Overcoming inefficiencies of patent licensing

Paper to be presented at DRUID21
Copenhagen Business School, Copenhagen, Denmark
October 18-20, 2021

Overcoming inefficiencies of patent licensing: A method to assess patent’s essentiality for technical standards

Rudi Bekkers
Eindhoven University of Technology
School of Innovation Sciences
R.N.A.Bekkers@tue.nl

Elena M. Tur
Eindhoven University of Technology
Technology, Innovation & Society
e.m.mas.tur@tue.nl

Joachim Henkel
Technical University of Munich
TUM School of Management / Innovation & Entrepreneurship
henkel@wi.tum.de

Tommy Vorst, van der
Dialogic innovation & interacttion
Innovation
vandervorst@dialogic.nl

Menno Driesse
Dialogic Innovatie & Interactie
Innovatie
driesse@dialogic.nl

Jorge L. Contreras
University of Utah
S.J. Quinney College of Law
jorge.contreras@law.utah.edu

Abstract

The market for patent licenses is of highest importance for technological innovation, yet riddled with inefficiencies. A particular problem in the context of technical standards such as LTE is a lack of information regarding which patents are ?essential’ to implement the standard. This information is crucial because it simplifies the determination of infringement and implies specific, ?FRAND’ licensing rules. However, while many standards-developing organisations stipulate that such patents be explicitly declared, little is known about which of these are actually essential. The absence of publicly available information on essentiality has significant social costs due to frictions it causes in the licensing market. Responding to calls from industry, courts, and policy makers, several commercial and a few academic studies have attempted such assessments, but each has limitations. This paper reports on the technical feasibility of a system of expert assessments for patent essentiality. In an experiment, based on a factorial design, twenty-eight experts, including many patent examiners, performed 109 assessments, spending a total of over 100 working days. Comparing their outcomes to a high-quality reference point, we conclude that sufficiently accurate assessments, at a price level that allows large scale testing, are certainly technically feasible, and identify routes to further improvement.

Overcoming inefficiencies of patent licensing: A method to assess patent’s essentiality for technical standards

Abstract

The market for patent licenses is of highest importance for technological innovation, yet riddled with inefficiencies. A particular problem in the context of technical standards such as LTE is a lack of information regarding which patents are ‘essential’ to implement the standard. This information is crucial because it simplifies the determination of infringement and implies specific, ‘FRAND’ licensing rules. However, while many standards-developing organisations stipulate that such patents be explicitly declared, little is known about which of these are actually essential. The absence of publicly available information on essentiality has significant social costs due to frictions it causes in the licensing market. Responding to calls from industry, courts, and policy makers, several commercial and a few academic studies have attempted such assessments, but each has limitations. This paper reports on the technical feasibility of a system of expert assessments for patent essentiality. In an experiment, based on a factorial design, twenty-eight experts, including many patent examiners, performed 109 assessments, spending a total of over 100 working days. Comparing their outcomes to a high-quality reference point, we conclude that sufficiently accurate assessments, at a price level that allows large scale testing, are certainly technically feasible, and identify routes to further improvement.

Disclaimer / funding

(** this part is anonymized for peer review)

Acknowledgements

(** this part is anonymized for peer review)

1 Introduction

In the field of high-tech products, academic studies pay considerable attention not only to the product market but also to the markets for knowledge and the related market for patents (Arora et al., 2001a, 2001b). In this field, an extensive body of literature has emerged on patents required to implement technical standards. Such 'standard-essential patents' (SEPs) are of particular interest: unlike regular patents, there is no possibility for a party making products that incorporate these standards (for instance, for a mobile phone, video player or WiFi device) to design around such patents, creating an unusually strong bargaining position for the owners of such patents.

Implementers of such standards often face thousands of patents claimed by their owners to be potentially standard essential. The market for licenses for these patents exhibits several market imperfections, among other things related to transaction costs and information asymmetry. Increasingly calls can be heard for increased transparency in this area to address these market inefficiencies. In its 2017 communication on Standard Essential Patents (European Commission, 2017), the European Commission took the position that 'There is therefore a need for a higher degree of scrutiny on essentiality claims. This would require scrutiny being performed by an independent party with technical capabilities and market recognition, at the right point in time.' At the same time, the Commission recognizes that such a system must be balanced against the costs. While highly accurate assessments for single patents are indeed possible if cost and time are no issue (as is often the case in the context of litigation), a large-scale system that systematically assesses the essentiality of patents for a given standard, which would require highly qualified individuals to perform, would necessitate appropriate costs and time.

Motivated by the above, this paper investigates whether essentiality assessments can be made sufficiently efficient (in terms of time and costs) as well as sufficiently accurate, that one could set up a large-scale system of such assessments and thus overcome important inefficiencies in the market for SEP licensing. Considering the complexity of the assessment task, but also the need for efficiency, this paper focuses on assessments that require about eight hours of work each (i.e. a budget of one day), acknowledging that existing assessments, while considered to be accurate, spend approximately 5 to 10 times as many resources (see Section 3.2, below).

We formulate the following research questions:

> Q1. Given an average time budget of eight hours per patent, how accurately can qualified assessors determine the essentiality of a given patent to a specific standard?
>
> Q2. How do the availability of claim charts and the chosen definition of essentiality affect the above impact?

Note that we focus on the 'technical feasibility' of a large-scale system of essentiality testing. For questions concerning 'institutional feasibility', like who should set up such a system, who should carry out the assessments, who would finance it, and whether there is sufficient support by stakeholders, we refer to the complementary work presented in (**source anonymized for peer review**).

In an experiment, based on a factorial design, twenty-eight experts, including many patent examiners, performed 109 assessments, spending a total of almost 100 working days. We used patent pool data as a high-quality reference point, and were able to obtain not only data on patents accepted by pools, but also on those rejected by pools as well as the full information set that was provided at that time to pools – including claim charts prepared by the patent owner. This setting allows us to replicate the earlier pool assessment as closely as possible. On the basis of our analysis, we conclude that sufficiently accurate assessments, at a price level that allows large scale testing, are certainly technically feasible, and identify routes to further improvement.

The remainder of this paper is organized as follows. In Section 2 we discuss market inefficiencies in the licensing of patents, with a particular focus on SEPs. In Section 3, we discuss various existing attempts to determine essentiality, including patent pools, commercial and academic studies, and court case analyses, and furthermore review recent endeavours using AI approaches. In Section 4 we introduce the experiment design and data for our research, and Section 5 presents the results. Section 6 concludes and provides a discussion.

2 Market inefficiencies in the licensing of standard-essential patents

2.1 The imperfect market for patent licenses

As economists have agreed since Adam Smith (1774), well-functioning markets increase efficiency through specialization and a division of labour. However, efficient markets require parties to have full information. This assumption is frequently violated by information asymmetry, a lack of transparency, and uncertainty about future events. The classical example is Akerlof's (1970) 'Market for Lemons.'

In the market for technologies (MFT) (Arora et. al., 2001a, 2001b), potential efficiency gains, but also market imperfections, are particularly pronounced. Efficiency gains arise, for instance, in the division of labour between startups that generate new technology and large incumbents that commercialize them (e.g., Baumol, 2010), or through general-purpose technologies that can be fruitfully employed in many markets (Bresnahan and Trajtenberg, 1995). Accordingly, the MFT has acquired a high relevance for innovative industries.

A challenge of trading knowledge lies in the fact that it can easily be expropriated. Intellectual property rights, patents in particular, can alleviate this problem by increasing the appropriability of knowledge (Lamoreaux and Sokoloff, 1999; Arora et al., 2001a, 2001b; Arora and Ceccagnoli, 2006; Gans et al., 2008). Accordingly, the market for patent licenses (under which we subsume, for simplicity, both licenses and assignments of patents) has acquired a high importance in its own right (Madiès et al., 2014), and scholars often measure transactions on the MFT through patent transactions (Lamoreaux and Sokoloff, 1999; Gambardella et al., 2007; Serrano, 2010). In fact, for publicly available but patent-protected knowledge, the relevant market is not the MFT, but the market for patents (Fischer and Henkel, 2012: 1531).

