ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 709 April 2012

Patent Pendency, Learning Effects, and Innovation Importance at the US Patent Office

Samrawit Mariam, Pierre Regibeau and Katharine Rockett

Patent Pendency, Learning Effects, and Innovation Importance at the US Patent Office

Samrawit Mariam*, Pierre Regibeau** and Katharine Rockett***

April, 2012

Abstract

We replicate and extend the results of Regibeau and Rockett (2010) on a new data set. We confirm that importance of a patent and the delay in approval at the patent office are negatively related. This relation survives even if we do not control for learning effects and so suggests that carefully defining the technology is sufficient to recover the negative relation. We use new measures to test for the existence of patent "thickets", thereby ruling out some strategic considerations in delay behaviour. It appears that delay is attributable to patent office, not filer, behaviour in our sample. A more careful analysis of the possible effect of examiner workload finds that larger workloads per examiner are associated with shorter approval time, lending credence to Lemley and Shapiro's concern that heavy workloads force examiners to devote too little time to each patent review.

JEL Codes: L5, O3

Keywords: Patent Office, Examiner Delay, Patent Thicket, Survival Analysis

* CRA International

** CRA International and Imperial College, London

*** University of Essex and CEPR. Corresponding author at kerock@essex.ac.uk

We thank Diana Terrazas for excellent research assistance. The views expressed here reflect our opinions only and do not reflect those of our organisations.

Introduction

A growing body of work has investigated the workings of patent and trademark offices (PTOs) in a variety of countries. Concern about the quality of patent office reviews has been particularly strong in the US, (FTC, 2003, NAS, 2004, Jaffe and Lerner, 2004). Aspects that have been studied include “rational ignorance” (Lemley, 2001), examiner incentives to provide high quality patent review (Schuett, 2011), disclosure of prior art (Langinier and Marcoul, 2007), patent overload (Caillaud and Duchene,2011), patent re-examination (Graham and Mowery, 2004), and potential biases in the review process (Johnson and Popp, 2003).

The last work, by Johnson and Popp, provides empirical evidence that more important patents are reviewed systematically more slowly by the US PTO. This result was questioned in recent work by Regibeau and Rockett (2010), who pointed out that when broad technology groupings are used in econometric work, innovations at different points in their “technology cycles” are combined together. If major innovations occur disproportionately at the beginning of such cycles, then they occur at a point where patent examiners may be learning about how to evaluate patents in such a novel area. Hence, importance and learning effects on delay can be confounded. They argue that when learning effects are significant this could lead to a positive but spurious empirical correlation between patent importance and pendency.

This paper presents a similar but extended analysis to confirm and elaborate on the Regibeau and Rockett results. Specifically, Regibeau and Rockett (2010) analysed a single and somewhat restricted data set. Confirmation on an independent data set that their effects remain would increase the confidence one could put in their results. That paper also did not test that data for the presence of potential strategic effects outside their theoretical model and, more precisely, it did not test the data for the presence of patent thicket effects. Strategic effects arise and can affect the results because a patent that is part of a thicket may generate very large rents even if it is not “important”. For example, a “defensive” patent is a useful bargaining tool to extract profits from infringers if it merely blocks other patents. Such strategic interactions among patent applicants can affect patent submission strategies and grant times, affecting the correlation between any single patent’s importance and its filing behaviour. New patent thicket measures have become available recently, namely a measure by Clarkson (2004), which this paper applies to its data set before conducting empirical analysis.

This paper finds that its data set passes these tests so that we reject the presence of a distinct patent thicket within our data. We then move on to an empirical analysis to confirm the relation between the importance of a new innovation and its pendency at the patent office. While confirming the main result of Regibeau and Rockett (2010) that, contrary to Johnson and Popp (2003) and consistent with Harhoff and Wagner (2009), more important patents are approved more quickly than less important patents the results in this paper are distinct in a number of ways. The data set exhibits a much more muted learning effect over the technology cycle than the data of Regibeau and Rockett (2010). Indeed, while the broad technology class to which our patents belong exhibits a positive relation between patent pendency and importance in the Johnson and Popp analysis, merely by specifying a narrower technology this paper is able to recover a negative relation between pendency and importance. The smaller technology learning effect is not surprising given the nature of the technology involved, in fact: our earlier data set involved a radically new technology (genetic modification), whereas this data set involves a much more modest step (stent development). Hence, the source of the basic sign of the relation is quite different in this paper: here, the negative relation can be recovered merely by defining a technology narrowly enough that it represents a single technological innovation. Separating out the learning effect once the technology has been isolated is not crucial to obtaining the correct sign, as it was in Regibeau and Rockett (2010). Indeed, Harhoff and Wagner's (2009) model does not separate out learning effects but does specify technologies more precisely than Johnson and Popp (2003) and achieves results linking importance and delay that are consistent with ours.

