Lawyer rankings either do not matter for litigation outcomes or are redundant*

Abstract

I investigate the success of litigants in tax cases in England and Wales between 1996 and 2010. I explore the effect upon success of having better-ranked legal representation, according to rankings of barristers published by *Chambers*. I find that, for a variety of model specifications, there is no significant positive effect of having better-ranked legal representation. After conducting a sensitivity analysis, I conclude that better-ranked legal representation might have a positive effect on litigation outcomes, but only if better-ranked lawyers receive cases that are substantially more difficult to win. However, if better-ranked lawyers receive substantially more difficult cases, this suggests consumers of legal representation are sophisticated enough to dispense with legal rankings.

* The data used in this article are available (anonymously) online: see `https://gist.github.com/anonymous/933e185f013c73301c48`. The `R` code used is also available online, at `https://gist.github.com/anonymous/e02e67feddedc2a8cff6`

I. Introduction

Rankings are a common feature of markets for professional services, particularly markets for credence goods (goods the quality of which cannot be assessed by the consumer). This is true for legal education, as the literature on the ranking of US law schools attests (Sauder and Lancaster, 2006; Espeland and Sauder, 2007). It is also true for legal representation. In both the United States and the United Kingdom, there is a competitive market for rankings of lawyers (Carter, 2007). In the United Kingdom, this market is further divided between those companies which produce rankings for all lawyers, and those which produce specific rankings for barristers – lawyers who specialize in litigation.

Such rankings can be studied as the outcome of a particular social process. This, for example, is the approach taken by Espeland and Sauder (2007). It is, however, arguably more important to study whether these rankings matter for outcomes that people care about. It matters whether better-ranked law schools improve their students' chances of post-graduation employment. If there is no systematic relationship between better ranking and better outcomes *ceteris paribus*, it calls into question the usefulness of rankings as measures of product quality.

I offer a first test of the relationship between lawyer rankings and litigation outcomes by analyzing outcomes in tax cases in England and Wales between 1996 and 2010. I thereby contribute both to the literature on the relationship between attorney skill and outcomes, and the literature examining legal rankings in general, and attorney rankings in particular. My findings are important for researchers who might be tempted to use reputation-based measures such as lawyer rankings as a measure of attorney skill, and for legal consumers who might hire more expensive legal representation on the basis of better rankings.

The article proceeds as follows. In Section II I review the literature on attorney skill and outcomes, before moving on to discuss (in Section III) reasons for believing that rankings should (not) be associated with outcomes. I discuss the annual rankings in the *Chambers Guide to the*

Legal Profession in Section IV, before describing tax cases in the UK in Section V, and giving reasons why we should see a stronger association with outcomes in tax compared to other areas of law. I then present both cross-tabulations (Section VI) and multivariate regressions (Section VII) of litigation outcomes. In a subsequent, more technical section, I subject my findings to a sensitivity analysis, to assess the extent to which the results from sections VI and VII might be due to a tendency of better-ranked lawyers to receive more difficult cases. I close by formulating a dilemma for lawyer rankings: either they do not matter for litigation outcomes, or they do matter for litigation outcomes, but only if consumers are sufficiently expert to appreciate the *ex ante* probability of winning their appeal, in which case they are unlikely to need lawyer rankings.

II. Literature review

A small literature directly examines rankings of lawyers, principally with a view to describing the effect that such rankings have on the functioning of the market for legal services. Petroni (2007, 201–203) argues that lawyer rankings can improve efficiency in the market by aiding consumers with limited access to other sources of information about lawyer quality: specifically, minority and low-income clients. Zacharias (2008) is less sanguine about the effects of lawyer rankings, noting that measures of reputation based on customer evaluations may encourage behaviour inimical to professional ideals. Both authors note the possibility that ranking systems may be gamed, but neither attempts to assess the degree to which lawyer rankings measure aspects of lawyer quality.

Some pathologies of ranking systems more generally are discussed by the broader literature on ranking within the law. Scholars have shown how the objects of evaluation typically change their behaviour in response to being ranked (Espeland and Sauder, 2007); how the producers of rankings may demonstrate considerable 'stickiness' (Schultz et al., 2001; Grewal et al., 2008);

and how ranking systems which are based on peer or customer evaluation may suffer from conscious or subconscious biases against specific groups, in particular women (Gill et al., 2011; Sen, 2014).