The market for patent licenses is ripe with imperfections due to transaction costs, both motivation costs and coordination costs (e.g., Milgrom and Roberts, 1992: 29). Motivation costs cause inefficiencies through information asymmetry (Caves et al., 1983), which in the market for patent licenses is typically bilateral. The patent owner might have private information about the value of the invention, its limitations, or required complementary knowledge that are not evident from the patent text. In turn, the prospective licensee will generally know more about potential applications of the invention and its resulting economic value. To some extent, the consequences of such information asymmetries can be mitigated by a suitable choice of licensing terms (Gallini and Wright, 1990; Beggs, 1992; Macho-Stadler et al., 1996), but inefficiencies remain.

In addition, coordination costs arise due to uncertainty and a lack of transparency. To start with, patent applications are typically not disclosed to the public until 18 months after filing. This is often followed by a considerable time before a grant decision: at the EPO, a patent grant comes on average 6.0 years after the filing date[1], and only then are the final patent text and scope known. Even then, a granted patent may be 'latently invalid' due to prior art not found by the examiner or disclosed by the applicant (Farrell and Shapiro, 2008; Miller, 2013; Henkel and Zischka, 2019). Furthermore, being written in natural language, a patent leaves room for interpretation (Bessen and Meurer, 2005). In particular, for complex technologies such as information and communication technologies (ICT), it may be difficult to ascertain whether a given patent claims features of a particular product (Lemley and Shapiro, 2007; Magliocca, 2007). Finally, the sheer number of patents in a given field may make it difficult to identify those patents that are relevant to a given technology. This problem is aggravated by the issuance of patents on trivial (e.g., Reitzig et al., 2007: 147) or non-novel inventions (e.g., Graham and Mowery, 2003: 226). These difficulties complicate licensing negotiations and entail the related difficulty of identifying potential licensors in the first place.

The inefficiencies in the market for patent licenses can have serious welfare consequences, including the under-utilization of existing technologies, inadvertent or intentional infringement, and unfounded or, in the case of inadvertent infringement and subsequent lock-in, excessive royalty demands (e.g., Jaffe and Lerner, 2006; Lemley and Shapiro, 2007). These consequences are particularly severe when it comes to widely used, standardized technologies, as we discuss next.

2.2 Specific problems posed by licensing of Standard Essential Patents

A significant number of product categories in the ICT sector are characterized by the use of interoperability standards. Each smartphone, tablet and laptop sold today embodies hundreds of different standards (Biddle et al, 2007). Likewise, hundreds of different firms and research organizations have collaborated on the development of widely deployed standards such as UMTS (3G), LTE (4G), 5G, and Wi-Fi. At the same time, these standards are implemented by a large number of equipment makers. The rise of the Internet of Things (IoT) and smart technologies in the context of Grand Societal Challenges is expected to increase the number of implementers significantly.

Interoperability standards are typically developed under the aegis of voluntary associations known as standards-development organizations (SDOs), in which firms collaborate on the development of standards of interest to the industry. As we explain in more detail in Section 3, the policies of most SDOs in the ICT sector try to ensure that for any (known) patents that are required to implement their standards, the patent owner has committed to offer licenses on terms that are 'fair, reasonable and non-discriminatory' (FRAND). These FRAND licenses may be royalty-free or royalty-bearing. Royalty-free licenses, such as those used with respect to standards such as Bluetooth, HTTP, TCP/IP, and USB, are comparatively non-controversial. However, when FRAND licenses are royalty-bearing, SEP licensing takes on the characteristics of a market transaction.

The market for SEPs exhibits a number of peculiarities. First, although standardized products presumably infringe a large number of SEPs, licenses are often not sought or finalized until months or years after products have been placed on the market (Contreras, 2013: 59-62). Second, any given mobile telecommunication product is likely covered by thousands of patents that are declared to be (potential) SEPs (Baron and Pohlmann, 2018; Bekkers et al., 2020). Yet due to vague patent scope and the complexity of the technology, it is often unclear which patents, precisely, cover a given product, component or standard. Finally, it is often uncertain whether patents that have been declared as 'potentially essential' (see Section 3) to a given standard are, in fact, essential, even

[1] For patents applied for at the EPO between 2001 and 2008, the grant delay is on average 6.02 years (with yearly variations in the average of +/- 0.2 years). After that, as a result of truncation, the average gradually declines to 4.0 years for patents applied for in 2014. Source: author's own calculations on the basis of PATSTAT 2020a, comparing the filing date with the day the B1 publication was published.

assuming that the patents are otherwise valid and enforceable. This last question is a crucial one, as non-essential patents can often be worked around or omitted from a product, while essential patents by definition cannot if a product is to comply with the respective standard. Thus, reliable information about a patent's essentiality would go a long way toward creating transparency regarding its use in a given product.

Yet, as discussed in Section 3.1, studies have reported significant 'over-declaration' of SEPs relating to mobile telecommunications (see studies referenced in Contreras, 2019: 211-212), whether or not intentional. It is thus believed that many, if not the majority, of patents declared to be essential to standards in the mobile telecommunications industry are actually not essential. The widespread occurrence of SEP over-declaration creates significant inefficiencies in the market for SEPs. Among these is asymmetric information: Owners of potential SEPs usually have intimate knowledge about their own patented inventions and whether they are likely to be essential or not. Implementers, on the other hand, are typically confronted by dozens of SEP holders with thousands of patents, and typically have limited or no knowledge about the details of individual patents that are claimed to be SEPs. This asymmetry is complicated by long supply chains in which products implementing standards range from generic chips, application-specific chips and modules to intermediate products and end products.

Asymmetric information and associated uncertainties hamper licensing negotiations for SEPs and invite opportunistic behaviour, creating frictions on the MFT and reducing societal welfare. Recognizing these issues, the European Commission (2017: 5) states in its Communication 'Setting out the EU approach to Standard Essential Patents': *'Evidence points to the risk of broad over-declarations and makes a strong case for more reliability with respect to SEP essentiality. Stakeholders report that recorded declarations create a de facto presumption of essentiality in negotiations with licensees. This scenario places a high burden on any willing licensee, especially SMEs and start-ups, to check the essentiality of a large number of SEPs in licensing negotiations. There is therefore a need for a higher degree of scrutiny on essentiality claims.'*

A final inefficiency arising from declarations of patent essentiality is the lack of a consistent standard for defining essentiality among SDOs. Bekkers and Updegrove (2012: 35) identify thirteen different features of essentiality definitions among twelve major SDOs (see also Contreras, 2007: 12-13, discussing additional variants and exclusions). There are two major definitional axes along which SDO essentiality definitions differ: (1) the degree to which they cover optional portions of a standard, and (2) whether they speak in terms of 'technical' or 'commercial' essentiality (the latter meaning that alternatives to the patented solution exist but are unattractive for cost reasons). Yet these terms are vaguely defined (Contreras, 2017: 218-219), leading to uncertainty among patent holders regarding which patents to declare as essential, and causing implementers to question whether patents declared essential in one SDO are assessed as essential under the policies of another SDO.

The problems raised above are particularly serious given the uptake of the IoT, where compared to the smartphone market implementers are much more numerous and heterogeneous, yet much less knowledgeable about IoT technologies and the SEPs that cover them (Henkel, 2021). Thus, processes to reliably assess actual essentiality of declared SEPs are urgently needed to ensure an efficient SEP licensing market. In the following, we review existing attempts at large-scale essentiality assessment, before introducing our own study.

3 Existing attempts to determine essentiality

Recognizing the nature and consequences of inefficiencies in the markets for SEPs, a number of attempts have already been made to determine the essentiality of patents for given technical standards. This section discusses these attempts, and shows that while we can learn from them, none of them to date has provided the market with assessment of a known accuracy level *and* a cost per patent that would make an assessment of all patents potentially essential for a standard feasible. We first discuss how declarations of potential essentiality are made pursuant to SDO IPR policies (Section 3.1), and then discuss recent attempts to assess essentiality using expert-based approaches (Section 3.2) and Artificial Intelligence (AI) and other automated approaches (Section 3.3).