On the other hand, we find significant individual examiner effects in our data, which increase the magnitude of our main relation between importance and pendency. As our data is drawn from a narrow technology, we can compare our results to other studies that used much broader technology definitions to confirm their results with a stricter control for the type of patent reviewed. The issue of individual examiner effects was pointed out in Cockburn, Kortum, and Stern (2003), was present in more muted form in Regibeau and Rockett (2010) and was excluded from Johnson and Popp (2003). Indeed, as a rather efficient examiner handles a large proportion of our patents early in the data set, without this control we could find that examination times lengthened over the period of our data. Rising workloads of examiners and their effect on patent pendency has been discussed extensively, of course, in practice by Lemley and Shapiro (2005), FTC(2003), and NAS(2004) and in theory by Caillaud and Duchene (2011). While average workloads rise over the course of time in our data, we find that the effect of a change in workload at the level of individual examiners and within a technology are – perhaps surprisingly – reversed: lower workloads are associated with longer pendency times. This is consistent with Lemley and Shapiro's (2005) informal concern that heavy workloads force examiners to devote very little time to each patent review. These authors postulate the increased workloads at the patent office may have led to lower quality examining. While we have no evidence of a change in quality within our sample, our findings are consistent with their view that increased workloads may put pressure on examiners to spend less time on each patent examination. Our results are qualified, however, since they are not completely consistent across the sample. In keeping with Cockburn et al (2003), we find that examiner experience does not generally affect examination delay significantly, and when it does it has inconsistent effects.

Part I of the paper discusses the link between patent pendency and patent importance, with a focus on separating out the effect of importance itself rather than other effects that could be confounded with importance. Queries that will be addressed by this paper are the focus of this section. Part II reviews recently proposed measures of patent thickets and the results of these measures when applied to our data. Part III presents our empirical results for the relation between patent pendency, patent importance, and the stage in the innovation cycle. Part IV concludes.

I: Patent Pendency and Patent Importance

We focus on the relation between the approval delay involved in patent office review and the quality of the underlying innovation. While Johnson and Popp, 2003, presented empirical results suggesting that more important patents are reviewed systematically more slowly, contrasting empirical work was presented by Harhoff and Wagner (2009) and Regibeau and Rockett (2010). Regibeau and Rockett's model suggests that the delay observed for more important innovation is composed of two effects. First, delays rise when innovations are early in their "innovation cycle", in other words, when innovations are quite novel. For example, early in the years of genetic engineering, patents based on this new technology were uncharted territory for examiners and so would have required more review time. Once the technology was established and examiners had experience in such reviews, review time would fall as the work became more routine with better known markers for patent quality. If important innovations, in the sense of being highly cited, tend to occur early in innovation cycles, they will tend to have a disproportionate learning effect and so a disproportionate delay that is not directly due to importance. Second, delays rise when applicants or examiners are less diligent in the "give and take" that is involved in patent approval. This process involves queries raised by examiners and responses made by applicants (or vice versa). Similar to the refereeing process for academic papers, this process can be undertaken with more or less effort on both sides. If more important patents are those that generate more social welfare or private profits, then the effort incentive for the applicant and/or the patent office will tend to be greater, leading to quicker approval times. The paper tests these implications on a data set that traces a single and focussed technology (GM food) from inception up to 1999 and finds support for the result that higher delays tend to be due to a learning effect while lower delays tend to be due to importance *per se.* The importance of a patent is measured by its own forward citations, which is one of the standard measures used in the literature (see Lanjouw and Schankerman, 2004 for discussion).

Specifically, Rockett and Regibeau (2010) and also this paper test the model:

Approval delay = f(importance, location in innovation cycle, control variables)

Where approval delay is measured as the difference between grant date and application date for a patent, the location in the innovation cycle is measured either by a time trend or by yearly dummy variables, and controls include the number of claims in the patent as a measure of patent review complexity, examiner and attorney dummies to reflect the effects of differences in practice across individuals as identified in Cockburn et al (2003), and private firm dummies to reflect differences in behaviour across private firms and other types of organisations that patent (such as universities). Patent application complexity is important to include in such a specification as the complexity can clearly affect review time. Further, if more important patents also tend to have more complex applications, then the effects of importance and complexity can be confounded. The log of delay was used in initial OLS runs to correct for skewness of the distribution of delay, as noted in Caballero and Jaffe (1993). We report Cox proportional hazard (survival) model results, although a Weibull specification generated similar results. Finally, we run a second specification with formal controls for workload: elapsed time since the appearance of the new technology captures the effect of "learning the technology", examiner cumulative experience in reviewing patents captures the effect of "learning to be an examiner", and overall yearly workload to captures congestion effects.

The Regibeau and Rockett test is limited in several ways that are addressed here. First it is based data from a single technology. This is the result of the need to carefully specify what is meant by an "innovation cycle": even if a carefully constructed data set such as the Trajtenberg CT Scanner patents is used, this data combines relatively well-established technologies (such as those relating to the design of a CT scan examination table) with those of relatively new technologies (such as certain tracer drugs) so that the learning effect is difficult to isolate. Indeed, even relatively precise identification of relevant data with patent classes tend to result in data sets that involve a variety of technologies at different points in their innovation cycle. As a result, a time-consuming method of reading patent documents was adopted to isolate a single "technology cycle" involving genetic modification as applied to food. This resulted in rather clean data, but clearly was not suited to a large data set. Hence, a concern is whether the Regibeau and Rockett (2010) results are specific to their data set. This paper presents results based on a second, carefully collected, data set and confirms the earlier results, raising the confidence one can place in their conclusions.