Part of this broader literature tackles the question of what rankings might matter *for*. In the case of rankings of law schools, we presume that rankings matter for outcomes after graduation, and that if it were shown that, holding other things equal, rankings – or component parts of ranking systems – did not in fact matter for post-graduation outcomes, we would normally conclude that the fault lay with the method of ranking rather than the lack of a link between institutional quality, however defined, and the subsequent employment of graduates. Indeed, this is arguably true for the reputational element of the *US News & World Report* law school rankings (Morriss and Henderson, 2008, 818-819).

In many cases, however, identifying the outcomes for which rankings (ought to) matter is not easy. This can be seen by surveying the literature on lawyer quality. Despite skepticism regarding the quantification of legal quality (Carlson, 1976) there is some agreement on the number of procedures by which it may be assessed. According to Rosenthal (1976), quality may be assessed

1. through identifying the training and certificates held by lawyers, on the assumption that certified lawyers meet a minimum quality standard (and uncertified lawyers do not);

2. through measures of reputation, including the kind of rankings discussed here;

3. through measurement of certain desirable outcomes, such that higher-quality lawyers bring about better outcomes, including litigation outcomes of the kind discussed here;

4. by contrast with measures of negligence, such that quality lawyers are simply non-negligent ones;

5. through process tracing, such that higher-quality lawyers follow, or discharge better, certain procedures

We may therefore rephrase the question, 'do lawyer rankings matter?' as a question about whether reputational measures of lawyer quality match outcome-based measures of quality, under the assumption that lawyer rankings are primarily produced for end-consumers who care about the outcomes of the legal processes in which they are involved (rather than being produced 'as a reference for those already involved in the legal profession': Petroni 2007, 200).

A key question therefore concerns the kinds of outcomes for which we might expect to see evidence of lawyer skill. The literature in political science has concentrated on appellate outcomes, which can in most cases be reduced to a binary variable (either the appeal succeeds or it does not), and which has employed indirect measures of lawyer quality (for example: Justice Blackmun's letter grades assigned to counsel at oral argument (Johnson et al., 2006) or experience (Haynie and Sill, 2007) or senior counsel status (Szmer et al., 2007) or past record (Hanretty, 2013)) to test whether better counsel win more appeals. A more general literature in empirical legal studies has focused on continuous outcomes, sometimes monetary outcomes in personal injury cases (Moorhead et al., 1994; Rosenthal, 1974) but more often sentence lengths in criminal trials, where the stakes are particularly high for clients (Abrams and Yoon, 2007; Shinall, 2010).

This literature has proceeded tentatively, fully aware that the link between quality and outcomes is not strong and subject to confounding effects. As Moorhead et al. (2003, 25) note, "poor lawyers win cases, and good lawyers lose cases". One recent innovation has involved using random assignment of lawyers to cases to eliminate potential confounding. Path-breaking work by Abrams and Yoon (2007) has rapidly been followed by other articles exploiting random assignment (Anderson and Heaton, 2011; Huang et al., 2010). This emphasis on the identification of causal effects is welcome, but awareness of potential confounding needs to go beyond the simple assertion that 'lawyers matter'.

The present article is akin to the political science literature on lawyer effects, in that it concentrates on outcomes all of which may be loosely regarded as 'appellate' outcomes. It is like the most recent randomization literature on lawyer effects in that it explicitly deals with confounding, thought not in such an elegant way as through randomization. It represents an advance, however, in marrying this to the study of rankings. Because the link between reputation and outcomes has been queried in the literature, it is important to provide arguments both for and against the proposition that rank information might matter for litigation outcomes.

III. Why might we expect rankings to matter for outcomes?

The argument in favour of the proposition that rankings matter for outcomes is simply stated. First, lawyer rankings are based on lawyers' reputations amongst certain well-qualified groups, including other lawyers, judges, and (some) consumers. Second, these 'expert judgments' are a suitable proxy for quality, an unobserved characteristic of lawyers. Third, the better the quality of a lawyer, the better (on average, and *ceteris paribus*) the outcomes s/he secures. It therefore follows that there ought to be an association between rankings and outcomes. We know, for example, that judges' evaluations of lawyers' performances *in particular cases* are predictive of appellate outcomes (Johnson et al., 2006), which warrants the idea that judges' evaluations of lawyers' performances in general ought to be predictive in the same way.

A more pragmatic argument in favour of this proposition is simply that it would be difficult to explain the growth in market for rankings if they offered no guide to securing the better outcomes prized by end consumers. Both in the United States and the United Kingdom, the number of players in the market has grown recently. The monopoly held by Martindale-Hubbell directory, with its letter grades, was broken by The Best Lawyers in 1983, who were subsequently joined by Super Lawyers, which went national in 2003 (Carter, 2007). In the United Kingdom, Chambers & Partners (hereafter Chambers), which has produced rankings since 1993/4, and

which now also operates in the United States, faces competition from Legal 500 and from Super Lawyers.