3.1 Procedures for declaring SEPs and resulting over-declaration

In most SDOs, the disclosure of a patent as 'essential' to a standard is based entirely on the self-declaration of the patent holder. Thus, the fact that a patent has been declared essential to a standard does not imply that it is actually essential, and the studies discussed below have found significant 'over-declaration' of SEPs, particularly at SDOs focusing on mobile telecommunications (e.g., Contreras, 2019: 211-212, collecting studies). For example, studies of the 2G, 3G and 4G standards have found over-declaration rates between 8% and 58%, with individual patent holders over-declaring at rates as high as 82% (see Unwired Planet [2017 EWHC 711 at 324-329], citing numerous studies).

Such over-declarations may be unintentional. A party may, for instance, have submitted a technical proposal to an SDO and accurately declared patents covering the proposal as essential, after which the standard may have evolved to exclude the technology in the proposal. Or a party may have made a declaration on the basis of a patent application, and the patent that ultimately issues no longer claims technology included in the standard (see Lerner, 1994; Marco et al., 2019, observing that the scope of an issued patent is often significantly narrower than the original patent application). Or a patentee that is unsure whether a patent is essential may prefer to err on the side of over-declaration, given the significant legal consequences of not declaring an essential patent and the few legal consequences for over-declaration (Contreras, 2017: 223).

However, deliberate over-declarations may also occur. A firm's share in the overall royalties that can be charged for a standard is often approximated, for lack of a better criterion, by its numerical share of declared SEP families (Unwired Planet [2017] EWHC 711, at 182). Accordingly, patentees have an economic incentive to over-declare – a form of strategic behaviour (Dewatripont and Legros, 2013; Bekkers and West, 2009, and Aoki and Arai, 2018). As noted by Judge Birss in the *Unwired Planet v Huawei,* 'Very many more patents are declared to be essential than in fact are essential', and 'it must also be recognised that the fact that rates are negotiated by counting patents creates a perverse incentive to declare as many patents as possible, making over-declaration worse.'[2] Finally, even if a patent owner is convinced about the essentiality of a given patent, this might not actually be the case. Lemley and Simcoe (2019) find that a substantial share of declared essential patents are found non-essential even after they have been carefully chosen for litigation (and, arguably, their owner believes they are essential).

For the sake of completeness, we must also acknowledge that there are reasons why some essential patents are not declared as potentially essential. First, parties that are not participants in the relevant SDO are not under any obligation to declare potentially essential patents at that SDO (Contreras, 2016 (discussing the concept of standards 'outsiders')). Second, SDO disclosure policies are not 'absolute', and what needs to be declared may depend on actual participation in relevant working groups and on the knowledge of individual participants – even if room for 'manoeuvring'

[2] Unwired Planet v Huawei [2017] EWHC 2988 (Pat).

might be limited by the 'good faith' or other requirements that SDOs have in their policies or that are imposed by law (see Bekkers and Updegrove, 2013: 78-80, 82-85).

3.2 Large scale essentiality tests using expert-based approaches

Given the uncertainties and inefficiencies relating to the declaration of SEPs noted above, various mechanisms have been developed to assess the essentiality of sets of patents for a given technical standard. Such analyses have been conducted by different parties for different purposes. In this section, we review these existing approaches[3], and focus on larger scale assessments where hundreds or at least dozens of patents are assessed.

The first patent pools for technology standards, such as those for the MPEG-2 video compression standard and the DVD (digital video disc), started to appear in the 1990s (Uijl et al, 2013). To ensure that their operations would be compatible with competition (antitrust) law, some of these pools sought Business Review Letters from the U.S. Department of Justice Antitrust Division. The analysis conducted under these letters effectively concluded that in order to prevent anti-competitive effects, it was important to ensure that patents included in a pool were essential to the standards in question, and that the patented technologies included in the pool were complements rather than substitutes for one another (Gilbert, 2017).[4] As a result, almost all pools set up formal mechanisms for assessing the essentiality of patents that are proposed for inclusion in the pool. Typically, these assessment procedures (1) require patent owners to propose patents and submit claim charts that demonstrate why the proposed patents are indeed essential, (2) outsource the assessments to independent, external experts (usually at specialized law firms), and (3) have formal appeals procedures for patent owners and – sometimes – for other pool members and/or licensees. While details of the procedures used by specific pools are usually not public, an interesting exception is the 3GPP, also known as the 'WCDMA pool', or the '3G Patent Platform', whose initiators published an extensive book describing their approach (Goldstein & Kearsey, 2004). Among other things, this pool (which is part of the reference set we will discuss below) involves independent, parallel assessments and compare their results. A recent study commissioned by the EC includes a review of essentiality assessment mechanisms in patent pools (**source anonymized for peer review**; see also Merges & Mattioli, 2017).

Pool organizers have extensive experience with such essentiality assessments. Given the strong legal incentives to include only essential patents, pools may be expected to implement diligently performed, high-quality assessment mechanisms. Also, a pool licensing a portfolio of SEPs for a certain standard should have an incentive to build a reputation of licensing actual SEPs; and existing pool members should be unwilling to accept a dilution of their portfolio share through newly added patents unless these are actually essential (Merges & Mattioli 2017; for royalty allocating mechanisms in pools, see Layne-Farrar and Lerner, 2011). The resources they spend per patent are

[3] In this section, we do not review company in-house assessment mechanisms because information on such processes is not publicly shared.

[4] A rather special case is the situation in which a standard can only be implemented by making use of one of a set of two (or more) technologies that are each patented – actually making the patents in this set substitutes. In the definition of essentiality by ETSI (See Section 4.1), such patents are deemed essential after all. However, as indicated in the ETSI definition itself, these are exception cases; and neither during the 50+ hours of face to face talks we had we experts when carry out or work, or from the feedback we collected from the assessors in our experiment, there was any indication that this was happening in practice (even when we asked for it explicitly).

in line with such a high-quality assessment.[5] This does not mean that pool assessments are perfect. There is inherent uncertainty associated with essentiality assessment. In the case of uncertainty about the actual essentiality of a patent submitted by a member, patent pools may be subject to incentives to include that patent, which would lead to over-inclusion. Also, communication between a submitting member and the evaluator as well as appeal opportunities may lead to over-inclusion. And on an aggregate level, a pool might benefit from having more lenient inclusion criteria as it increases its portfolio size relative to that of other patent holders, thus helping to justify higher royalty demands.

In addition to essentiality assessments made in connection with the formation of patent pools, third parties have assessed the essentiality of patents to different standards. Sometimes, this work has been carried out by academics, who often publish it openly, but more often it is conducted by private consulting firms that make the results available only to parties that purchase their reports. One of the first (published) attempts to perform such a systematic essentiality assessment was that of Goodman & Myers (2005), which was executed in the context of a conflict between several companies over patent portfolio value (and also sponsored by one of these companies). Many later studies refer to this work, but it also received criticism (see Martin & De Meyer, 2006). From around 2007 on, a stream of commercial studies appeared that followed this approach, including Fairfield Resources International (Fairfield, 2007, 2009, 2010), which are continuations of the Goodman & Myers (2005) study, as well as studies from Article One Partners (2011), Cyber Creative Institute (2011), Jefferies & Company (2011), iRunway (2012), PA Consulting Group (2015) and Charles River Associates (2016). These studies mostly begin with lists of patents declared as potentially essential – for instance patents declared to ETSI – and perform manual assessments of essentiality. It is difficult to assess the quality of these efforts: the underlying methodology, working assumptions and data processing steps are not generally made public, and there is not any evaluation, such as a comparison of the results to a benchmark of known accuracy, or inter-rater consistency. Validity of outcomes is not extensively discussed in these works. It is also difficult to compare the outcomes of these studies to each other since they differ in terms of the standard that is investigated, data selection, and cut-off dates.

Finally, large-scale essentiality tests have been performed by economics experts in the context of patent litigation. Such cases include *Unwired Planet v Huawei, TCL v Ericsson*, and *In re Innovatio IP Ventures*.[6] The purpose of such analyses was usually to provide input to a court's top-down calculation of FRAND royalties, where the aggregate royalty for all SEPs covering a specific standard is first determined, and then allocated among individual SEPs and SEP holders (Siebrasse & Cotter, 2017). Compared to the analysis by consulting firms discussed above, litigation analyses are somewhat more transparent, yet their procedures vary considerably in design and parameters used. While the courts in the three cases mentioned above recognised that the large-scale essentiality assessments performed by experts were not perfect, they were useful to the courts in making their FRAND royalty determinations.

