Journal of Law and Economics 47*(1), 45–74.

Lemley, M. A. and C. Shapiro (2007). Patent Holdup and Royalty Stacking. *Texas Law Review 85*, 1991–2048.

Mejer, M. and B. van Pottelsberghe de la Potterie (2012). Economic Incongruities in the European Patent System. *European Journal of Law and Economics 34*(1), 215–234.

Moser, P. and A. Voena (2012). Compulsory Licensing: Evidence from the Trading with the Enemy Act. *The American Economic Review 102*(1), 396–427.

Murray, F. E. and S. Stern (2007). Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis. *Journal of Economic Behavior & Organization 63*(4), 648–687.

Sampat, B. and H. L. Williams (2015). How Do Patents Affect Follow-on Innovation? Evidence from the Human Genome.

Sampat, B. N. (2010). When do Applicants Search for Prior Art? *Journal of Law and Economics 53*(2), 399–416.

Schmoch, U. (2008). Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation (WIPO).

Schneider, C. (2011). The Battle for Patent Rights in Plant Biotechnology: Evidence from Opposition Filings. *The Journal of Technology Transfer 36*(5), 565–579.

Scotchmer, S. (1991). Standing on the Shoulders of Giants. *The Journal of Economic Perspectives 5*(1), 29–41.

Scotchmer, S. and J. Green (1990). Novelty and Disclosure in Patent Law. *The RAND Journal of Economics 21*(1), 131–146.

Shapiro, C. (2001). *Navigating the Patent Thicket: Cross Licensing, Patent Pools, and Standard Setting*, pp. 119–150. MIT Press.

Stock, J. H. and M. Yogo (2005). Testing for Weak Instruments in Linear IV Regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*.

Teece, D. J. (1986). Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy. *Research Policy 15*(1), 285–305.

Von Graevenitz, G., S. Wagner, and D. Harhoff (2011). How to Measure Patent Thickets – A Novel Approach. *Economics Letters 111*(1), 6–9.

Watzinger, M., T. Fackler, M. Nagler, and M. Schnitzer (2016). Antitrust, Patents, and Cumulative Innovation: Evidence from Bell Labs.

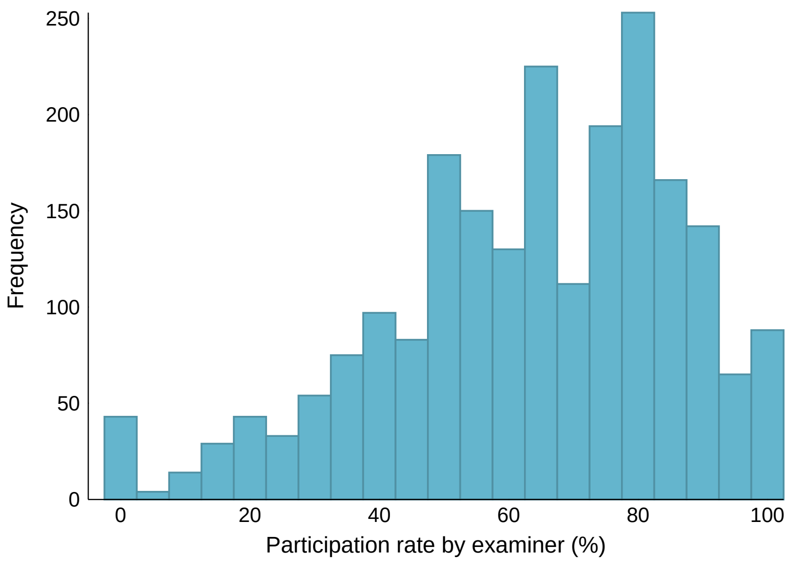
Williams, H. L. (2013). Intellectual Property Rights and Innovation: Evidence from the Human Genome. *Journal of Political Economy 121*(1), 1–27.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (Second ed.). MIT press.

Ziedonis, R. H. (2004). Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms. *Management Science 50*(6), 804–820.

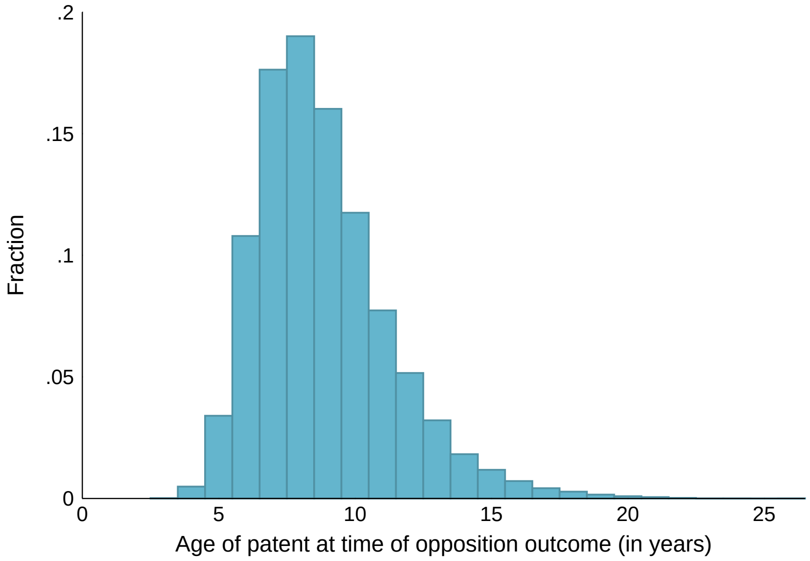
A Appendix: Figures

Figure A-1: Distribution of examiner-specific participation rates



Notes: This graph shows the distribution of participation rates at examiner level. Examiners with less than 3 observations excluded.

Figure A-2: Distribution of patent age



Notes: This graph includes all opposition proceedings (on patent level) which are part of our main sample of analysis.