Both the principled and pragmatic arguments are open to objections. We might challenge the idea that particular groups' repute for a lawyer is based on their quality, all things considered. If the measure of quality is a measure of reputation amongst judges (as is the case, at least in part, with Martindale-Hubbell), then measures of quality may attend to skill in courtroom advocacy and the making of fine points rather than skill in litigation more generally (Rosenthal, 1976, 263). If the measure of quality is a measure of reputation amongst other lawyers, then measures of quality may reflect the clubbishness of different lawyers, such that lawyers who rank well are not necessarily lawyers who do their job well, but rather who have an extensive network of contacts spread through different legal strata (Petroni, 2007, 197).

This objection has particular force when rankings are based in whole or in part on client evaluations. There are few good reasons to believe that clients are good judges of quality in the law. Law's claim to being a profession is in part based on the mastery of esoteric knowledge not accessible to those outside of the profession (Macdonald, 1995, 1). Clients who cannot understand the basis for lawyers acting in particular ways have difficulty in evaluating whether their lawyer's action was skilful or not. Given this inability, clients may judge on the basis of irrelevant characteristics, or "tangible aspects of personal service (has the lawyer treated them with respect, have they understood the clients' concerns, have they explained what is happening and so on)" (Moorhead et al., 2003, 19). Certainly, client satisfaction (which may form part of a reputation-based measure) is known to be poorly related to peer reviews of lawyers' performances (Moorhead et al., 2003, 16), and only weakly related to successful outcomes (Moorhead et al., 2003, 24). To the extent that reputational measures are based exclusively or predominantly on what clients say, there is considerable potential for reputational measures of quality to diverge considerably from outcome-based measures of quality.

The pragmatic argument from the growth in the size of the market for rankings is also less convincing the more lawyer rankings are produced not as aids to consumers, but as reference guides for those already in the profession. Whilst rankings produced for consumers might reward those aspects of quality which are related to outcomes, rankings produced for those in the legal profession might reward "lawyers' lawyers" – counsel who put forward ingenious and inventive arguments which are not necessarily efficacious.

These objections depend importantly on the particular rankings, and the ends they serve. In the section that follows, I discuss one provider of lawyer rankings in the UK – *Chambers* – and provide descriptive statistics concerning rankings in the field of tax law. I then go on to justify the choice of tax law as an area where clients are highly sophisticated and outcome oriented, and where objections based on consumer ignorance should have less force.

IV. Rankings in the UK

Chambers & Partners is the oldest provider of lawyer rankings in the UK, having published its *Guide to the Legal Profession* since 1994. These hefty annual volumes provide separate listings of solicitors and barristers. Rankings for barristers concern England and Wales exclusively. [1]

According to Chambers' description of their methodology (accessible through their website), there are two principal sources of information: information supplied by, and interviews with, barristers in chambers, and interviews with clients. Often, barristers' sets will submit information to Chambers in order to help them in the task of identifying clients in complex or important cases. Such submissions are "optional and not a prerequisite for inclusion".

1 Barristers are lawyers who enjoy rights of audience in all courts in England and Wales. Solicitors and in-house counsel may appear in lower courts, or in higher courts after meeting certain additional requirements. The tension between barristers and solicitors regarding rights of audience in higher courts is recounted by Abel (2003, 159–184)

It is not clear how these sources of information are combined. Chambers makes two claims: that "greater weighting is given to the opinions of clients", and "any biased viewpoints tend to cancel each other out". It is not clear from this, however, whether the greater weighting given to client opinions is a formal weight, a weight used to break ties, or an informal weight. Nor is it clear why "biased viewpoints tend to cancel each other out": there is no mention, for example, of the different strategies used by US suppliers of legal rankings to discount within-firm references or reciprocated references (Carter, 2007), and it is difficult to see how persistent societal biases (for example, gender or ethnic biases) could 'cancel out'.[2]

The criteria upon which the rankings are based is clear in part. Chambers writes that factors of 'primary' importance are:

- "The volume, complexity and importance of work undertaken in the year prior to publication"
- "Visibility and profile as reflected in the views of our many interviewees (clients, instructing solicitors and fellow barristers) who are asked to comment on … individual barristers' performance (whether on paper, in court or in conference)"

Rankings are awarded separately for junior and senior barristers (those who have been recognized by the state as being at the peak of their profession, and who have thus become Queen's Counsel). Rankings are awarded in different bands, where Band 1 is better than Band 2, and so on. Sometimes, extra rankings are awarded for star barristers and for new barristers (or new QCs). For tax law, on average 63 barristers are given a rank in any given year, though this figure has increased over time. Tax is one of the larger of almost fifty 'practice areas' covered by Chambers.