3.3 Large scale essentiality tests using AI and other automated approaches

Inspired by work on the computation of semantic similarity between patents (Younge & Kuhn, 2016; Arts et. al., 2017), Brachtendorf et al. (2020) investigated the semantic similarity between patents

[5] These resources are estimated to be € 5,000 to € 10,000 for a single European patent, and up to twice as much for a single US patent (see Merges & Mattioli, 2017 and (**source anonymized for peer review**).

[6] Unwired Planet v Huawei, [2017] EWHC 711 (Pat); TCL Commc'n Tech. Holdings, Ltd. v. Telefonaktiebolaget LM Ericsson, No. 8:14-cv-00341 (C.D. Cal. Dec. 21, 2017), rev'd on other grounds, 943 F.3d 1360 (Fed. Cir. 2019); Innovatio IP Ventures, LLC Patent Litigation, 2013 U.S. Dist. LEXIS 144061 (N.D. Ill. Sept 27, 2013).

and standards documents to assess actual essentiality of declared SEPs.[7] The algorithm is validated by comparing the findings with the results of the manual essentiality assessments for the TCL v Ericsson court case, which was mentioned above. At the individual patent level, the consistency this study finds between their own outcomes and the court case data is limited. From the set of 166 patents assessed to be essential by manual evaluators, the automated system predicted only 40 (24%) were essential. From the set of 236 patents assessed not to be essential by manual evaluators, the automated system predicted 216 (92%) were not essential. If we assume the reference point is perfect (which it may not be), then the automated system has many false negatives, and fewer false positives. Yet, the authors find strong and highly significant correlations between the experts' decisions on standard essentiality and their own measurement of semantic similarity, and good accuracy in predicting the share of actual SEPs in a larger portfolio. All in all, these first results are promising but do not yet seem to be satisfactory in terms of predicting essentiality on an individual patent basis.

An undeniable strength of automated approaches is their scalability. In potential, they would allow the analyses of very large sets in a relatively short time span, and at low costs. But they also come with several inherent limitations. Firstly, the meaning, interpretation, and precise scope of words and terminology (both in patents and standards) are dependent on the context, making it hard to automate. Second, semantic approaches can face difficulties dealing with changes in terminology over time. Third, the patent to be evaluated, or parts of it, may be written in a different (natural) language than the respective part of the standard. Furthermore, even with the same natural language, the vocabulary in patents (drafted by patent attorneys) often differs from that in standards (drafted by engineers). Fourth, a technology or solution required to implement the standard may not be explicitly mentioned in the standard's text but may still be required in order to satisfy the standard (i.e., being implied by the standard). Fifth, an essentiality analysis should consider possible alternatives to the patent under investigation that may also satisfy the standard. This means that an automated approach should not only look at the patent under investigation, but also all other patented and non-patented inventions.[8] Sixth, any automated system is prone to gaming, whereby patent owners, anticipating the workings of such a system, will adapt the wording in their patent applications (which might end up in the granted patent claims) to the wording employed in the standard documents.

In sum, the approaches to essentiality assessment discussed above provide useful input in terms of designing an essentiality testing mechanism, but do not yet answer the question of whether essentiality assessments can be made sufficiently efficient (in terms of time and costs) as well as sufficiently accurate, that one could set up a large-scale system of such assessments and thus overcome important inefficiencies in the market for SEP licensing.

[7] The authors identify standards documents on the basis of patent declarations at ETSI, resulting in 4,796 standards documents, and compare them with 37 million patent documents, considering patent claims as well as technological descriptions. The study uses an algorithm developed by Natterer (2016).

[8] The definition of essentiality at ETSI is explicit on this aspect: if alternatives exist that are not patented, the patent in question is not essential; if only alternatives exist that are also patented, then the focal patent is essential (as well as the patented alternatives). (ETSI, 2020: Annex 6, §15, Item 6; see also Contreras, 2017: 218-19). Rules at other SDOs may differ or are not explicit (Bekkers & Updegrove, 2013: 66-67; Contreras, 2017: 218-19).

4 Experiment design and data

In this section, we discuss the experimental design and the associated key choices and specifically elaborate on the definition of essentiality used, the reference points, the selection of assessment cases (and associated data collection), and the assessors, case allocation, and instructions.

The experiment's quantitative part follows a factorial design where the treatment is whether a patent has claim charts or not, and the block is whether the patent was considered essential by a pool. We also gathered qualitative outcomes by asking the assessors both closed questions and open feedback. We did so after each assessment, as well as at the end of all assessments. To ensure that our assessors would have deep expertise in the field of the standards and patents they had to review, we focused the experiment on a single technological area, namely the ETSI/3GPP 3G and 4G standards, and selected assessors accordingly. This technical area is, in fact, one of the primary areas in which the calls for essentiality testing have been made (European Commission, 2017). Furthermore, there are several patent pools active in this area, allowing us to use their essentiality decisions as reference points. There are unobserved variations between cases (technically, 'nuisances') that we cannot control, such as the level of difficulty, so we randomize the data in each cell to limit their impact in our results.

4.1 Definition of essentiality

SDOs and other organisations have adopted different definitions of essentiality (see Section 2.2). Since our experiment considers assessments of patent essentiality for ETSI standards, we in principle followed that organisation's definition of essentiality throughout the experiment: *'ESSENTIAL' as applied to IPR means that it is not possible on technical (but not commercial) grounds, taking into account normal technical practice and the state of the art generally available at the time of standardisation, to make, sell, lease, otherwise dispose of, repair, use or operate EQUIPMENT or METHODS which comply with a STANDARD without infringing that IPR. For the avoidance of doubt in exceptional cases where a STANDARD can only be implemented by technical solutions, all of which are infringements of IPRs, all such IPRs shall be considered ESSENTIAL.'* (ETSI, 2020).

While discussing this definition with the various involved patent offices in preparation for the experiment, some offices raised concerns that patent examiners are not trained in determining infringement and asked whether the assignment could be re-phrased. Together with these offices, we developed an alternative that we call a 'novelty-based test', based on the following thought experiment: *In the hypothetical case that the standard document had already been published before the priority date of the patent, would that document have been novelty-destroying?* We stress that we are *not* asking patent examiners to determine whether the patent is *valid* or not; after all, the text of the standard document is in reality published *after* the priority date of the patent. The different definitions were included as an additional block in our experiment. Several stakeholders indicated that they did not expect the specific definition to make a difference in essentiality assessments; nonetheless, in our analysis we compare the essentiality assessments based on the ETSI definition to those based on the novelty-based test.

4.2 Reference assessments

To determine the accuracy of the assessments in this experiment, a reference is required to compare our results with. The ultimate, authorised decision concerning essentiality (and infringements) lies with competent courts. While some courts have indeed issued (public) verdicts on the essentiality of patents, the number of data points is very limited and may be based on different definitions of essentiality than assumed here (e.g., take only infringement into account). Moreover, we have little insight into the exact information that was used to arrive at that court verdict - making it hard to ensure our assessors would work from the same information. Moreover, there is the risk that our assessors are aware of these court verdicts, thereby creating a possible bias.

For our experiment, we used what we believe to be the most accurate assessment points existing *outside* of a litigation context: the assessment by patent pools. To comply with competition (antitrust) law, these pools have developed diligent and sophisticated procedures where patents submitted to the pool are scrutinised by external, independent parties (usually law firms or patent attorneys specialised in this task), as we discussed in Section 3.2 above. While pool assessments cannot be regarded as 'perfect', they are considered by almost all stakeholders as the gold standard, and we believe they are appropriate as a reference point for our study. If an expert can replicate the pools' assessments, this is a strong indicator of a high accuracy level. Furthermore, by collaborating with patent holders that submitted their patents to pools, we were able to ensure that the assessors in our experiment would receive no more information than the pools used, and are assessing patents against precisely the same version/release of the standard, etc.