Figure A-3: Quantile-Quantile Plot: EP/WO examiner citations vs US citations

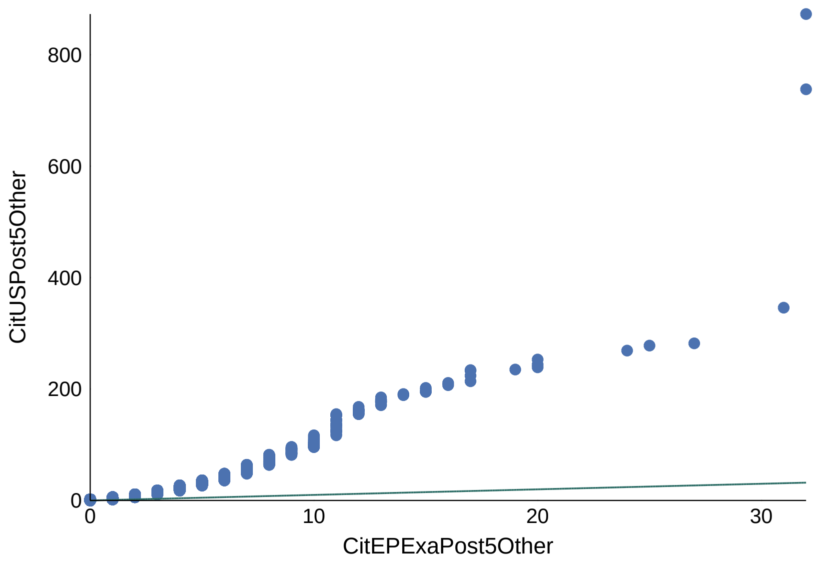
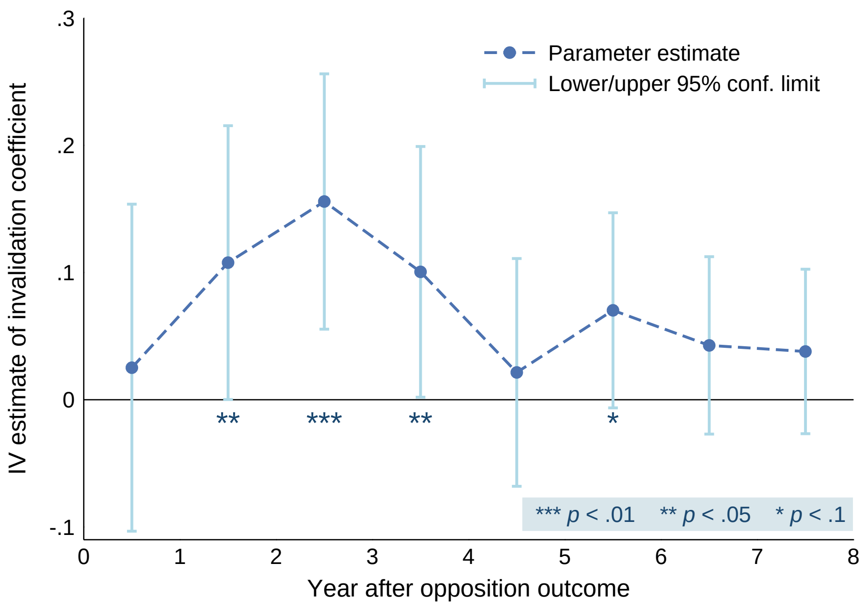
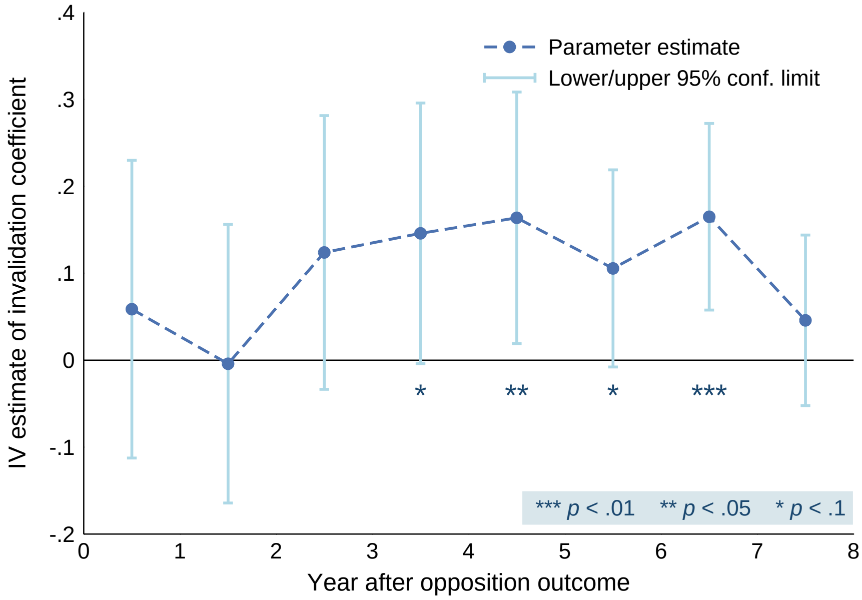


Figure A-4: Timing of the invalidation effect – Chemistry subsample



Notes: This figure is in direct analogy to Fig. 4 in the main text, but with the sample restricted to chemistry patents. Blue points depict the coefficients of invalidation resulting from IV regressions for each year after opposition outcome, where as the dependent variable we use a dummy citation variable indicating whether or not a patent has been cited in the respective time span. The usual independent citation control variables (Pre3Self and Pre3Other) are also replaced by dummies. Error bars show the corresponding lower and upper 95% confidence limits. The significance levels are indicated by stars below each bar.

Figure A-5: Timing of the invalidation effect – Electr. Engineering / Instruments subsample



Notes: The IV regressions underlying this figure are restricted to electrical engineering and instrument patents, in direct analogy to Fig. A-4.