2 Note, however, that of the counsel who argued a case in 2010, and who thus feature in the data discussed below, men were not significantly more likely to be listed in Chambers: $\chi^2 = 0.0329$ on 1 d.f., $p = 0.86$. It is difficult to test for ethnic biases given the ethnically homogeneous character of the English bar.

Table 1. Chambers & Partners Ranks as awarded

Juniors	QCs
7. Star Junior (0.8%)	14. Star Silk (4.6%)
6. Band 1 (12.4%)	13. Band 1 (16.3%)
5. Band 2 (17.5%)	12. Band 2 (15.5%)
4. Band 3 (7.7%)	11. Band 3 (11.6%)
3. Band 4 (0.8%)	10. Band 4 (4%)
2. Up-and-coming (4.5%)	9. New Silks (4.3%)
1. [unranked junior]	8. [unranked QC]

Note: percentages in parentheses refer to the distribution of all awarded ranks over the period.

Table 1 shows the distribution of ranks awarded by Chambers for the editions from 1996/7, 2000/1, 2002/3, 2003/4, 2006, 2007, 2008, 2009, 2011 and 2013.[3] The top 'ordinary' bands for juniors and seniors are generally the most numerous.

The distribution of ranks in the population of lawyers involved in tax cases is different from the distribution shown in Table 1 for one minor and one major reason. The minor reason is that some lawyers have multiple ranks, even within tax law, and these multiple ranks must be reduced to a single figure. Separate rankings are often given for different areas within tax law – for example, in corporate tax, private client work, and indirect taxation (value-added taxes, customs and properties taxes). I have collapsed multiple rankings by taking the highest rank assigned, under the assumption that barristers concentrate on areas of comparative advantage. Each barrister is therefore associated with a single rank for a given year.[4]

Second, and more importantly, many lawyers who participate in tax cases are not ranked by Chambers. Table 2 shows the distribution of ranks for counsel who participated in the tax cases described in the following section. Almost two-fifths of lawyers participating in tax cases in a given year did not feature in the edition of Chambers for that same year; of these, unranked

3 Some years could not be sourced due to difficulties locating the print editions of the guide.

4 An appendix explores the results if lagged ranks or all-career averages are used instead of the rank in the given year. The results are similar.

Table 2. Chambers & Partners Ranks in tax cases

Juniors	QCs
7. Star Junior (1.1%)	14. Star Silk (2.2%)
6. Band 1 (6.7%)	13. Band 1 (9%)
5. Band 2 (5.4%)	12. Band 2 (5.9%)
4. Band 3 (2.1%)	11. Band 3 (4.2%)
3. Band 4 (4.4%)	10. Band 4 (1.2%)
2. Up-and-coming (16.5%)	9. New Silks (1.1%)
1. [unranked junior] (29.1%)	8. [unranked QC] (11.2%)

Note: percentages in parentheses refer to the distribution of ranks amongst counsel in tax cases

juniors were far more common than unranked senior counsel.

Tables 1 and 2 both map ranks to numbers, such that a Band 1 QC = 13, and Band 1 junior = 6, and so on. To simplify the analysis I shall sometimes treat these ranks as if they were interval rather than ordinal data. This is most likely false: the gap in quality between unranked junior counsel and Band 4 junior counsel is not equal to, and is likely much larger than, the gap between Band 4 junior counsel and Band 2 junior counsel. One might even argue that the data are not even ordinal, in that star junior counsel might be higher quality than unranked senior counsel. I shall therefore sometimes also treat rank data as though only the fact of receiving a rank were relevant.

V. What kind of outcomes? Tax outcomes in the UK

There are three reasons why tax law represents a good area in which to test the rankings-outcomes link. First, tax cases have comparable outcomes which, on the face of it, advantage neither party. The majority of tax cases in England and Wales are appellate cases in that they are appeals made against determinations of Her Majesty's Revenue and Customs – that is, decisions by the tax authority – , or appeals from the different first-tier courts which heard such

appeals. As such, we may reduce the outcome to a simple dichotomous variable: either the appeal succeeds, or it does not. A comparatively large proportion of appeals do in fact succeed (39%, cf. Eisenberg and Heise (2009) for statistics on US state courts), suggesting that the scales of justice are not so grossly unbalanced that they preclude a role for counsel.