4.3 Selection of assessment cases and associated data collection

To perform the experiment, we developed a sample of cases, where 'case' refers to a combination of a granted patent document and a (specific release of a) standard document (e.g. TS 25.211 V2.5.0). While 'positive' reference cases can be easily identified using public information by pools on which patents were determined to be essential, 'negative' references cases required another approach. Ideally, we want to know which patents were actually submitted to pools, but then rejected. To obtain such information, we sought collaboration with patent holders, and, after negotiations, several patent holders that are participating in pools were willing to share that (private) information with us. Moreover, we also found them willing to share the claim charts that they actually submitted to the pools, for both accepted and rejected patents, allowing us to provide exactly the same information to our assessors as was provided to the pools. The very confidential nature of these claim charts did require non-disclosure agreements to be conducted between all the involved parties (here, it helped that companies have high confidence in the professionalism and confidentiality of the patent examiners that were part of our experiment).

Our experiment focuses on patents essential to the ETSI/3GPP 3G and 4G standards, and uses data from the following pools, which all have a licensing program for these patents: the 'WCDMA' patent pool, the Sisvel LTE/LTE-A patent pool, the Via Licensing LTE patent pool, and the Avanci patent pool. Given that we engaged patent examiners from European patent offices, we only included EPO patents in the experiment. While granted EPO patents always have claims in the English language, the other text in the document may be in any of the EPO's three official languages (English, German and French). Our selected patents also reflected that.

Our final data set has four categories, as shown in Table 1. Category I and II are based on the data provided by patent owners, discussed above. There are two important things to be discussed. Concerning category II. Firstly, because companies usually internally review their patents and only submit patents to pools which they believe likely to succeed, this set is smaller than Category 1. Secondly, because of this preselection, these patents may be relatively more difficult to assess. (We will come back to this later in our analysis and conclusion.) Category III and IV are based on public data, and complement the above data to ensure we have a the required number of patents for our factorial design. Category III are patents that are publicly disclosed by the pool to be essential. We have no claim charts for them, so we use them for the cases where we did not plan to provide claim chart to our assessors. Category IV is the most challenging one, because patent pools do not publish the identity of patents that were submitted but then rejected. We had to reconstruct this category, by creating a set of patents very similar to the one in the actual pool, using a series of defined criteria.[9] By having to reconstruct this set, we acknowledge that the data quality may be lower than that in the other three categories. (As discussed in Section 4.5, we paid special attention to any signals of possible issues with patents from this category.)

Table 1. Data sources

	Patent included in pool	Patent not included in pool
Based on data provided by patent owner	**Category I** - Data sources: pool acceptance information supplied by patent owner (verified by public pool information); claim charts supplied by patent owner - Data quality: very high - Assessment difficulty: average	**Category II** - Data sources: pool rejection information supplied by patent owner; claim charts supplied by patent owner - Data quality: very high - Assessment difficulty: high
Based on public data	**Category III** - Data source: pool inclusion information from pool publication - Data quality: very high - Assessment difficulty: average	**Category IV** - Data source: pool non-inclusion data reconstructed - Data quality: medium - Assessment difficulty: average

In total, the experiment involved 45 unique patents and 48 unique standards documents. Cases were randomly allocated to assessors, while ensuring that a single assessor never received more than one case on the same patent or on the same standard document (to prevent unobserved learning effects).

4.4 Assessors, case allocation, and instructions

In the experiment, patent examiners employed by six different European patent offices participated as assessors. While the management of patent offices themselves was closely involved in study design and operationalisation, none of the ultimate assessors was given any of this information. In total, 20 patent examiners from six different patent offices participated, selected to have considerable expertise related to the technical areas of our cases (ETSI/3GPP 3G and 4G standards). Each assessor received eight different cases, uniquely assigned to this participant and equally distributed in the combination of essential/non-essential and with or without claim charts, yielding 40 cases per combination (see Table 2). Assessors were instructed not to assume any particular distribution of cases. Since each assessment by our experts can be either consistent or inconsistent with the assessment of the pools, we can assign a value of 0/1 to each of them. Thus, each assessment is an independent, identically distributed observation from a Bernoulli distribution, so the final observation in each cell is an observation of a Binomial distribution.

[9] These criteria are as follows: (1) the patent owner is a member of the WCDMA pool, (2) the patent was declared to ETSI as potentially essential for the relevant standards, (3) the ETSI declaration included information on the specific standards documents for which the standard was potentially essential, (4) the ETSI declaration was within a time window in which the declaring firm declared most of its patents that eventually became WCDMA pool patents, (5) the patent itself is not part of WCDMA pool patents nor of an INPADOC family containing other patents that are among WCDMA pool patents, and (6) the patent was applied for at the EPO and granted.

Table 2. Allocation of assessments

		Block		
	Claim chart	**Essential**	**Non-essential**	**Total**
Treatment	**With**	40	40	80
	Without	40	40	80
Total		80	80	160

Each participant received eight different cases, where a case, as indicated above, refers to a combination of a granted patent document and a (specific release of a) standard document (e.g. TS 25.211 V2.5.0). Standards and patents (essential to them) have an *n:m* relationship, and also in our dataset, some cases shared a patent or shared a standard. Yet, to avoid biases from learning, individual assessors considered every patent and every standard only once through the whole experiment. With the above restriction in mind, both the allotment of the cases and the order in which the assessor processed the cases was randomised (and our logistics ensured they were indeed evaluated in that order).

Assessors were provided with an extensive set of instructions, which were developed together with the patent office management departments, and were pre-tested for clarity (see Appendix 1). Among other things, assessors were instructed not to look for any information that was not provided by us so that their assessment was solely based on the patent text and the standard document we provided. They were not allowed to discuss cases with others for the entire duration of the experiment. Also, they were only allowed to look up *technical* information from other sources if such would be necessary to understand the technology described in the patent or standard (e.g. a technical handbook or a standards document in the same 3GPP series). Patent documents were anonymised by removing patent number and assignee information, and assessors were instructed not to look up information on the patent specifically (e.g. by searching on the title). For additional verification, assessors were asked to indicate if they had a suspicion about the identity of the patent owner and/or the patent itself. Finally, after the experiment was completed, all assessors were provided with a feedback and debriefing form.

4.5 Data verification

Before we carried out the data analyses, we verified the assessment data for factors that could potentially have a confounding effect on the experiment. During debriefings, we understood that one group of participants had not respected all the elements in the instructions. While they did so with good intentions, they did not realise this was at odds with our research design, and we had to exclude the associated observations from our quantitative analysis, but we still used their feedback in the qualitative analysis. Furthermore, in a few cases, participants reported they had seen the patent before (possibly as an examiner) and/or informed us they knew (or thought they know) who the patent owner was. These cases were also discarded from the quantitative (but not the qualitative) analysis. Finally, studying the feedback we received from assessors, we identified 19 observations (all from data Category IV) for which assessors reported specific issues and where there may be doubt about the cases or the reference assessment (for instance, a patent might not match the specific standards document provided even though that same set was previously provided to the pool). While we kept these cases in our dataset, we performed an additional analysis excluding these cases, and this analysis did not reveal qualitative difference to our conclusions. Altogether, our final analysis includes 109 valid observations.

5 Results

In this section, we present the results of our experiment. First, we address the accuracy level, second, the impact of claim chart availability on this accuracy, and finally, the differences between the original and the alternative (novelty-based) definition of essentiality. For each of these, we discuss both the quantitative and qualitative outcomes.

5.1 Quantitative findings on overall accuracy of assessments

Table 3 shows the assessment outcomes in the experiment, compared to the reference points used. In 74% of the cases, the outcome of the assessment was consistent with the reference. This percentage is higher for (according to the pools) essential patents (83%) than for the non-essential patents (62%), indicating some difference between the false positives and the false negatives. Note that here we do not yet distinguish differences across assessments with or without a claim chart, neither across the different essentiality definitions.