The difference in the test goes beyond the specific data sets analysed, however. First, the effects of examiner workload and ability are particularly important to control for in the data set of this paper. While examiner workload and ability were issues in Regibeau and Rockett (2010), workload did not vary significantly for that data set and examiner ability did not interact significantly with the year of review. For this data set we have several differences. First, the workload of the examiners involved rose remarkably in many cases, sometimes as much as ten times in the course of a few years. Second, the average speed of review differs significantly across examiners and the relative proportion of our data set handled by these quicker or slower examiners also differs by year. Hence, controlling for the elapsed time since the beginning of the technology cycle (to allow for "learning the technology") and examiner career effects (to allow for "learning to be an examiner", reflected as in Cockburn et al (2003) by cumulative patents handled by the examiner), and overall yearly workload (to allow for congestion effects) is important to capture the various types of learning.

Third, recent work has indicated that where patent thickets are present, strategic effects may affect behaviour. Indeed, "defensive" patenting for the sole purpose of blocking other patents has been discussed extensively in the literature (Cohen et al, 2000). For our purposes, a patent that is part of a thicket may generate very large rents to its owner even if in itself it is not "important": it merely needs to block other patents owned by other firms in order to be a useful bargaining tool to extract profits from devices that infringe the patents. For example, a single profit-generating device may involve a large number of complementary patents. Strategic behaviour in filing patents that are to be used as bargaining chips rather than to generate profits based on their own importance is not part of Regibeau and Rockett's model: profit streams in their model are generated by single patents and not by those patents used in combination. If patents generate profit only when used in combination across patent-holders then the link between profitability and importance of a single patent can be broken: a patent may generate high profits because it blocks another important patent, not because its own measure of importance is high. It is, then, necessary to check that the patent data set used to test their model does not suffer from thicket issues in order to interpret a negative link between importance and pendency as maximising welfare, as do Regibeau and Rockett(2010).

More precisely, Clarkson (2004) suggests that standard network density measures can be adapted to test for patent thickets. Summarising Clarkson's presentation, a standard network density equation for a directed network with g nodes is the total number of ties (or linkages) in a network divided by the total number of possible ties. For example, if x_{ij} is the value of a tie between nodes *I* and *j*, we would have density Δ defined as:

$$\Delta = \frac{\sum_{i=1}^{g}\sum_{j=1}^{g} x_{ij}}{g(g-1)}$$

Applying this to a network of patents, we could have *g* patents, and we could measure linkages or ties as citations or references of other patents (equivalent to "outdegrees" in a network model). An assumption of the standard density equations, however, is that each node can be linked to each other node. As patents are ranked chronologically, with the potential citations pool of each patent determined by the relevant population that is available at the time of citing, this is not the case. Hence, Clarkson proceeds to modify the standard density measure to account for this change in potential citations pool over time. For example, later patents cannot be referenced by earlier patents.

List the patents in the network chronologically, then, so that patent (or "node") *n* can reference the *n-1* patents earlier in the network. The local patent density for each patent, *n*, summed up over all g patents is the number of linkages of that patent divided by its relevant total possible number of linkages. If we wish the average density, then we can divide this total by *g-1*. Hence, Clarkson defines average density measures for either outdegrees or indegrees as:

$$\Delta_p = \frac{\sum_{n=2}^{g}\sum_{j=1}^{g} \frac{x_{nj}}{n-1}}{g-1}$$

Where the difference between outdegrees and indegrees is whether the linkages measure references cited or citations received. Notice that the oldest patent's local density is discarded for the formula when outdegrees are measured (as no references within the network could be made) and the youngest patent's local density is discarded when indegrees are measured.

When testing the properties of this measure on incompletely connected networks, Clarkson notes that weighting each local density by the possible number of citations has the advantage of returning sensible results for complete networks and also means that the measure is not affected by citation placement (so that the total number of citations and not their distribution is what matters to the measure). He therefore modifies the measure one last time to generate:

$$\Delta_{pfinal} = \frac{\sum_{n=1}^{g}\sum_{j=1}^{g} x_{nj}}{g(g-1)/2}$$

Notice that the weighting applied cancels out the denominator *n-1* in this expression and the denominator regains its similarity to the original network density expression, above, but the potential number of links is modified to be only those in a directed network. This is his final measure. He suggests that the measure be calculated both for the patent network of interest and for a near universe of patents so that a "thicket", if it exists, can be detached from its surrounding universe and identified as a separate thicket. The purpose of this is to identify not just classes of patents that may be interlinked, but membership in the potential thicket. Where the patent network density is higher than that of the near universe, a thicket can be "detached" from the surrounding universe and analysed as a separate pool. Clarkson then goes on to test his measures on two thickets that were identified by means of third part evidence or litigation. His measure confirms that both of the pools did, in fact, involve a thicket.