Second, tax cases are well-reported. I use *Simon's Tax Cases*, which reports on "all tax cases decided in the Upper Tribunal, the High Court, the Court of Appeal, Court of Session in Scotland, Court of Appeal in Northern Ireland and the Supreme Court and selected cases decided in the Privy Council, the European Court of Justice and the European Court of Human Rights".[5] Consequently, above the level of the First Tier Tax Tribunal (in which there are occasionally reported decisions), there is no issue with selective reporting.

Third, and most importantly, tax represents a disconfirmatory crucial case (Gerring, 2007) for the argument that rankings are measures of quality which should *prima facie* be associated with better outcomes. Tax clients are often either very often 'sophisticated corporate clients, who know what they want and actively shop for counsel' (Zacharias, 2008), and who may also be repeat-players, or governments who are inevitably repeat-players. As such, their ability to judge the efficacy of legal representation is greater than that of most other clients. Insofar as the rankings for tax lawyers are based on the evaluations of these highly-sophisticated clients, we should expect them to be free of some of the sources of bias that initially led us to be skeptical about the link between reputation- and outcome-based measures of quality.

My data includes information on the outcomes of 462 tax cases heard in 1996, 2000, 2002, 2003, 2005, 2006, 2007, 2008 and 2010, matching the years for which rank information is available. This represents all tax cases reported in *Simon's Tax Cases* for these years, which were heard below the European level, for which a clear outcome was delivered, and in which counsel appeared for both sides (rather than litigants appearing in-person). The number of decisions (438) is less

5 Description taken from `http://www.lexisnexis.co.uk/store/uk/Simons-Tax-Cases-Set/product`

than the number of cases, because some cases were joined. These cases are treated separately in the analysis, because in joined cases it was nevertheless possible for some appellants to be successful whilst other appellants lost.[6]

Most cases (286) were heard at the level of the High Court or the Upper Tribunal, the court of equivalent status created in 2008 to deal with appeals against administrative tribunals. 32 reached the relevant final court (the House of Lords [HoL], the UK Supreme Court [UKSC], or the Privy Council [PC]), with the remainder heard in the Court of Appeal. Surprisingly, cases heard in higher courts are not significantly more complex, using the length of words as a proxy for case complexity.[7]

Most cases (282) involved more than two lawyers; although the average number of lawyers was almost identical for appellants and respondents (1.54). The government was the most common appellant (169), followed by limited companies (137) and individuals (76), with public limited companies less common (38). Sole traders and trusts made limited appearances. Summary statistics for these and other variables are reported in Table 3.

VI. Bivariate analysis

Having argued for the suitability of outcomes in tax litigation as a critical area in which to test the claim that lawyer rankings matter for outcomes, I provide an initial analysis which cross-tabulates outcomes according to whether or not the appellant had the better legal representation, according to the rankings. This precedes the more complicated analysis which follows (Sections VII, VIII).

6 Of the 6 decisions in joined cases, split outcomes occurred in 4: [2003] STC 35, [2008] STC 3090, [2008] STC 885 and [2010] STC 1965.

7 An appendix uses an alternative measure of complexity, namely the number of cases cited in the judgment. The results are similar.

Table 3. Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
Govt. is appellant	462	0.366	0.482	0	1
Limited company is appellant	462	0.297	0.457	0	1
Trust is appellant	462	0.039	0.194	0	1
Publicly-traded company is appellant	462	0.082	0.275	0	1
Individual or sole trader is appellant	462	0.210	0.408	0	1
App. lawyers - resp. lawyers	462	0.011	0.644	−3	2
Complexity (wordcount)	462	11,919.840	7,677.376	1,536	55,501
High Court or equiv.	462	0.619	0.486	0	1
Final court (HoL/UKSC/PC)	462	0.069	0.254	0	1
Post 2008	462	0.091	0.288	0	1

Even a simple cross-tabulation poses certain problems. What does it mean to have better legal representation, according to lawyer rankings? The scenario where a single lawyer represents each side is the simplest, but still poses certain problems. We might claim that a single lawyer is better than another lawyer, according to rankings, if one lawyer has a higher rank than the other lawyer, following the mapping used in Tables 1 and 2. Thus, a Band 1 QC would be better than an unranked QC, who in turn would be better than a Star Junior. Or, we might claim that a single lawyer is better than another lawyer, according to rankings, if if one lawyer is listed in the rankings whilst the other lawyer is not. Thus, a Band 1 QC would be better than an unranked QC, but so would a Star Junior, or a Band 4 Junior.