B Appendix: Descriptive tables

Table B-1: Overview and definition of subsamples

Sample definition	N	%
All patents with filed opposition and grant date 1993-2011	49,938	100.00%
– destroyed files	8	0.02%
– unavailable files	150	0.30%
⇒ **available in online file inspection register**	49,780	99.68%
– no readable examiner information	1,203	2.41%
⇒ **with (primary) examiner information**	48,577	97.27%
– patent holder requests revocation	2,031	4.07%
– patent holder withdraws patent	514	1.03%
– opponent withdraws opposition	3,863	7.74%
– no readable opposition information	338	0.68%
– opposition proceeding still pending	470	0.94%
⇒ **with opposition division information**	41,358	82.82%
– first decision after 2011	8,283	16.59%
⇒ **sample of analysis**	33,075	66.23%

Table B-2: Opposition outcomes and appeals

	Oppositions	Appeal rate	Reversal rate
Electrical engineering			
Outcome: valid	982	0.39	0.18
Outcome: invalid	2,458	0.45	0.02
Instruments			
Outcome: valid	1,136	0.46	0.23
Outcome: invalid	3,086	0.50	0.03
Chemistry			
Outcome: valid	3,277	0.43	0.22
Outcome: invalid	9,734	0.50	0.02
Mechanical engineering			
Outcome: valid	3,496	0.43	0.21
Outcome: invalid	6,890	0.45	0.02
Other Fields			
Outcome: valid	743	0.48	0.17
Outcome: invalid	1,273	0.46	0.02

Notes: Reversal rate unconditional on appeal.

Table B-3: Groups of control variables

Group name	**Variables in group**
Year effects	Dummies for grant year Dummies for opposition outcome year
Age effects	Dummies for age in years
Technology effects	Dummies for technology class (34)
Patent characteristics	Dummy for PCT application Dummy for accelerated examination Dummy for examination in Munich Dummies for publication language Size of docdb family Number of IPC classes Number of inventors log(1 + Number of patent literature references) log(1 + Number of non-patent literature references)
Patent holder characteristics	Number of applicants Dummies for patent holder country Dummy for patent holder corporation Dummies for patent holder patent portfolio size: tertiles within technology: small – medium – large
Opponent characteristics	Number of opponents Dummies for opponent country Dummy for opponent corporation
Examination characteristics	Duration of examination Duration of wait until examination

C Appendix: Robustness

C.1 EP/WO citations – Robustness regressions

Table R-1: Probit regressions of instrumental dummy variable "Examiner participation in opposition proceeding" on patent and examination characteristics

	(1)	(2)	(3)	(4)
Estimation method	Probit	Probit	Probit	Probit
Dependent variable	Exam. part.	Exam. part.	Exam. part.	Exam. part.
log(No of claims)	−0.031***	−0.020***	−0.004	−0.005
	(0.004)	(0.004)	(0.004)	(0.004)
log(CitEPExaPre3Other)	−0.008†	−0.004	0.001	0.001
	(0.005)	(0.005)	(0.005)	(0.005)
log(CitEPExaPre3Self)	0.003	0.008	0.005	0.006
	(0.006)	(0.006)	(0.007)	(0.007)
PCT application (d)		−0.045***	−0.002	−0.003
		(0.006)	(0.007)	(0.007)
Accelerated examination (d)		−0.012	0.022**	0.020*
		(0.009)	(0.008)	(0.009)
Examined in Munich (d)		0.120***	0.017*	0.017*
		(0.008)	(0.008)	(0.008)
Publication language: German (d)		0.013	0.010	0.010
		(0.010)	(0.010)	(0.010)
Publication language: English (d)		0.032**	0.013	0.002
		(0.010)	(0.010)	(0.011)
Docdb family size		−0.001***	0.000	0.000
		(0.000)	(0.000)	(0.000)
No of IPC classifications		0.006***	0.001	0.001
		(0.001)	(0.001)	(0.001)
No of inventors		0.002	0.003†	0.003†
		(0.002)	(0.002)	(0.002)
log(Patent backward references)		−0.021***	−0.005	−0.005
		(0.005)	(0.005)	(0.005)
log(Non-patent backward references)		0.006	0.015***	0.013**
		(0.004)	(0.004)	(0.004)
Duration of examination (yr)		−0.015***	0.005	0.004
		(0.002)	(0.006)	(0.007)
Duration of wait (yr)		−0.014***	0.008	0.007
		(0.003)	(0.007)	(0.007)
Year effects	No	No	Yes***	Yes***
Age effects	No	No	Yes*	Yes*
Technology effects	No	No	Yes***	Yes***
Patent holder characteristics	No	No	No	Yes†
Opponent characteristics	No	No	No	Yes
Model degrees of freedom	3	15	97	112
χ^2-statistic	73.7	596.1	2,757.0	2,777.4
Pseudo-R^2	0.002	0.015	0.073	0.073
Observations	33,075	33,075	33,075	33,075

Marginal effects; Robust standard errors in parentheses
(d) for discrete change of dummy variable from 0 to 1
† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: One is added to all citation variables before taking the logarithm to include patents with no citations. A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-2: Impact of invalidation as opposition outcome on EP/WO citations – by sizes of focal and citing patent holders – chemistry subsample

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExa…Post5Other)	Large	Non-large	Large	Non-large
Patent holder subsample	Large	Large	Non-large	Non-large
Invalidated (d)	0.039	−0.079	0.137*	0.239**
	(0.104)	(0.126)	(0.058)	(0.088)
log(No of claims)	0.026*	0.029**	0.015*	0.031***
	(0.010)	(0.011)	(0.006)	(0.009)
log(CitEPExaPre3Other)	0.033**	0.063***	0.052***	0.070***
	(0.011)	(0.012)	(0.007)	(0.010)
log(CitEPExaPre3Self)	0.020†	−0.010	0.027*	−0.004
	(0.012)	(0.013)	(0.011)	(0.013)
Year effects	Yes**	Yes***	Yes***	Yes***
Age effects	Yes	Yes	Yes***	Yes
Technology effects	Yes*	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes*	Yes	Yes	Yes
Patent holder characteristics	Yes*	Yes*	Yes†	Yes***
Opponent characteristics	Yes	Yes	Yes†	Yes†
Underidentification test	36.0	36.0	142.3	142.3
Weak identification test	93.0	93.0	240.5	240.5
Observations	4,321	4,321	8,677	8,677