While an advantage of Clarkson's approach compared to either that of the Von Graevenitz et al (2011) or the Cohen et al (2000) alternative measures is that it proposes an objective "standard" by which his patent thicket measure can be judged, a difficulty is that the comparison group may not have the same technological spread as the original data set that is being tested for thickets. Clearly, the more narrowly defined a particular technology, the more linkages will occur as "background" simply due to the requirement to cite technological antecedents. A higher measure on a more narrowly defined technology may not, then, indicate a thicket problem. If we define the comparison group for our relatively narrow patent data set to be the technology class to which it belongs, following Clarkson's methodology, we compare a narrow patent data set (stents) to a broad class (prosthetics, roughly). We are likely to find a significant difference in these measures simply because the comparison group is not another data set that follows a single narrow technology from inception to the current day, with our data showing a higher thicket measure. Hence, a difficulty with the Clarkson methodology is that it can return extremely small figures and still represent a statistically detachable patent thicket compared to its near universe. For example, in Clarkson (2004), a patent pool that has been externally verified as presenting a thicket is identified with a density measure of .03. For a measure that varies between zero and one, this low measure as an indication of a potential thicket "problem" is in itself problematic. Hence, if we use a broad patent class as our comparison group, we note that this means that the Clarkson thicket measure may be quite biased towards returning a thicket where none exists.

As Clarkson himself points out, patents linked by citations do not necessarily generate thicket problems: while his measure may identify a thicket in a narrow sense of a technical linkage, it may not identify a thicket that matters in any sense of creating a bargaining problem. Indeed, if the linkages are sparse enough, parties should be able to bargain around those linkages in many cases. Hence, the bargaining problem that underlies a patent thicket problem requires a dearth of choice in partnering that is not well captured by ignoring the absolute magnitude of the density or the parties that are linked. Hence, the measure also is noisy as a reflection of thickets. For our purposes, we are concerned with whether representing our overall data set with a model of non-strategic patenting behaviour is a good reflection of the "average" behaviour of our data. If it is either the case that a low density measure reflects patents are sparsely connected (so that bargaining or other strategic problems are unlikely to be bad in an overall sense), or just a few patents are very highly linked (so that strategic behaviour affects a small part of our sample), our approach for the data set as a whole will be valid. Indeed, our preferred interpretation of this density measure is that the larger the magnitude of the statistic, the higher the probability that a thicket and its associated strategic behaviour would arise. This is similar to the way we would interpret a high concentration ratio in an industry: the higher it is, the more we would be vigilant to problems, but problems need not necessarily arise[1].

As a way of generating a more appropriate comparison group, one possibility is to compare our density figures with the density of the narrowly defined data sets of Clarkson's paper. These two data sets – the MPEG-3 and PRK data sets – are associated with patent thickets that have been identified externally by legal evaluation as thickets. Clarkson's adjusted density measures for these data sets are, respectively, 0.029 and 0.203.

These figures differ by an order of magnitude of ten, so more investigation of which is the more appropriate figure is in order. First, we could consider the technologies involved to pick the data set that is closer to our technology class. MPEG-3 involves (roughly speaking) electronics while PRK involves a medical technology. In this sense, the latter might be more appropriate. Still, neither is very close to prosthetics as a technology class. We can add to this, however, Newberg's (2001) detailed outline of the difference between these two sets of patents and the thicket issues each presents. The first is a set of patents that present thicket issues because of complementarity. In other words, the technology they relate to requires a large set of complementary patents in order to

[1] As a final note, it is not clear that Clarkson uses net citations in his calculations. Citations of a firm to its own patents should not create a problem of hold up, as we would associate with thickets nor would it create strategic issues between firms. Hence, the magnitude of his measures -- as he calculates them -- should overstate the thicket problem. On the other hand, the control data set in his work also is not based on net citations, it seems, to it is not clear that the difference between the two is affected by the use of all citations.

be able to perform a function that is valued by consumers. Indeed, the MPEG-3 pool was aimed at creating a set of standard essential patents covering a wide set of technologies. Hence, the patents underlying this data set are not necessarily technically related, but they are related legally. This would not necessarily generate citation linkages between the underlying patents, however, so that Clarkson's citations-based measure is probably a poor reflection of the underlying strategic issues linking these patents. Hence, it is perhaps not surprising that this identified thicket returns such a low figure by Clarkson's index. The second – those related to the PRK patent pool -- refer to a set of patents that by and large are related by substitutability (of two methods of performing a medical procedure). The rationale for the PRK patent pool was, according to Newberg, primarily to resolve infringement issues among these patents so that the procedures could be used flexibly by practitioners. This is much more likely to be an issue that will be well-reflected by patent citation patterns. Accordingly, this thicket returns a much higher figure by Clarkson's methodology: his method is capturing a larger proportion of the thicket issues in the second case.

Our data set does not concern a complex technology, so that any strategic issues would be expected to be infringement issues rather than standards or other complementarity issues. As such, patent citations would be a relatively good measure of strategic issues among patents. As our data also relates to medical technology, we take the PRK database as the more appropriate comparison if we want a narrow technology group. We also compare our dataset against the broad technology class within which most of its citations fall in order to follow Clarkson's methodology.