Table 3: Discrimination between essential and non-essential patents

	Experiment outcome compared to reference point		
Essentiality status according to the reference point	**Consistent**	**Inconsistent**	**Total**
Essential	53 (83%)	11 (17%)	64 (100%)
Non-essential	28 (62%)	17 (38%)	45 (100%)
Total	81 (74%)	28 (26%)	109 (100%)

Note: Cells show the number of observations and percentage of row total.

The first test we perform is whether *participants can differentiate essential from non-essential patents (i.e. consistent with the reference).* To do so, we compare the assessments with a (hypothetical) set of random assessments with a probability of an assessment as 'essential' equal to the share of essential (according to the respective pool) observations in the sample. The chi-squared test of proportions indicates that the assessors are significantly better than random in differentiating essential from non-essential patents ($\chi^2 = 23.32, p = 1.37 \cdot 10^{-6}$). With a share of 74.3% consistent assessments the outcome is far better than random (51.5%).[10] We will show below that this score can be increased further.

5.2 Quantitative findings on impact of claim charts and essentiality definition on accuracy

Table 4 shows the result of the assessments depending on the availability of claim charts (52 of the 109 observations included a claim chart). The percentage of assessments inconsistent with the reference was twice as large without claim charts (33%) than with claim charts (17%).[11] The chi-

[10] A 'random' assessment would yield 'essential' with a probability corresponding to the share of essential (according to the pool) cases in the sample, 64/109. Thus, the expected share of consistent assessments in a 'random' assessment is given by $(64/109)^2 + (45/109)^2 = 0.515$.

[11] Note here, again, that our non-essential cases with claim chart might have been the most difficult to assess, since patent holders would not create a claim chart for a case in which they did not believe the patent to be essential. Thus, for the overall population of patents that might be candidates for an assessment procedure, the difference might be bigger than Table 4 indicates.

squared test of proportions shows that availability of claim charts improves the outcome of the assessment significantly ($\chi^2 = 3.66, p = 0.056$). Importantly, providing claim charts increases the share of consistent assessments to 83%.

Table 4: The effect of the availability of claim charts

	Experiment outcome compared to reference point		
Claim chart availability	**Consistent**	**Inconsistent**	**Total**
No claim chart	38 (67%)	19 (33%)	57 (100%)
Claim chart	43 (83%)	9 (17%)	52 (100%)
Total	81 (74%)	28 (26%)	109 (100%)

Note: Cells show the number of observations and percentage of row total

As mentioned in Section 4.1, most of the assessors applied the ETSI definition of essentiality, while others followed what we called a novelty-based test. So far, the results shown included the data points from both. To check that this is indeed valid, we verify whether the results of the 'ETSI-based' essentiality assessments and those based on the novelty-based tests are comparable.[12] Table 5 shows the results.

Table 5: Novelty-based vs. regular essentiality definition

	Experiment outcome compared to reference point		
Type of essentiality definition	**Consistent**	**Inconsistent**	**Total**
Regular	59 (73%)	22 (27%)	81 (100%)
Novelty-based	22 (79%)	6 (21%)	28 (100%)
Total	81 (74%)	28 (26%)	109 (100%)

Note: Cells show the number of observations and percentage of row total

A chi-squared test of proportions does not indicate a significant difference between the distributions in the rows of Table 5. ($\chi^2 = 0.358, p = 0.55$). If anything, we see that the outcome of novelty-based assessments appears slightly more often consistent with the reference (79%) than the regular assessments (73%). This result has an important implication. Even though most assessors felt qualified to perform the assessments (and they expressed this in their open feedback), patent examiners are not always trained to perform infringement analyses, and infringement partly depends on the respective national law. However, since patent examiners are trained to perform novelty analyses, they can directly perform essentiality assessments under the novelty-based definition. This is relevant given that stakeholders expressed their confidence in the reputation of patent offices as trustworthy, independent third parties that were a qualified candidate to perform these tests on a large scale.

Finally, we look at the combined effects of claim chart availability and different essentiality definitions. Table 6 shows that claim charts seem to improve the degree of consistency of regular assessments considerably (from 63% to 84%), while consistency of novelty-based assessments is

[12] Note that we ran this analysis *before* the others, but only discuss this analysis now for readability reasons.

unchanged (at 79%). However, the chi-squared test of proportions shows that the differences are not significant ($\chi^2 = 1.34, p = 0.25$).

Table 6: Separated results for claim chart availability and novelty-based assessments

		Experiment outcome compared to reference point		
Claim chart provided	**Type of assessment**	**Consistent**	**Inconsistent**	**Total**
No	Regular	27 (63%)	16 (37%)	43 (100%)
No	Novelty-based	11 (79%)	3 (21%)	14 (100%)
Yes	Regular	32 (84%)	6 (16%)	38 (100%)
Yes	Novelty-based	11 (79%)	3 (21%)	14 (100%)
All		**81 (74%)**	**28 (26%)**	**109 (100%)**

Note: Cells show the number of observations and percentage of row total

5.3 False positives vs. false negatives

An intriguing aspect of our experiment is the comparison between false positives and false negatives. As discussed above, the assessment accuracy might differ between presumed essential and presumed non-essential patents. We therefore test whether *there is a difference between false positives (assessors claim essential, reference is not essential) and false negatives (expert claim not essential, reference is essential).* The percentage of non-essential patents assessed as essential (inconsistent with the reference point) is 38%, more than twice the percentage of essential patents assessed as non-essential (17%). We compare this result with an expected outcome that would give the same likelihood to false positives and false negatives. The chi-squared test of proportions indicates that there are indeed significant differences between them ($\chi^2 = 5.87, p = 0.015$).

There are two possible interpretations of this result. First, it might be the case that assessors are more inclined to assess a patent as essential than non-essential, as a form of confirmation bias. Secondly, it might be that our non-essential cases (according to the reference) were 'more difficult' to assess than the essential cases. The non-essential patents with claim charts are, as discussed in Section 4.3, cases for which the patent holders thought they were close enough to being essential that it was worth investing the effort and cost of creating a claim chart. Thus, some of the non-essential cases might be more difficult to assess. Relevant to both explanations is that in their qualitative feedback, assessors did indicate that proving a patent's essentiality typically took less effort than proving non-essentiality, particularly when no claim charts were provided.

5.4 Qualitative findings

We also gather extensive qualitative feedback from the assessors, both for each individual assessment they performed, and after having finalized all their work. They expressed their strong belief that the task given to them required a thorough knowledge of the standard documents. Such knowledge, they felt, could be gained by practice, for example, by specialising in essentiality assessments. For example, in one case the patent referred to a feature that was not available at all in the specific standards document provided to the assessor but may have been elsewhere in the standard (3GPP standards together cover hundreds of separate documents). In another case, an assessor expressed the suspicion that a patent would be essential for a newer release of the relevant standards document, whereas the provided version did not require the use of the patented technology. Assessors indicated that improved searching tools could help a less experienced assessor, especially when the claimed essential features were spread over a combination of standard documents. Moreover, assessors felt they would have benefited from access to additional information about the patent, such as written opinions from patent offices, claim trees, external knowledge, and interaction with stakeholders. Additionally, while the text of the standard (and, where applicable, the claim charts) we provided was always in the English language, and the claims in the patent publication were also always in English, as we have already mentioned we had cases in which the rest of the text in the patent publication was in German or French. In response, some examiners expressed that they felt less confident about their interpretation of the patent coverage. Implementing the above suggestions should help to increase the share of consistent assessments further.

On average, assessors reported spending 7 hours per case – which means than some took less than the eight hours we communicated that they could spend on the assessment. For nine observations, participants reported spending 'much more' time than anticipated, for five observations 'much less' and for other observations only a little more or less than expected.

Participants appeared relatively confident in their evaluations, labelling them as 'very certain' (25 observations) or 'quite certain' (101). In the remaining observations participants felt 'undecided' (16), 'quite uncertain' (12) or 'very uncertain' (6). Participants generally (in qualitative feedback) indicated they felt qualified to perform the assessment, even while the task was new, and a few cases were noted to be outside of the regular field of expertise of the assessor (e.g., at a lower-level technical 'layer'). In 137 cases, participants reported that their skill level increased 'slightly' or more, which may indicate a learning effect could be present (on this note, see (**source anonymized for peer review**) which elaborates on the learning effect observed in the earlier experiment).