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the impact of invalidation on EP/WO examiner citations *in the chemistry subsample* with respect to the differences in size between the holder of the citing patent (dependent variable) and the holder of the focal patent (subsample). Columns (1) and (2) show the effect of invalidation on citations to patents held by large and non-large patent owners, respectively, for the subsample of patents held by large patent owners, columns (3) and (4) analogously for the subsample of patents held by non-large patent owners. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-3: Impact of invalidation as opposition outcome on EP/WO citations – patent thickets and patent fences – chemistry subsample

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Other	Other
Subsample	Thicket	No thicket	Fence	No fence
Invalidated (d)	−0.209	0.231*	0.126	0.298**
	(0.157)	(0.096)	(0.161)	(0.096)
log(No of claims)	0.050*	0.047***	0.027	0.048***
	(0.021)	(0.009)	(0.017)	(0.010)
log(CitEPExaPre3Other)	0.084***	0.099***	0.086***	0.104***
	(0.023)	(0.010)	(0.017)	(0.011)
log(CitEPExaPre3Self)	0.002	0.000	0.004	0.018
	(0.027)	(0.013)	(0.018)	(0.014)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes	Yes*	Yes	Yes
Technology effects	Yes**	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes
Patent holder characteristics	Yes*	Yes*	Yes	Yes†
Opponent characteristics	Yes	Yes	Yes	Yes
Underidentification test	63.8	96.7	42.5	137.8
Weak identification test	82.3	249.6	78.9	253.6
Observations	1,613	10,786	3,629	9,364

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the different effects of invalidation on EP/WO examiner citations *in the chemistry subsample* in the presence or absence of patent thickets and patent fences. Columns (1) and (2) represent a sample split with respect to the presence of a patent thicket in the focal patent's technology area. We consider a thicket to be present if the area triples variable derived by Von Graevenitz et al. (2011) lies at or above the 90th percentile in the full sample. Columns (3) and (4) show the effect of invalidation for a sample split with respect to the presence of a patent fence erected by the holder of the focal patent. We consider a fence to be present if we find at least one similar patent by the focal patent owner prior to opposition. The similarity measure we use is sensitive to the title, the claims, the technology area and the full text of the patent. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-4: Impact of invalidation as opposition outcome on EP/WO citations – by sizes of focal and citing patent holders – electrical engineering / instruments subsample

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExa...Post5Other)	Large	Non-large	Large	Non-large
Patent holder subsample	Large	Large	Non-large	Non-large
Invalidated (d)	−0.006	0.055	0.039	0.402**
	(0.147)	(0.249)	(0.100)	(0.140)
log(No of claims)	0.006	0.023	0.026**	0.069***
	(0.013)	(0.016)	(0.008)	(0.012)
log(CitEPExaPre3Other)	0.080***	0.104***	0.077***	0.112***
	(0.015)	(0.016)	(0.010)	(0.014)
log(CitEPExaPre3Self)	0.039†	0.002	0.016	0.079**
	(0.021)	(0.023)	(0.018)	(0.024)
Year effects	Yes†	Yes***	Yes***	Yes***
Age effects	Yes	Yes	Yes**	Yes
Technology effects	Yes*	Yes***	Yes***	Yes***
Patent characteristics	Yes†	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes
Patent holder characteristics	Yes	Yes*	Yes	Yes*
Opponent characteristics	Yes	Yes**	Yes	Yes
Underidentification test	39.1	39.1	66.3	66.3
Weak identification test	57.1	57.1	113.3	113.3
Observations	2,543	2,543	5,109	5,109

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the impact of invalidation on EP/WO examiner citations *in the electrical engineering / instruments subsample* with respect to the differences in size between the holder of the citing patent (dependent variable) and the holder of the focal patent (subsample). Columns (1) and (2) show the effect of invalidation on citations to patents held by large and non-large patent owners, respectively, for the subsample of patents held by large patent owners, columns (3) and (4) analogously for the subsample of patents held by non-large patent owners. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-5: Impact of invalidation as opposition outcome on EP/WO citations – patent thickets and patent fences – electrical engineering / instruments subsample

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Other	Other
Subsample	Thicket	No thicket	Fence	No fence
Invalidated (d)	−0.152	0.437*	0.188	0.416*
	(0.196)	(0.215)	(0.178)	(0.210)
log(No of claims)	0.075**	0.062***	0.062***	0.058***
	(0.026)	(0.012)	(0.019)	(0.013)
log(CitEPExaPre3Other)	0.170***	0.157***	0.147***	0.156***
	(0.026)	(0.015)	(0.022)	(0.015)
log(CitEPExaPre3Self)	0.107*	0.029	0.021	0.094***
	(0.055)	(0.021)	(0.027)	(0.027)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes	Yes	Yes
Technology effects	Yes***	Yes***	Yes*	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes
Patent holder characteristics	Yes***	Yes*	Yes*	Yes*
Opponent characteristics	Yes	Yes*	Yes*	Yes**
Underidentification test	37.7	73.2	48.8	58.2
Weak identification test	51.1	120.0	53.0	107.7
Observations	1,097	6,200	1,844	5,798

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the different effects of invalidation on EP/WO examiner citations *in the electrical engineering / instruments subsample* in the presence or absence of patent thickets and patent fences. Columns (1) and (2) represent a sample split with respect to the presence of a patent thicket in the focal patent's technology area. We consider a thicket to be present if the area triples variable derived by Von Graevenitz et al. (2011) lies at or above the 90th percentile in the full sample. Columns (3) and (4) show the effect of invalidation for a sample split with respect to the presence of a patent fence erected by the holder of the focal patent. We consider a fence to be present if we find at least one similar patent by the focal patent owner prior to opposition. The similarity measure we use is sensitive to the title, the claims, the technology area and the full text of the patent. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-6: Impact of invalidation as opposition outcome on EP/WO citations – exclusion of particular cases