In sum, we implement both Clarkson's unadjusted and adjusted measures as he calculates them for the purposes of this paper; however, we note that a very small figure for the density measure would give us less confidence in the existence of a thicket than a higher figure. In other words, even if the measure is statistically different from zero, or from the background measure, we would view the magnitude of the measure as important in gauging the threat that a set of patent linkages forms. As our magnitude should reflect the full extent of the patent thicket, similar to the PRK case, we would compare our figures in the first instance to the network density in that case in order to judge our data by a similarly narrow technology group. A low reading in comparison to the PRK case or absolutely would suggest that we should proceed with the non-strategic interpretation of our overall data set.

Our near universe for the Clarkson methodology uses the technology classes of the bulk of the stent data, which is primary US classification 623. In some cases class 623 only enters as a secondary classification, however. We then take a near universe as patents with technology class 623 either as a primary or a secondary designation. Most patents within this sample refer to prosthetics of some type. We eliminate patents that occur within our original data set. As performing the Clarkson calculation can be onerous for large data sets, we use a random sample of our near universe and assume this random sample reflects the characteristics of the sample as a whole.

In addition to Clarkson's statistic, we also verify Cohen et al (2000)'s intuitive classification of technology classes into discrete and complex technologies and von Graevenitz et al's (2011) empirical implementation on EPO data for the applicable broad technology class. These three measures get at different aspects of patent thickets: Clarkson's measures reflects the potential for patent interference of the technology for this particular data set only; von Graevenitz et al's results reflect the same potential for a broader technology class containing our technology; Cohen et al's reflects of how much conflict and interference actually is felt to occur for the technology class that contains our data.

II: Data Set

We gathered data on 887 US patents related to medical stents granted between 1976 and 1999. The starting year reflected the first appearance of medical stent patents at the US PTO and the end date was chosen to minimise truncation bias and to minimise the effects of changes in the patent system itself introduced under the American Inventors Protection Act. Also, after 1999 technological changes in patent examination became prominent, affecting review times (see Chin, 2009). Only those patents granted prior to 1996 were included in the final study to minimise any selection bias that could result from patents at the end of the data set being precisely those that were approved the most quickly. This leaves us with a final sample of 658 patents. As most patents in the data set have approval delays of less than three years, a three year window was viewed as appropriate to minimise selection effects while maintaining as large a data set as possible. Finally, as our latest patents had a fifteen year citation history, virtually no citations were still accumulating for our latest patents. Hence, we do not truncate our citations, as was done in Regibeau and Rockett (2010) to "level the playing field" across time: we assume we have virtually the entire citation distribution for all our patents.

The data set involves what is probably best phrased as a technology application or an incremental technology. This distinguishes our data set from our earlier data on GM food. The earlier data involved not only a radically different methodology from earlier methods but also a technology where patentability itself was in question. Indeed, the questions involving how patents should be written for genetic modification applications were not resolved fully until after the final date of our earlier data set. The data set underlying this paper involves arterial stents. This technology is conceptually closer to its precedents in the broadly defined field of prosthetics than genetic modification. While many new issues arise in our stent patents, they are those that an examiner would be familiar enough with that wrestling with how to write the patent and whether the material was patentable in the first place should not have involved any fundamental challenges for the examiners. Hence, we expect a much more muted learning effect in this data set compared to our earlier work.

Our comparison data set for the purposes of calculating our thicket measure is a random sample drawn from patents with 623 as one of the patent classes. This sample of 154 patents mainly consists of prosthesis patents, although some other categories are present as well. This sample was chosen as it captures the patent categories that both our stent data set and the set of patents that they cite falls into. We also looked at measures where 623 was the primary patent class, although the results did not change greatly.

III: Results

Referring to Cohen et al (2000) classification, medical equipment is probably the most relevant category for our data. Cohen et al classify this as a complex technology, using their system of dividing their sample at ISIC code 2900 into discrete (less than 2900) or complex (2900 or above). They acknowledge, however, that this system undoubtedly glosses over some heterogeneity within groups. Von Graevenitz, Wagner, and Harhoff (2011) find, using EPO data and a complexity classification measure adapted to EPO procedures, that medical equipment has a rather low complexity score comparable to other ISIC codes classified as discrete by Cohen et al (2000). While Von Graevenitz et al have no absolute cut-off of when their measure indicates the likelihood of a problematic patent thicket and when it doesn't, similar to our discussion they seem to assume that higher values of their measure raise the likelihood that thickets and their associated strategic behaviour will arise. Hence, these two methods differ on whether the medical equipment class is or is not complex so that we are left unsure from these measures whether our technology should be considered complex – and hence subject to thickets – or not. We turn to Clarkson to resolve this.