In the more complicated scenario where one or both sides have more than one lawyer, what does it mean to say that one side has better quality legal representation? Where one or both sides have teams of lawyers, we may say either that one side has better quality legal representation, according to rankings, if the average rank of one side's lawyers is higher than the average rank of the other side. (Note that this requires treating these ranks as interval data). Call this the *rank-based measure*. Or, we may say that one side has better quality legal representation, according

Table 4. Cross-tabulation using rank-based measure

	Appellant loses		Appellant wins	
	n	Percent	n	Percent
Appellant counsel better	108	65.06	58	34.94
Appellant counsel not better	173	58.45	123	41.55

Table 5. Cross-tabulation using listing-based measure

	Appellant loses		Appellant wins	
	n	Percent	n	Percent
Appellant counsel better	67	64.42	37	35.58
Appellant counsel not better	214	59.78	144	40.22

to rankings, if a greater proportion of one side's lawyers are listed in the rankings. Call this the *listing-based measure.*

Accordingly, I cross-tabulate outcomes according to whether or not the appellant's legal representation was better on either the rank-based measure or the listing-based measure. Tables 4 and 5 show that the percentage of cases won by appellants actually decreases given better counsel, on either the rank-based or listing-based measure. In neither case are these proportions significantly different from what we would expect given the marginal distributions (rank-based measure: $\chi^2 = 1.68$ on 1 d.f., $p = 0.19$; listing-based measure: $\chi^2 = 0.55$ on 1 d.f., $p = 0.46$).

This cross-tabulation therefore shows no evidence for the claim that rankings matter for outcomes. However, it is highly likely that the kinds of cases argued by better lawyers acting on behalf of the appellant are systematically different in terms of their observed and unobserved characteristics than other cases. These initial results are particularly susceptible to the objection that better lawyers receive cases that are simply harder. In other to account for this kind of confounding, we must formalize our intuitions about the effect of rankings we wish to identify.

VII. Multivariate analysis

I use y_i to stand for the outcome in case i, such that $y_i = 1$ if and only if the appeal succeeds, and $y_i = 0$ in all other instances. Let T_i be a treatment indicator, such that $T_i = 1$ if and only if the appellant in case i received better legal counsel according to the rankings (on either the rank-based or list-based measure), and $T_i = 0$ in all other instances. For appellants faced with a respondent with known legal representation, the quantity of interest is typically the difference between two potential outcomes: the outcome between the scenario in which they hire better-ranked counsel than the respondent, and the outcome in the scenario in which they hire counsel who are equal or worse according to the rankings. That is, they are interested in the 'treatment effect' for their particular value of i,

$$\tau_i = y_i^1 - y_i^0$$

where y_i^1 indicates the outcome with the 'treatment', and y_i^0 the outcome without.

Fortunately for lawyers, it is not possible to run the same case twice with and without better counsel. We observe either y_i^1 or y_i^0, but not both. We must therefore settle for estimating the *average* treatment effect.

$$\tau = E[Y^1 - Y^0] = E[Y^1] - E[Y^0]$$

where $E[\cdot]$ indicates the expected value over all values of i, including values of i where we do not observe a value for y.

Where cases are a random sample from a broader population, and where the treatment T is assigned randomly to cases, then we can recover $E[Y^1]$ by simply averaging over the observed values of y in the sample which were treated – that is, by taking $\bar{y}^1$. This recovers $E[Y^1]$ by construction: were the sample average to be different from the expected value, the sample would

not be random. We might therefore recover the average treatment effect either by taking a simple difference of means, $\bar{y}^1 - \bar{y}^0$, or in a regression framework, by estimating the coefficient θ in the equation.

$$y_i = \alpha + \theta T_i + \beta X_i + \epsilon$$

where additional pre-test covariates (X_i) might be included solely to improve the accuracy of the estimates of θ.

Unfortunately, with rare exceptions (Abrams and Yoon, 2007), treatment T is not assigned randomly. Regression will therefore not recover the causal effect, because differences between $\bar{y}^1$ and $\bar{y}^0$ may be due to other systemic differences between units which receive the treatment and units which did not. Additional covariates are thus typically included not to improve the accuracy of the estimate of θ (as in the case above), but to 'control for' these additional sources of systemic differences between the two groups. The addition of these controls eventually allows researchers to explicitly assert or to implicitly rely upon an assumption of ignorability, namely that

$$y^0, y^1 \perp T|X$$

or that 'the distribution of the potential outcomes (y^0, y^1) is the same across levels of the treatment variable T, once we condition on confounding variates X' (Gelman and Hill, 2007, 183). In the case of tax cases, an ignorability assumption would involve claiming that the cases received by better or lower quality counsel are not 'easier' or 'harder' after controlling for certain features which might make them such, including overall complexity, the level of the court system, or the identity of the litigants.