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Other	Other
Subsample	No dead patents	No acc exam	No rev appeals	No appeals
Invalidated (d)	0.205***	0.176*	0.214***	0.133
	(0.062)	(0.072)	(0.058)	(0.082)
log(No of claims)	0.055***	0.053***	0.055***	0.052***
	(0.005)	(0.005)	(0.005)	(0.007)
log(CitEPExaPre3Other)	0.131***	0.125***	0.130***	0.142***
	(0.006)	(0.007)	(0.006)	(0.008)
log(CitEPExaPre3Self)	0.020*	0.020*	0.017*	0.011
	(0.008)	(0.009)	(0.008)	(0.011)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes*	Yes*	Yes**
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes**	Yes***	Yes**	Yes*
Patent holder characteristics	Yes**	Yes**	Yes**	Yes
Opponent characteristics	Yes**	Yes***	Yes**	Yes
Underidentification test	348.6	258.4	461.1	94.7
Weak identification test	766.4	588.8	970.2	316.1
Observations	30,347	29,389	30,620	17,653

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the “Invalidated” dummy is instrumented with the corresponding probability predicted by a probit regression on the “Examiner participation” dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata’s ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-7: Impact of invalidation as opposition outcome on EP/WO citation dummy variables

	(1)	(2)	(3)	(4)
Estimation method	OLS	IV	IV	IV
Dep var: (CitEPExaPost5...) > 0 (d)	Other	Other	Self	Total
Invalidated (d)	−0.005	0.203***	0.037	0.215***
	(0.006)	(0.051)	(0.030)	(0.052)
log(No of claims)	0.050***	0.044***	0.016***	0.047***
	(0.004)	(0.004)	(0.003)	(0.005)
CitEPExaPre3Other > 0 (d)	0.101***	0.101***	0.009**	0.095***
	(0.006)	(0.006)	(0.003)	(0.006)
CitEPExaPre3Self > 0 (d)	0.015*	0.017*	0.037***	0.034***
	(0.007)	(0.007)	(0.004)	(0.007)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes*	Yes	Yes**
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes***	Yes***	Yes	Yes***
Patent holder characteristics	Yes†	Yes†	Yes***	Yes
Opponent characteristics	Yes***	Yes*	Yes	Yes**
Underidentification test		345.2	345.2	345.2
Weak identification test		701.4	701.4	701.4
Observations	33,075	33,075	33,075	33,075

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table is analogous to Table 6, but has all citation variables replaced with the corresponding dummies indicating at least one citation. Columns (1) and (2) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on EP/WO examiner citations to patents held by other parties than the focal patent owner, as measured by EP/WO examiner forward citations in a 5-year window following the decision of the opposition proceeding. Columns (2)–(4) show 2SLS regressions for the impact of invalidation on the number of follow-on patents held by other parties than the focal patent owner, on the number of follow-on patents held by the focal patent owner himself and on the total number of follow-on patents, respectively. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-8: Impact of invalidation as opposition outcome on EP/WO citations – alternative treatment of "amended" patents

	(1)	(2)	(3)	(4)
Estimation method	OLS	IV	IV	IV
Dep var: log(CitEPExaPost5...)	Other	Other	Self	Total
Invalidated (d)	-0.020^{***}	0.219^{**}	0.040	0.228^{**}
	(0.006)	(0.072)	(0.031)	(0.075)
log(No of claims)	0.060^{***}	0.062^{***}	0.016^{***}	0.070^{***}
	(0.005)	(0.005)	(0.002)	(0.005)
log(CitEPExaPre3Other)	0.129^{***}	0.130^{***}	0.006^{*}	0.130^{***}
	(0.006)	(0.006)	(0.003)	(0.006)
log(CitEPExaPre3Self)	0.019^{*}	0.021^{*}	0.047^{***}	0.050^{***}
	(0.008)	(0.008)	(0.005)	(0.009)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes**	Yes*	Yes	Yes**
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes***	Yes**	Yes	Yes**
Patent holder characteristics	Yes**	Yes**	Yes***	Yes
Opponent characteristics	Yes***	Yes**	Yes	Yes**
Underidentification test		99.1	99.1	99.1
Weak identification test		428.2	428.2	428.2
Observations	33,075	33,075	33,075	33,075

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table is analogous to Table 6, but has cases where the patent remained valid in amended form with fewer claims lost than the global median treated as valid. Columns (1) and (2) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on EP/WO examiner citations to patents held by other parties than the focal patent owner, as measured by EP/WO examiner forward citations in a 5-year window following the decision of the opposition proceeding. Columns (2)–(4) show 2SLS regressions for the impact of invalidation on the number of follow-on patents held by other parties than the focal patent owner, on the number of follow-on patents held by the focal patent owner himself and on the total number of follow-on patents, respectively. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-9: Impact of invalidation as opposition outcome on EP/WO citations – technology and size – alternative treatment of “amended” patents

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	IV	IV	IV	IV	IV	IV
Dep var: log(CitEPExaPost5…)	Other	Other	Other	Other	Other	Other
Subsample	Complex	Discrete	Large	Non-large	Complex or large	Discrete, non-large
Invalidated (d)	0.070	0.240^{**}	0.077	0.298^{***}	0.129	0.276^{**}
	(0.167)	(0.076)	(0.126)	(0.086)	(0.111)	(0.093)
log(No of claims)	0.071^{***}	0.051^{***}	0.048^{***}	0.068^{***}	0.065^{***}	0.053^{***}
	(0.008)	(0.006)	(0.008)	(0.006)	(0.006)	(0.008)
log(CitEPExaPre3Other)	0.155^{***}	0.107^{***}	0.113^{***}	0.139^{***}	0.135^{***}	0.117^{***}
	(0.010)	(0.008)	(0.010)	(0.008)	(0.008)	(0.010)
log(CitEPExaPre3Self)	0.031^{*}	$0.017^{\dagger}$	$0.019^{\dagger}$	0.033^{**}	0.019^{*}	$0.027^{\dagger}$
	(0.014)	(0.010)	(0.011)	(0.012)	(0.010)	(0.014)
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes	Yes	Yes**	Yes†	Yes
Technology effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes*	Yes*	Yes	Yes***	Yes	Yes**
Patent holder characteristics	Yes*	Yes*	Yes	Yes	Yes†	Yes
Opponent characteristics	Yes*	Yes†	Yes*	Yes	Yes**	Yes
Underidentification test	22.5	87.0	28.2	79.6	36.9	85.7
Weak identification test	76.8	386.5	98.8	342.1	157.8	287.6
Observations	14,946	18,129	11,014	22,061	20,911	12,164