Participants indicated (in a closed question form) that the claim chart was 'very helpful' (31 observations) or even 'extremely helpful' (28 observations).[13] In their open feedback, assessors indicated that claim charts were useful for two reasons: claim charts saved them time and made them feel more confident about the outcome. They also commented that absent claim charts, the procedure of reading the patent description, isolating the parts that are truly reflected in the claims, and then doing the same for the standard document and matching both parts took them a lot more effort.

6 Conclusion and discussion

The purpose of this study was to investigates whether essentiality assessments can be made sufficiently efficient (in terms of time and costs) as well as sufficiently accurate, that one could set up a large-scale system of such assessments and thus overcome important inefficiencies in the market for SEP licensing.

In our experiment, where assessors were instructed to spend an average of eight hours per case, we find that overall 74% of the outcomes are consistent with the essentiality assessments of patent pools (which we use as the reference point), and when assessors in the experiment are provided with claim charts as input and use the 'regular' essentiality definition, consistency increases to 84%.

Further improvements should be achievable if our approach is implemented in practice. Firstly, in our experiment, we introduced several limitations to ensure a proper research design. Among other things, our assessors were not allowed to work in teams or exchange information, could not look up patent prosecution history or additional, possibly relevant information. In a practical implementation of the approach, such limitations would not be imposed, likely increasing performance. Second, even though the assessors in our experiment were selected on expertise with the relevant ETSI/3GPP 3G and 4G standards, this is still a relatively broad area. In real life, a larger assessment team could include specialists in relevant subfields (switching, radio protocols, etc.), and

[13] These numbers include qualitative feedback on cases that we had to exclude from the quantitative analysis.

patents could be allocated to assessors according to their key technological competences, improving performance further. Thirdly, our data set included a significant number of relatively 'difficult' cases, such as patents previously submitted to a pool but subsequently rejected. Patent pool assessments are costly, and we must assume that the patent owner had reasonable expectations that the patent would have been essential. In contrast, in a large-scale implementation of our approach, assessments may start from all patents declared as potentially essential to SSOs, implying that there will be many more 'easy' cases. Again, performance should increase. Fourth, we see many learning opportunities, both on an individual basis (progressing experience and knowledge) and in a team setting. Indeed, the assessors participating in our experiment commented that eight cases were not enough to generate learning effects. Finally, a practically implemented system could allow parties – patent owners as well as third parties – to appeal the result of the assessment. This measure, too, should help to make assessments more accurate, though we recognize such procedures need to be designed carefully to avoid the potential for misuse.

Given the outcome of our experiment and the above opportunities to improve performance further, we believe that a large-scale system of essentiality assessments based on our approach can achieve a high degree of accuracy at affordable cost, and thus overcome important inefficiencies in the market for SEP licensing.

Our experiment has several limitations. First, in our assessment, we used ETSI patents and thus the ETSI essentiality definition. Doing such assessments for standards developed by other SSOs could be more challenging, especially if these SSOs have essentiality definitions that include, for instance, commercial essentiality. Second, while ETSI requires parties to disclose which specific patents they believe to be potentially essential, other SSOs, including ITU, IEEE, and ISO/IEC, allow parties to submit 'blanket' declarations that do not indicate specific patents. While a large-scale essentiality test mechanism does not necessarily need to rely on declarations made at SSOs (it may also start by patent owners proposing their patents for assessment), this may limit system design options. Third, the availability of input claim charts, where we observed the highest degree of consistency, will depend on the willingness of patent owners to make such information available – and, in turn, on the incentives patent owners see to do so. As indicated above, such questions concerning 'institutional feasibility' are not in the scope of this paper but are addressed in complementary work presented in (**source anonymized for peer review**). Finally, while we believe that patent pool assessments are a very appropriate reference point for this study, they do not represent an absolute reference point, and such a reference point does not exist. Therefore, our findings are necessarily limited to observing consistency, not accuracy.

There are ample opportunities for future research in this area, especially since the European Commission, in its IP Action Plan of November 2020, announced it will explore the creation of an independent system of third-party essentiality checks in view of improving legal certainty and reducing litigation (European Commission, 2020:13). One of these opportunities is research to understand how AI-based systems, while not replacing human assessors (see Section 3.3), can complement human assessments.

References

Akerlof, G. A., 1970. The market for 'lemons': Quality uncertainty and the market mechanism. Quarterly Journal of Economics 84(3), 488-500.

Aoki, R., Arai, Y. (2018). Strategic Declaration of Standard Essential Patents. RIETI Discussion Paper Series 18-E-035. Retrieved from https://www.rieti.go.jp/jp/publications/dp/18e035.pdf

Arora, A., Ceccagnoli, M., 2006. Patent protection, complementary assets, and firms' incentives for technology licensing. Management Science 52 (2), 293–308.

Arora, A., Fosfuri, A., & Gambardella, A. (2001a). Markets for technology and their implications for corporate strategy. Industrial and corporate change, 10(2), 419-451.

Arora, A., Fosfuri, A., Gambardella, A., 2001b. Markets for Technology: The Economics of Innovation and Corporate Strategy. MIT Press, Cambridge, MA.

Article One Partners. 2011. LTE Standard Essential Patents now and in the future. Retrieved January 30, 2019, from http://newsletters.articleonepartners.com/news_f1317eac-ee13-5a66-d0f5-38ea99a4c1eeLTE-Standard-Essential-Patents-Now-and-in-the-Future.pdf

Arts, S., Cassiman, B. & Gomez, J.C. 2017. Text matching to measure patent similarity. Strategic Management Journal 39 (1), 62-84.

Baumol, W.J., 2010. The Microtheory of Innovative Entrepreneurship. Princeton University Press, Princeton, NJ.

Beggs, A.W., 1992. The licensing of patents under asymmetric information. International Journal of Industrial Organization 10(2), 171-191.

Bekkers, R., Catalini, C., Martinelli, A., Righi, C., Simcoe, T. (2017). Disclosure rules and declared essential patents. NBER Working Paper 23627, available at http://www.nber.org/papers/w23627. National Bureau of Economic Research, Cambridge (MA).

Bekkers, R., Raiteri, E., Martinelli, A., M. Tur, E. (2020) Landscape Study of Potentially Essential Patents Disclosed to ETSI. Nikolaus Thumm (ed.). Publications Office of the European Union, Luxembourg, JRC121411, doi:10.2760/313626, ISBN 978-92-76-20786-3.

Bekkers, R., Updegrove, A. 2013. A study of IPR policies and practices of a representative group of Standards Setting Organizations Worldwide. Washington: National Academies of Science. Retrieved from http://www.nap.edu/catalog.php?record_id=18510

Bekkers, R., West, J. 2009. The limits to IPR standardization policies as evidenced by strategic patenting in UMTS. Telecommunications Policy, 33(1-2), 80-97.

Bessen, J.E., Meurer, M.J., 2008. Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk. Princeton University Press, Princeton, PA.

Brachtendorf, L., Gaessler, F., Harhoff, D. 2020. Truly standard-essential patents? A semantics-based analysis. CEPR Discussion Paper No. DP14726. https://ssrn.com/abstract=3603956.

Bresnahan, T. F., Trajtenberg, M., 1995. General purpose technologies 'Engines of growth'? Journal of econometrics, 65(1), 83-108.

Caves, R. E., Crookell, H., Killing, J. P., 1983. The imperfect market for technology licenses. Oxford Bulletin of Economics and Statistics, 45(3), 249-267.

Charles River Associates (CRA). 2016. Transparency, predictability, and efficiency of SSO-based standardization and SEP licensing. A Report for the European Commission. Retrieved from https://ec.europa.eu/growth/content/study-transparency-predictability-and-efficiency-sso-based-standardization-and-sep-0_en

Contreras, J, 2013. Fixing FRAND: A Pseudo-Pool Approach to Standards-Based Patent Licensing, Antitrust Law Journal 79, pages 47-97.

Contreras, J, 2016. When A Stranger Calls: Standards Outsiders And Unencumbered Patents, Journal of Competition Law and Economics 12(3), pages 507-540.