All medical equipment is a much larger technological classification than our data set, which is restricted to a single and rather narrow technological application. Implementing Clarkson's unadjusted network density measure to this data, we find that our data set exhibits a measure of 0.0154 (sd 0.0306) if we use outdegrees (backwards citations) and 0.0082 (sd 0.0173) if we use indegrees (forward citations). Both of these measures are extremely small, and while they are not statistically different from zero, they are significantly different from the background measures of

0.0013(sd 0.0045) in outdegrees and 0.0007 (sd 0.0039) in indegrees, using a standard t-test. Given that the density measures have a range of 0 to 1, these figures are unlikely to be associated with any effect on strategic behaviour. By this measure, this subset of medical technology is not a prime candidate to suffer significant strategic patenting problems.

When we use the adjusted network density measure, while the control group's figure changes relatively little, rising to 0.0025, the figure for the stent data set rises considerably to 0.0247. While this figure still is very small, and likely to create no strategic effects if one looks at the absolute magnitude, the rise in the figure is of concern. The change is due to a large effect of controlling for the citing population on our narrow data set since it follows a single technology from birth – where the citing population starts at zero and rises greatly) – compared to little effect on the control data set (where the change in population belonging to this technology class is not large over the measurement period). The citing population, then, changes drastically for our data set and very little for the control group. As a result, it is hardly surprising that the adjusted measure rises much more for our data set than for the control. On the other hand, it points out that our control is not really the same type of data set as our original: it does not follow a single technology from inception. The narrowness of our technology compared to our control group may be making it appear that we have a problematic thicket where one does not exist. While we are following Clarkson's methodology, we are not necessarily following the right methodology.

Comparing our measure to Clarkson's PRK measure, a similarly narrow technology, we see that ours (0.025) is about a tenth of the PRK figure (0.203). Comparing it to the PRK case, then, the measure for our data set is considerably (ten times) smaller.

Taking these together, we tentatively conclude on the basis of the absolute magnitude of our measure and on the basis of a comparison with a similarly narrow patent data set that our data is no likely to present strategic patenting behaviour as a result of thicket considerations. We note, however, that the thicket measures clearly need more work: patent citations are not necessarily capturing important thicket issues and are noisy measures of strategic behaviour in any case. We leave a full analysis of thicket measures to future work and now move on to the rest of our analysis.

Importance and cumulative age of the technology are negatively and significantly related in our sample[2] while complexity of the patent application rises with the cumulative age. We observe, however, that combining these two opposing effects of rising potential learning with rising potential

[2] Relating importance to age of the technology, we obtain a coefficient of -0.0153***, while relating complexity to age of the technology we obtain a coefficient of 0.00062**.

complexity of the task, patent examiner delay does not fall over time significantly in simple tabulations[3].

With this as background, we can now replicate the earlier results of Regibeau and Rockett (2010) on the Stent dataset. We present our results in tables 1 and 2, in equations controlling for examiner and attorney effects but not reporting these coefficients[4]. Several features stand out.

Tables 1 and 2 about here

First, our main result that importance of a patent, as measured by citations, is negatively related to patent pendency is supported by this data. Second, claims enter into pendency with significant and positive coefficients, reflecting the effect on pendency of the complexity of the task facing the examiner. Both of these results confirm those of Regibeau and Rockett (2010).

The effect of the evolution of the technology is insignificant on the truncated data set, suggesting that learning effects are small; however the technology is not a conceptually large step from earlier work so this is not surprising. Indeed, importance and delay are negatively related even if the data is taken in cross section[5], confirming that when learning is modest defining the technology narrowly is sufficient to recover the negative relation between importance and delay. We note that Harhoff and Wagner (2009) present a model where they define technology more narrowly than Johnson and Popp (2003) but where learning effects are not a focus. They recover the negative relation between importance and delay, as we do. Our results suggest that their approach can be valid if learning effects are not too large.

The effect of applicant type also is insignificant in our data, as private firms do not behave significantly differently from other types of applicant. Hence, while Regibeau and Rockett postulated that patent filer behaviour, such as eagerness and promptness to satisfy examiner requests, could be the driving force behind their importance-patent delay relation, and that those incentives would be stronger for private firms, no such effects appear to be present in this data set despite the presence in the earlier analysis. Hence, our results are perhaps more consistent with the

[3] Relating examiner delay to age of the technology, we obtain an insignificant coefficient of 0.00003.
[4] Full results are available from the authors. While some examiner and attorney coefficients are significant – and as a group they are significant – they do not present much interest in themselves.
[5] Not surprisingly, the coefficient on the cross section is barely different from the reported results at -0.000482*** when claims are controlled for.

model variant where delays are attributable to patent office -- rather than patent filer – behaviour for this data. Lacking data in the US on examiner and applicant delays in the examination process, this is probably the closest we can come to determining empirically which side of the negotiations is driving the pattern of behaviour we observe. Indeed, Palangkaraya et al (2008) find evidence of strategic delay by applicants in an international patent matched sample, but takes applicant delay at the US patent office as zero. Given that we analyse US data, it is perhaps unsurprising that we find most delay attributable to patent office behaviour, then.