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table is analogous to Table 8, but has cases where the patent remained valid in amended form with fewer claims lost than the global median treated as valid. Columns (1) and (2) compare the effect in complex technologies to that in discrete technologies, columns (3) and (4) compare the effect for patents held by large patent holders to that for patents held by non-large patent holders and columns (5) and (6) compare the effect for patents which are in complex technologies or held by a large patent holder to that for patents which are in discrete technologies and held by a non-large patent holder. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the “Invalidated” dummy is instrumented with the corresponding probability predicted by a probit regression on the “Examiner participation” dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata’s ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table R-10: Impact of invalidation as opposition outcome on EP/WO citations – citations added by non-focal examiners

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitEPOtExaPost5...)	Other	Other	Other	Other
Technology area	Electr Eng	Instruments	Chemistry	Mech Eng
Invalidated (d)	0.154	0.387†	0.306*	0.086
	(0.262)	(0.209)	(0.143)	(0.161)
log(No of claims)	0.002	0.081**	0.018	0.041***
	(0.026)	(0.028)	(0.014)	(0.012)
log(CitEPOtExaPre3Other)	0.139**	0.192***	0.027	0.127***
	(0.048)	(0.057)	(0.023)	(0.027)
log(CitEPOtExaPre3Self)	−0.037	0.043	−0.021	0.062†
	(0.065)	(0.063)	(0.025)	(0.035)
Year effects	Yes	Yes**	Yes***	Yes***
Age effects	Yes	Yes**	Yes***	Yes
Technology effects	Yes	Yes*	Yes	Yes***
Patent characteristics	Yes	Yes**	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes
Patent holder characteristics	Yes	Yes	Yes***	Yes
Opponent characteristics	Yes	Yes	Yes	Yes
Underidentification test	13.5	27.7	40.9	23.4
Weak identification test	15.8	28.0	57.2	37.6
Observations	576	725	2,596	2,674

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: While in close analogy to Table 7, the EP examiner citation variables (both dependent and independent) in the IV regressions above include only those citations, for which we can exclude that they were made by the focal patent's examiner. Due to resulting data restrictions we have to limit the sample to patents with an application filing year ≥ 2001. While this reduces the number of observations and the citation count, the coefficients closely reproduce those of Table 7, ruling out potentially modified powers of recall when a focal examiner involved in the opposition proceeding is compiling subsequent search reports as a main driver of the observed effect. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

C.2 US citations

Note: The numbering of the tables is analogous to the one for EP/WO citations in the main text.

Table US-4: Characteristics of US forward citations by relationship to cited patent

	Self citations				Other citations			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Publication authority								
US (d)	1	0.00	1	1	1	0.00	1	1
Citation characteristics								
Citation lag (yr)	10.59	3.38	0	28	10.90	3.83	0	30
Docdb family size	7.28	6.60	1	134	5.77	6.22	1	254
Sector (citing applicant)								
Company (d)	0.99	0.11	0	1	0.86	0.35	0	1
Country of residence (citing applicant)								
EPC (excl. GB) (d)	0.31	0.46	0	1	0.20	0.40	0	1
GB (d)	0.01	0.11	0	1	0.02	0.13	0	1
US (d)	0.58	0.49	0	1	0.62	0.49	0	1
JP (d)	0.09	0.28	0	1	0.08	0.27	0	1
Other (d)	0.02	0.12	0	1	0.08	0.27	0	1
Size (citing applicant)								
Large (d)	0.60	0.49	0	1	0.27	0.45	0	1
Medium (d)	0.28	0.45	0	1	0.23	0.42	0	1
Small (d)	0.12	0.32	0	1	0.50	0.50	0	1
Observations	18,315				137,592			

Notes: This table includes all forward citations of US applications to patents subject to opposition proceedings in our main sample of analysis. The unit of observation is the citation. In case of multiple citing applicants, we give preference according to the ordering of sector, country of residence, and size. Size categories are proxied by the number of patents (incl. applications) filed during the last five years prior to the opposition decision (large: 200 and more patents, medium: 20 and more patents, small: fewer than 20 patents.

Table US-5: Examiner participation and opposition outcome (US citations)

	(1)	(2)	(3)
Estimation method	Probit	Probit	Probit
Dependent variable	Invalidated (d)	Invalidated (d)	Examiner participation (d)
Exam. participation (d)	-0.066^{***}	-0.040^{***}	
	(0.005)	(0.005)	
log(No of claims)		0.030^{***}	-0.005
		(0.004)	(0.004)
log(CitUSPre3Other)		0.008^{*}	0.001
		(0.004)	(0.004)
log(CitUSPre3Self)		-0.011^{*}	0.004
		(0.005)	(0.005)
Duration of examination (yr)		-0.007	0.004
		(0.006)	(0.006)
Duration of wait (yr)		0.009	0.007
		(0.007)	(0.007)
Year effects	No	Yes***	Yes***
Age effects	No	Yes*	Yes*
Technology effects	No	Yes***	Yes***
Patent characteristics	No	Yes***	Yes**
Patent holder characteristics	No	Yes*	Yes†
Opponent characteristics	No	Yes***	Yes
Model degrees of freedom	1	113	112
χ^2-statistic	154.3	2,407.5	2,777.4
Pseudo-R^2	0.004	0.083	0.073
Observations	33,075	33,075	33,075