My strategy in this article is not (principally) to argue for ignorability conditional on certain covariates – although I do include all the control variables mentioned in Section V – but rather to follow the approach suggested by Blackwell (2013). The key idea is that whilst we can never

demonstrate the truth of the ignorability assumption, we can stipulate a certain degree of departure from ignorability, or a certain degree of *confounding*, caused by better–ranked counsel receiving different types of cases. If we re-estimate our models given a stipulated degree of confounding, we may be able to test whether the identified (non-) effects are still present even given that degree of confounding. In the case of tax outcomes, we might be able to ask whether the non-significant negative outcomes of instructing better-ranked counsel, demonstrated in Tables 4 and 5 cease to be negative, and start to be both positive and significant, for a given degree of confounding – specifically, for a given propensity of better-ranked counsel to receive more difficult cases.

Blackwell's strategy is to characterise the degree of confounding as a function. This function, q, takes as inputs the data, and the values of the treatment, and has as a sole parameter, α. In the article, Blackwell discusses a number of different confounding functions, of which the simplest function is an identity

$$q(a, x; a) = \alpha$$

Here, α is equal to the difference in potential outcomes between treated and untreated units. α is defined in the same units as the outcome. In our case, because we will eventually be using a linear probability model on binary data, α will become the probability of appellate success. If α = 0.1, we would say that the cases where appellants receive better counsel (that is, are 'treated') are 10% more likely to win *ex ante*. It is more likely in our situation that α is negative – that cases where appellants receive (or seek out) better counsel are cases that are ex ante more difficult to win.

Having defined the confounding function, Blackwell is then able to show that, *if we are prepared to specify a given degree of confounding* (that is, set a level of alpha), we can recover the causal effect of the treatment by running a regression on an adjusted value of the outcome vari-

able, which is equal to the observed value, minus the confounding function times the propensity of not receiving the treatment. Where there is absolutely no confounding, the adjusted value is the same. The greater the confounding in a positive direction, the smaller the adjusted value.

This approach requires us to estimate three models: an outcome model (equation 1 below), which deals with the unadjusted outcome; a selection model, which models the propensity of receiving the treatment (equation 2 below), and an adjusted-outcome model, which corrects for the specified degree of confounding (equation 3). These are as follows:

$$y_i = \alpha + \theta T_i + \beta X_i \tag{1}$$

$$Pr[T_i = 1|X] = logit^{-1}(\alpha + \beta X_i) \tag{2}$$

$$y_i - q(T_i, X_i)Pr[1 - T_i|X_i] = \alpha + \theta T_i + \beta X_i \tag{3}$$

except that equation (3) is estimated for multiple values of the confounding parameter α.

Table 6 shows the results for an outcome (equation 1) and a selection model (equation 2) using as the treatment having better-ranked counsel on the rank-based measure. Continuous variables (complexity, relative number of lawyers) have been standardized to have mean zero and standard deviation of one.

Taking the outcome model first, we can see that some of our controls do have significant effects on the probability of appellant success. Governments have slightly better odds of success than the baseline category (limited companies), as we might expect given the literature on government advantage in litigation (McGuire, 1998; McCormick, 1993; Hanretty, 2013; cf. Smyth, 2000). Individuals have slightly worse odds. More complex cases are more likely to result in victories for the appellant, matching the findings of Clermont and Eisenberg (2002). There is no significant effect of the level of the appeal, nor of the difference in the raw number of lawyers. Most

importantly, after controlling for these factors, appellants that receive better-ranked lawyers actually fare worse, rather than better, strengthening the statistically non-significant findings from Tables 4 and 5.

The selection model does show that exposure to the treatment is affected by certain case characteristics. Notably, public limited companies and appellants in more complex cases are more likely to receive better counsel.

If we change the treatment from 'better-ranked lawyers on the rank-based measure' to 'better-ranked lawyers on the listing-based measure', we find little change in either the outcome or the selection models. Everything which is significant at the ten percent level or higher according to the rank-based measure is also significant at the same level or greater according to the listing-based measure, with the exception of the odds of PLCs acting as appellants receiving better-quality counsel.

These results show the effect of better-ranked counsel under the assumption that there is no confounding – that is, that appellants who receive better-ranked lawyers face cases that are not systematically more or less difficult to win that appellants who do not receive better-ranked lawyers. In the following section, I allow for different degrees of confounding, and investigate the changes in the effect of receiving better-ranked counsel.