Moving on to the examiner and attorney dummies, while these are not reported for space reasons a large number of the dummies are significant. Indeed, the effects of individual examiners and attorneys is of higher magnitude than in our earlier data set. Including the examiner and attorney effects in the regression affects the magnitude of the coefficients, but only modestly however.

We report the results of a workload analysis in table three.

Table 3 about here

Focussing on individual examiner workload effects and individual examiner learning about the "job of being an examiner", we find inconsistent effects of congestion effects within our data set -- and hence, controlling for the technology reviewed. These effects are found controlling for the age of the technology, and hence the familiarity of the underlying concepts. For the majority of examiners, the effect of congestion is negative (although they are not always significant): the higher the workload, the less time the examiner puts into each patent, all else equal. This is consistent the concern of Lemley and Shapiro (2005) that increased workloads might result in less time (and in their view, perhaps sloppier) reviews. Our support for their contention is qualified, however. First, we do not have any evidence that reviews decrease in quality in this data set, however, as there is no legal case history related to this dataset to date. Second, the effect is not consistent: for one examiner, a a heavier workload increases time spent per patent significantly, and for some others it has no significant effect. It may be, then, that examiners react in different ways to increased demands from the job and that this concern is not particularly serious across all examiners. Finally, we note that our data is drawn from a time when technological change in the process of examining had not really taken effect so that increased proficiency learning the new computerised system is probably not at the heart of our results.

The effect of experience in the job also is mixed. We take only consider the effects of learning within the time an examiner served as primary examiner. This usually occurs after a significant period of serving as secondary examiner, so that one would expect that most learning already would have occurred for our examiners. By and large this is true, with few significant coefficients and with a minimal effect even when the coefficient is significant. Most significant coefficients are negative as one would expect for an examiner "learning the job", but for one examiner the effect is significant and positive. Probably the right interpretation is to treat this examiner as an exception: we must assume that our controls are insufficient to take into account this person's circumstances.

Conclusions

We replicate and confirm that patent importance and patent examination delay are negatively related. We confirm that a narrow definition of a technology cycle can be sufficient to recover this relation where learning effects are modest. We use new measures of patent thickets to confirm that strategic filing behaviours probably are not affecting our results. That being said, while our measure of patent interactions is quite small in an absolute sense we do note that more work needs to be done on thicket measures. In particular, defining the correct comparison group appears quite important for the methodology we implement. We leave a full re-working of thicket measures to future work, however, and proceed with our discussion based on Regibeau and Rockett (2010). We note that in our data set, filer behaviour is not supported as the main cause our importance-delay relation since private firms and non-private entities appear to be insignificantly different. Examiner fixed effects are strong and including them increases the magnitude of our main relation modestly. Examiner workload results are consistent with the view that overloaded examiners devote little time to any single patent, but this relation does not seem to hold consistently across the entire examiner group. This suggests that while this is a concern, it may not be a concern that has serious effects on patent examination as a whole.

References

Caballero, R., and A. Jaffe (1993) "How High are the Giant's Shoulders: Am Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth" in O. Blanchard and S. Fischer, Eds, *National Bureau of Economic Research Macroeconomics Annual* 8, 15-86. MIT Press: Cambridge, MA.

Caillaud, B. and A. Duchene (2011) "Patent Office in Innovation Policy: Nobody's Perfect" *International Journal of Industrial Organization* 29(2), 242-252.

Clarkson, G. (2011) "Patent Informatics for Patent Thicket Detection: A Network Analytic Approach for Measuring the Density of Patent Space", Stern School Working Paper available at http://w4.stern.nyu.edu/emplibrary/ACFltbnmV.pdf.

Cohen et al (2000) "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)" *NBER Working Paper 7552.*

Federal Trade Commission (2003) "To Promote Innovation: A Proper Balance of Competition and Patent Law and Policy" available at http://www.ftc.gov/opa/2003/10/cpreport.htm

Graham, S. and D. Mowery (2004) "Submarines in Software? Continuation in US Software Patenting in the 1980s and 1990s" *Economics of Innovation and New Technology* 13(5), 443-456.

Harhoff, D. and S. Wagner (2009) "Modeling the Duration of Patent Examination at the European Patent Office" *Management Science*

Jaffe, A., J. Lerner (2004) *Innovation and its Discontents: How Our Broken Patent System is Endangering Innovation and Progress and What to Do About It.* Princeton, N.J: Princeton University Press.

Johnson, D. K. N. and D. Popp (2003) "Forced Out of the Closet: The Impact of the American Inventors' Protection Act on the Timing of Patent Disclosure" *Rand Journal of Economics* 34(1), 96-112.

Cockburn, I, S. Kortum, S. Stern (2003) "Are All Patent Examiners Equal? Examiners, Patent Characteristics, and Litigation Outcomes" in W. M. Cohen and S. A. Merrill*: Patents in the Knowledge-Based Economy.* Washington: National Academies Press, 17-53.

Langinier, C. and P. Marcoul (2007) "Patents, Search of Prior Art and Revelation of Information" *Staff General Research Papers* 10489 Iowa State University.

Lemley, M. (2001) "Rational Ignorance at the Patent Office" *Northwestern University Law Review* 95(4), 1495.