Marginal effects; Robust standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The probit regressions in columns (1) and (2) illuminate the relevance of the "Examiner participation" dummy for the outcome of the opposition proceeding. The invalidation predictions of the probit regression in column (2)—or equivalent predictions for subsamples and other citation measures—are used as the instrument in the 2SLS instrumental variables regressions throughout the paper. Column (3) shows the probit regression of the "Examiner participation" dummy on the other exogenous variables. One is added to all citation variables before taking the logarithm to include patents with no citations. A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table US-6: Impact of invalidation as opposition outcome on US citations

	(1)	(2)	(3)	(4)
Estimation method	OLS	IV	IV	IV
Dep var: log(CitUSPost5...)	Other	Other	Self	Total
Invalidated (d)	−0.034***	0.304**	0.147**	0.353***
	(0.010)	(0.100)	(0.052)	(0.103)
log(No of claims)	0.099***	0.089***	0.020***	0.093***
	(0.007)	(0.008)	(0.004)	(0.008)
log(CitUSPre3Other)	0.437***	0.435***	0.038***	0.430***
	(0.007)	(0.007)	(0.004)	(0.007)
log(CitUSPre3Self)	0.154***	0.158***	0.176***	0.217***
	(0.010)	(0.010)	(0.008)	(0.011)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes†	Yes†	Yes†
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes**	Yes†	Yes	Yes†
Patent holder characteristics	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes***	Yes**	Yes	Yes**
Underidentification test		345.8	345.8	345.8
Weak identification test		700.6	700.6	700.6
Observations	33,075	33,075	33,075	33,075

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Columns (1) and (2) provide a comparison between the OLS and the 2SLS regressions for the impact of invalidation on US forward citations to patents held by other parties than the focal patent owner, as measured by US forward citations in a 5-year window following the decision of the opposition proceeding. Columns (2)–(4) show 2SLS regressions for the impact of invalidation on the number of follow-on patents held by other parties than the focal patent owner, on the number of follow-on patents held by the focal patent owner himself and on the total number of follow-on patents, respectively. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table US-7: Impact of invalidation as opposition outcome on US citations – technology main areas

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitUSPost5...)	Other	Other	Other	Other
Technology area	Electr Eng	Instruments	Chemistry	Mech Eng
Invalidated (d)	−0.113	0.536	0.313*	−0.067
	(0.265)	(0.359)	(0.147)	(0.191)
log(No of claims)	0.134***	0.098***	0.099***	0.082***
	(0.024)	(0.025)	(0.013)	(0.012)
log(CitUSPre3Other)	0.505***	0.569***	0.349***	0.432***
	(0.021)	(0.019)	(0.011)	(0.013)
log(CitUSPre3Self)	0.182***	0.135***	0.155***	0.167***
	(0.035)	(0.028)	(0.014)	(0.019)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes***	Yes†	Yes	Yes***
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes*	Yes	Yes*
Patent holder characteristics	Yes	Yes*	Yes***	Yes***
Opponent characteristics	Yes	Yes*	Yes†	Yes†
Underidentification test	51.8	54.2	166.2	76.0
Weak identification test	90.9	76.3	327.1	115.4
Observations	3,432	4,220	13,011	10,384

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Columns (1)–(4) show the impact of invalidation on US forward citations to patents held by parties other than the focal patent holder for the technology main area subsamples Electrical Engineering, Instruments, Chemistry and Mechanical Engineering, respectively. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table US-8: Impact of invalidation as opposition outcome on US citations – technology and size

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	IV	IV	IV	IV	IV	IV
Dep var: log(CitUSPost5...)	Other	Other	Other	Other	Other	Other
Subsample	Complex	Discrete	Large	Non-large	Complex or large	Discrete, non-large
Invalidated (d)	$0.280^{\dagger}$	0.301^{*}	0.099	0.395^{***}	$0.262^{\dagger}$	0.307^{*}
	(0.161)	(0.126)	(0.178)	(0.117)	(0.136)	(0.141)
log(No of claims)	0.088^{***}	0.094^{***}	0.089^{***}	0.084^{***}	0.096^{***}	0.080^{***}
	(0.012)	(0.011)	(0.014)	(0.010)	(0.010)	(0.013)
log(CitUSPre3Other)	0.511^{***}	0.361^{***}	0.396^{***}	0.449^{***}	0.462^{***}	0.382^{***}
	(0.011)	(0.010)	(0.012)	(0.009)	(0.009)	(0.012)
log(CitUSPre3Self)	0.162^{***}	0.156^{***}	0.170^{***}	0.162^{***}	0.165^{***}	0.144^{***}
	(0.017)	(0.013)	(0.015)	(0.014)	(0.013)	(0.017)
Year effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Age effects	Yes*	Yes	Yes*	Yes*	Yes†	Yes
Technology effects	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Patent holder characteristics	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes*	Yes**	Yes***	Yes	Yes**	Yes
Underidentification test	141.1	201.9	102.6	264.6	182.0	170.0
Weak identification test	277.6	429.7	226.8	478.4	397.6	321.6
Observations	14,946	18,129	11,014	22,061	20,911	12,164

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the impact of invalidation on US forward citations to patents held by parties other than the focal patent holder for different sample splits. Columns (1) and (2) compare the effect in complex technologies to that in discrete technologies, columns (3) and (4) compare the effect for patents held by large patent holders to that for patents held by non-large patent holders and columns (5) and (6) compare the effect for patents which are in complex technologies or held by a large patent holder to that for patents which are in discrete technologies and held by a non-large patent holder. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table US-9: Impact of invalidation as opposition outcome on US citations – by sizes of focal and citing patent holders