Contreras, J. (ed). (2007). Standards Development Patent Policy Manual. American Bar Association (ABA), Committee on Technical Standardization, Section of Science & Technology Law. Chicago: ABA Publishing.

Contreras, J., 2017. Essentiality and Standards-Essential Patents, In J. Contreras (Ed.), The Cambridge Handbook of Technical Standardization Law: Competition, Antitrust, and Patents (pp. 209-230). New York: Cambridge University Press. doi:10.1017/9781316416723.016

Contreras, J., 2019. Technical Standards, Standards-Setting Organizations and Intellectual Property: A Survey of the Literature (With an Emphasis on Empirical Approaches), in Menell, P.S. and Schwartz, D.L. (Eds.), Research Handbooks on the Economics of Intellectual Property Law, Vol 2 - Analytical Methods (pp. 185-235). Cheltenham and Northampton: Edward Elgar.

Cyber Creative Institute. 2011. Evaluation of LTE essential patents declared to ETSI. Retrieved January 30, 2019, from http://www.cybersoken.com/file/lte01EN.pdf.

Dewatripont, M., Legros, P. 2013. 'Essential' patents, FRAND royalties and technological standards. Journal of Industrial Economics 61(4), 913-937.

European Commission (2017) Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee: Setting out the EU approach to Standard Essential Patents. COM(2017) 712. https://ec.europa.eu/docsroom/documents/26583.

European Commission (2020). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Making the most of the EU's innovative potential: An intellectual property action plan to support the EU's recovery and resilience. Brussels, 25.11.2020 COM(2020) 760 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0760

European Telecom. Standards Inst. (ETSI), 2020. Rules of Procedure of the European Telecommunications Standards Institute, Version approved by the extraordinary SCM of the General Assembly on 3 September 2020. ETSI' Sofia Antipolis.

Fairfield Resources International. 2007. Analysis of Patents Declared as Essential to GSM as of June 6, 2007. Retrieved January 30, 2019, from http://www.frlicense.com/recent.html

Fairfield Resources International. 2009. Review of Patents Declared as Essential to WCDMA Through December, 2008. Retrieved January 30, 2019, from http://www.frlicense.com/recent.html

Fairfield Resources International. 2010. Review of patents declared as essential to LTE and SAE (4G Wireless Standards) through June 30, 2009. Retrieved January 30, 2019, from http://www.frlicense.com/recent.html

Farrell, J., & Shapiro, C. 2008) How strong are weak patents? American Economic Review, 98(4), 1347–1369.

Gallini, N.T., Wright, B.D., 1990. Technology transfer under asymmetric information. RAND Journal of Economics 21(1), 147-160.

Gambardella, A., Giuri, P., Luzzi, A., 2007. The market for patents in Europe. Research Policy 36, 1163–1183.

Gans, J.S., Hsu, D.H., Stern, S., 2008. The impact of uncertain intellectual property rights on the

Goldstein, L.M., and Kearsey, B. 2004. Technology Patent Licensing. Aspatore Books.

Goodman, D.J., Myers, R.A., 2005. 3GPP cellular standards and patents. IEEE WirelessCom 2005.

Graham, S.J.H., Mowery, D.C., 2003. Intellectual property protection in the U.S. software industry. In: Cohen, W.M., Merrill, S.A. (Eds.), Patents in the Knowledge-Based Economy. National Academies Press, Washington, DC, pp. 219–258.

Henkel, J. 2021. How to license SEPs to promote innovation and entrepreneurship in the IoT. Working paper, TUM School of Management. https://ssrn.com/abstract=3808987.

Henkel, J., Zischka, H. 2019. How many patents are truly valid? Extent, causes, and remedies for latent patent invalidity. European Journal of Law and Economics, 48(2), 195-239.

iRunway. 2012. Patent & landscape analysis of 4G-LTE technology. Retrieved January 30, 2019, from https://www.i-runway.com/images/pdf/iRunway - Patent & Landscape Analysis of 4G-LTE.pdf

Jaffe, A. B., & Lerner, J. (2006). Innovation and its discontents. NBER Innovation Policy and the Economy, 6(1), 27–65.

Jefferies & Company. 2011. Research in Motion (RIMM) Limited patent value; cut target to salvage value of $21 as we wait for QNX. Retrieved January 30, 2019, from https://ipcloseup.files.wordpress.com/2011/10/rimm.pdf

Lamoreaux, N.R., Sokoloff, K.L., 1999. Inventive Activity and the Market for Technology in the United States, 1840–1920. NBER Working Paper 7107. National Bureau of Economics Research, Inc., Cambridge, MA.

Layne-Farrar, A., Lerner, J. 2011. To Join or Not to Join: Examining Patent Pool Participation and Rent Sharing Rules. International Journal of Industrial Organization 29(2) 294–303. doi:10.1016/j.ijindorg.2010.08.006

Lemley, M.A., Shapiro, C., 2007. Patent holdup and royalty stacking. Texas Law Review 85, 1991–2048.

Lemley, M.A., Simcoe, T. 2019. How Essential Are Standard-Essential Patents? Cornell Law Review 104(3): 607-642.

Lerner, J. (1994). The importance of patent scope: An empirical analysis. RAND J. Econ. 25 (2), 319–333.

Macho-Stadler, I., Martinez-Giralt, X., Perez-Castrillo, J.D., 1996. The role of information in licensing contract design. Research Policy 25(1), 43-57.

Madiès, T., Guellec, D., Prager, J. C. (eds.), 2014. Patent Markets in the Global Knowledge Economy: Theory, Empirics and Public Policy Implications. Cambridge University Press.

Magliocca, G., 2007. Blackberries and barnyards: patent trolls and the perils of innovation. Notre Dame Law Review 82 (5), 1809–1838.

Marco, A.C., Sarnoff, J.D., de Grazia, C.A.W. (2019). Patent claims and patent scope. Research Policy 48(9). https://doi.org/10.1016/j.respol.2019.04.014

Martin, D. L., De Meyer, C. 2006. Patent counting, a misleading index of patent value: A critique of Goodman & Myers and its uses. Available at SSRN: https://ssrn.com/abstract=949439

Merges, R.P., Mattioli, M. (2017). Measuring the Costs and Benefits of Patent Pools. Ohio State Law Journal 78(2), 282-347.

Milgrom, P., Roberts, J., 1992. Economics, Organization and Management. Englewood Cliffs, NJ: Prentice-Hall.

Miller, S. P. (2013). Where's the innovation? An analysis of the quantity and qualities of anticipated and obvious patents. Virginia Journal of Law and Technology, 18(01), 1–58.

Natterer, M. 2016. Ähnlichkeit von Patenten: Entwicklung, empirische Validierung und ökonomische Anwendung eines textbasierten Ähnlichkeitsmaßes. Verlag für Nationalökonomie, Management und Politikberatung.

PA Consulting Group. 2015. LTE Essential IPR. PA's 3GPP-LTE Database and Report.

Reitzig, M., Henkel, J., Heath, C.H., 2007. On sharks, trolls, and their patent prey – unrealistic damage awards and firms' strategies of 'being infringed'. Research Policy 36 (1), 134–154.

Serrano, C.J., 2010. The dynamics of the transfer and renewal of patents. RAND Journal of Economics 41 (4), 686–708.

Siebrasse, N. V., Cotter, T. F. (2017). Judicially determined FRAND royalties. In Contreras, J.L. (ed.), The Cambridge Handbook of Technical Standardization Law: Competition, Antitrust, and Patents (pp. 365-388). Cambridge University Press. https://doi.org/10.1017/9781316416723.026

Uijl, S. D., Bekkers, R., & de Vries, H. J. D. (2013). Managing Intellectual Property Using Patent Pools: Lessons from Three Generations of Pools in the Optical Disc Industry. California Management Review, 55(4), 31-50. doi:10.1525/cmr.2013.55.4.31

Younge, K.A., Kuhn, J.M. 2016. Patent-to-patent similarity: A vector space model. Social Science Research Network. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2709238

Appendix 1: Instructions for assessors

The below instruction was provided to assessment that were doing regular (ETSI-definition based) essentiality tests.

(source anonymized for peer review**).**