Lemley, M. and C. Shapiro (2005) "Probabilistic Patents" *Journal of Economic Perspectives* 19(2), 75-98.

Mariam, S. (2011) "Innovation Importance and Patent Grant Delay: Are Valuable Inventions Approved More Quickly? " University of Essex, Department of Economics unpublished MSc thesis.

National Academy of Science (2004) “A Patent System for the 21st Century” available at http://www.nap/edu/books/0309089107/html

Newberg, J. A. (2001) “Competition, Patent Pools, and the Management of Uncertainty “, *Atlantic Law Journal 3*, 1-27.

Palangkaraya, A., P. Jensen, E. Webster (2008) "Applicant Behaviour in Patent Examination Request Lags", *Economics Letters* 101(3), 243-245.

Regibeau, P., K. Rockett (2010) “Innovation Cycles and Learning at the Patent Office: Does the Early Patent Get the Delay?” *Journal of Industrial Economics* 58(2), 222-246.

Lanjouw, J. and M. Schankerman (2004) “Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators” *The Economic Journal* 114, 441-465.

Schuett, F. (2011) “Patent Quality and Incentives at the Patent Office “, mimeo, University of Toulouse.

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

Tables

Table 1: OLS, Cox and Weibull Results for Complete and Truncated Samples using linear time trend variable (File Date)†

	OLS Complete	OLS truncated	Cox Complete	Cox Truncated	Weibull Complete	Weibull Truncated
Constant	7.049*** (0.160)	6.511*** (0.200)	NA	NA	NA	NA
File Date	-0.0000458*** (0.0000126)	0.00000 (0.0000)	1.000168*** (0.0000466)	1.000052 (0.000047)	1.000272*** (.0000365)	1.000153** (0.0000753)
Citations	-0.000365*** (0.000114)	-0.00044*** (0.00002)	1.00101*** (0.000334)	1.00101*** (0.00034)	1.001443*** (0.000344)	1.001515*** (0.000423)
Claims	0.0046*** (0.001)	0.00424*** (0.00137)	0.9893335*** (0.00266)	0.98863*** (0.0035)	0.9892634*** (0.00295)	0.98838*** (0.00389)
Private Firm Dummy	0.0315 (0.0306)	0.0563 (0.0375)	0.9446 (0.0846)	0.92 (0.096)	0.906 (0.0786)	0.866 (0.103)

Standard errors in paretheses; * p<0.10, **p<0.05, ***p<0.01. Coefficients greater than 1 in survival analysis correspond to negative OLS coefficients in our specification. See Regibeau and Rockett (2010).

†Examiner and attorney dummies not reported.

Table 2: OLS, Cox and Weibull results for complete and truncated samples using yearly time dummies†

	OLS Complete	OLS truncated	Cox Complete	Cox Truncated	Weibull Complete	Weibull Truncated
Constant	6.505*** (0.0439)	6.485*** (0.052)	NA	NA	NA	NA
Citations	-0.00045*** (0.00012)	-0.00047*** (0.000131)	1.001463*** (0.000376)	1.00136*** (0.0004)	1.001843*** (0.000394)	1.0016*** (0.000413)
Claims	0.00443*** (0.001)	0.00428*** (0.00134)	0.987733*** (0.00241)	0.98738*** (0.00355)	0.987842*** (0.00274)	0.98732*** (0.00391)
Private Firm Dummy	0.0361 (0.03)	0.0520 (0.0356)	0.9289 (0.0824)	0.926 (0.0968)	0.911 (0.0885)	0.911 (0.102)

Standard errors in paretheses; * p<0.10, **p<0.05, ***p<0.01. Coefficients greater than 1 in survival analysis correspond to negative OLS coefficients in our specification. See Regibeau and Rockett (2010).

†Examiner and attorney dummies not reported

Table 3: Congestion and Experience effect per examiner (only top 15 examiners in terms of patents reviewed within the sample are included)

Examiner	Congestion effect	Experience effect
1	-0.0154** (0.0072)	-0.00051 (0.0004)
2	-0.0021** (0.001)	-0.00048 (0.0004)
3	0.00737 (0.006)	-0.00239* (0.00146)
4	0.00005 (0.0047)	0.00019 (0.0005)
5	-0.00219 (0.005)	-0.00003 (0.0007)
6	-0.00745 (0.005)	0.00196 (0.0017)
7	-0.00384*** (0.0012)	0.0001 (0.0002)
8	-0.0044* (0.0025)	0.0003 (0.00022)
9	-0.00284 (0.0055)	0.000174 (0.0014)
10	0.00168 (0.0011)	-0.00128*** (0.00039)
11	0.00036 (0.00087)	-0.00016* (0.00008)
12	-0.00107 (0.00082)	-0.000054 (0.000056)
13	-0.0376*** (0.0115)	0.01896*** (0.00536)
14	-0.00234** (0.001)	0.00005* (0.00003)
15	0.0871*** (0.008)	-0.0025*** (0.00029)

Standard errors in paretheses; * p<0.10, **p<0.05, ***p<0.01