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitUS...Post5Other)	Large	Non-large	Large	Non-large
Patent holder subsample	Large	Large	Non-large	Non-large
Invalidated (d)	0.085	0.046	0.208^{**}	0.276^{*}
	(0.142)	(0.169)	(0.078)	(0.112)
log(No of claims)	0.039^{***}	0.085^{***}	0.030^{***}	0.081^{***}
	(0.011)	(0.013)	(0.007)	(0.009)
log(CitUSPre3Other)	0.218^{***}	0.337^{***}	0.222^{***}	0.398^{***}
	(0.010)	(0.011)	(0.007)	(0.009)
log(CitUSPre3Self)	0.120^{***}	0.131^{***}	0.076^{***}	0.144^{***}
	(0.013)	(0.014)	(0.011)	(0.013)
Year effects	Yes^{***}	Yes^{***}	Yes^{***}	Yes^{***}
Age effects	Yes^{*}	Yes	Yes	$\text{Yes}^{\dagger}$
Technology effects	Yes^{***}	Yes^{***}	Yes^{***}	Yes^{***}
Patent characteristics	Yes^{***}	Yes^{***}	Yes^{***}	Yes^{***}
Examination characteristics	Yes	Yes	Yes	$\text{Yes}^{\dagger}$
Patent holder characteristics	Yes^{***}	Yes^{***}	Yes^{***}	Yes^{***}
Opponent characteristics	Yes^{***}	Yes^{*}	Yes	$\text{Yes}^{\dagger}$
Underidentification test	102.6	102.6	264.6	264.6
Weak identification test	226.8	226.8	478.4	478.4
Observations	11,014	11,014	22,061	22,061

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the impact of invalidation on US citations with respect to the differences in size between the holder of the citing patent (dependent variable) and the holder of the focal patent (subsample). Columns (1) and (2) show the effect of invalidation on citations to patents held by large and non-large patent owners, respectively, for the subsample of patents held by large patent owners, columns (3) and (4) analogously for the subsample of patents held by non-large patent owners. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

Table US-10: Impact of invalidation as opposition outcome on US citations – patent thickets and patent fences

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dep var: log(CitUSPost5...)	Other	Other	Other	Other
Subsample	Thicket	No thicket	Fence	No fence
Invalidated (d)	−0.190	0.185	0.182	0.379***
	(0.209)	(0.113)	(0.197)	(0.114)
log(No of claims)	0.119***	0.092***	0.066***	0.092***
	(0.024)	(0.009)	(0.016)	(0.009)
log(CitUSPre3Other)	0.427***	0.440***	0.399***	0.445***
	(0.018)	(0.008)	(0.013)	(0.009)
log(CitUSPre3Self)	0.134***	0.152***	0.166***	0.172***
	(0.026)	(0.011)	(0.016)	(0.013)
Year effects	Yes***	Yes***	Yes***	Yes***
Age effects	Yes	Yes†	Yes*	Yes†
Technology effects	Yes***	Yes***	Yes***	Yes***
Patent characteristics	Yes***	Yes***	Yes***	Yes***
Examination characteristics	Yes	Yes†	Yes	Yes*
Patent holder characteristics	Yes***	Yes***	Yes***	Yes***
Opponent characteristics	Yes	Yes**	Yes	Yes***
Underidentification test	101.5	245.2	88.5	273.4
Weak identification test	126.5	555.1	158.7	530.8
Observations	3,239	28,494	8,826	24,233

Robust standard errors in parentheses

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table explores the different effects of invalidation on US citations in the presence or absence of patent thickets and patent fences. Columns (1) and (2) represent a sample split with respect to the presence of a patent thicket in the focal patent's technology area. We consider a thicket to be present if the area triples variable derived by Von Graevenitz et al. (2011) lies at or above the 90th percentile in the full sample. Columns (3) and (4) show the effect of invalidation for a sample split with respect to the presence of a patent fence erected by the holder of the focal patent. We consider a fence to be present if we find at least one similar patent by the focal patent owner prior to opposition. The similarity measure we use is sensitive to the title, the claims, the technology area and the full text of the patent. One is added to all citation variables before taking the logarithm to include patents with no forward citations. In each 2SLS regression the "Invalidated" dummy is instrumented with the corresponding probability predicted by a probit regression on the "Examiner participation" dummy and all other exogenous variables. The underidentification and weak identification tests are the heteroskedasticity-robust Kleibergen and Paap (2006) rk LM and Wald F statistics, respectively, as reported by Stata's ivreg2 command (Baum et al., 2010). A comprehensive list of the control variables contained in the indicated groups can be found in Table B-3 in the appendix.

D Technical Appendix: Construction of the Examiner Participation Dummy Variable

As explained in section 4, we use the presence or absence of the primary examiner on the opposition board as an instrument to allow for causal inference concerning follow-on innovation for the sample of all opposed EP patents between 1993 and 2011. For this purpose, we first identify the relevant set of patents by the EPO PATSTAT Register – 2015 Autumn Edition. Second, to determine the names of the examination and opposition division's members, we download three types of (scanned) pdf-documents from the EPO database for each of the identified patents: the grant decision for the examination division and the minutes of the oral proceedings as well as the opposition outcome decision for the opposition division. We use two types of documents for the latter to reduce the likelihood of errors. Third, we extract and pre-process the image files included in the pdf-files and read the contained information to txt-files using optical character recognition (OCR) software. Fourth, using a keyword search specific to each document type and language, we identify and parse the names of the respective division's members to a standardized format with first and last names separated. Fifth, we check whether one person is a member of both the examination and the opposition division by comparing the names of both divisions with different string similarity measures.

Two aspects are worth noting. First, the use of both the minutes of the oral proceedings and the opposition decision document to identify the opposition division is legitimate, since the division holding the oral proceedings must be the same as the opposition division rendering the decision in writing, otherwise the decision is deemed to be void.[23] Second, in some cases we are unable to identify all relevant members, for example because the EPO database holds the wrong document under the specific link, and in some cases we might erroneously identify the substantive examiner as being present or absent, for example because the scanned document and thus the OCR is of poor quality. However, the read-out quality and success do not depend on the outcome of the opposition, since the corresponding decision document has the same format across all three outcomes, and thus does not affect identification.

[23] See for instance T 390/86 with a decision from 17 November 1987.