VIII. Sensitivity analysis

The sensitivity analysis I carry out – which is based on the confounding function described in the previous section, and involves repeatedly estimating equation 3 – is best presented visually. Figure 1 shows the results of two sensitivity analyses carried out using two different versions of the treatment – the rank-based measure and the listings-based measure. In the upper-half of the diagram, the solid line plots the values of the coefficient on the treatment variable (on the y-axis), for a specified level of confounding (on the x-axis); the grey shaded area indicates the 95%

Table 6. Regression models, rank-based measure

	Dependent variable:	
	Appellant wins	Treated
	OLS	*logistic*
	(1)	(2)
Better-ranked lawyers	-0.096^{**}	
	(0.047)	
Appellant: Government	0.137^{**}	−0.087
	(0.054)	(0.245)
Appellant: Trust	−0.030	0.221
	(0.120)	(0.528)
Appellant: Public limited company	0.021	0.714^{*}
	(0.088)	(0.378)
Appellant: Individual or sole trader	-0.167^{***}	−0.318
	(0.063)	(0.291)
No. appellant lawyers - no. respondent lawyers	0.0002	
	(0.022)	
log(Complexity)	0.066^{***}	0.383^{***}
	(0.023)	(0.104)
High Court v. Court of Appeal	0.008	0.095
	(0.050)	(0.226)
HoL/UKSC/PC v. Court of Appeal	0.106	0.167
	(0.095)	(0.427)
Post 2008	−0.034	0.607^{*}
	(0.080)	(0.341)
Constant	0.402^{***}	-0.705^{***}
	(0.058)	(0.252)
Observations	462	462
R^2	0.080	
Adjusted R^2	0.059	
Log Likelihood		−288.513
Akaike Inf. Crit.		595.026
Residual Std. Error	0.474 (df = 451)	
F Statistic	3.901^{***} (df = 10; 451)	
Note:		$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 7. Regression models, listing-based measure

	Dependent variable:	
	Appellant wins	Treated
	OLS	*logistic*
	(1)	(2)
Better-ranked lawyers	-0.112^{*}	
	(0.066)	
Appellant: Government	0.141^{**}	0.187
	(0.054)	(0.274)
Appellant: Trust	−0.028	0.019
	(0.120)	(0.615)
Appellant: Public limited company	0.009	−0.274
	(0.087)	(0.470)
Appellant: Individual or sole trader	-0.163^{**}	−0.107
	(0.063)	(0.329)
No. appellant lawyers - no. respondent lawyers	−0.028	
	(0.028)	
log(Complexity)	0.063^{***}	0.295^{**}
	(0.023)	(0.116)
High Court v. Court of Appeal	0.003	−0.186
	(0.050)	(0.249)
HoL/UKSC/PC v. Court of Appeal	0.098	−0.175
	(0.095)	(0.485)
Post 2008	−0.052	0.240
	(0.079)	(0.377)
Constant	0.396^{***}	-1.189^{***}
	(0.058)	(0.280)
Observations	462	462
R^2	0.077	
Adjusted R^2	0.057	
Log Likelihood		−241.559
Akaike Inf. Crit.		501.118
Residual Std. Error	0.475 (df = 451)	
F Statistic	3.766^{***} (df = 10; 451)	
Note:		$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

confidence interval for the coefficient. In the lower half of the diagram, the same information is plotted, except here the value of the coefficient on the treatment variable is plotted for a specified level of confounding expressed not an an absolute value, but in terms of the variance explained by the confounding.

The figure recapitulates the information from tables 6 and 7 in that it shows a negative and statistically significant, and negative and statistically insignificant effect on outcomes respectively, under zero confounding. As α becomes negative – that is, as the cases received by better-ranked lawyers become more difficult, after controlling for other covariates, the value of the coefficient associated with better-ranked legal representation increases. At certain values of α (-0.19 for the rank-based measure, -0.24 for the list-based measure), the effect of better-ranked representation becomes positive and statistically significant.

What is important is not whether or not the coefficient becomes positive and statistically significant at some level of α, for this is guaranteed; what matters is whether one can plausibly argue that this amount of confounding actually obtains in the real world. One way of proceeding is to ask: how does the amount of confounding compare when set against the magnitude of the other controls in the model? If the degree of confounding required for the treatment effect to be significant and positive is very much larger than the (standardized) magnitude of any of the coefficients in the model, then we might reasonably conclude that there is in fact no significant treatment effect. This is indeed the case, as the top half of the figure also plots ($\times$) the absolute positive and negative values of the coefficients from Tables 6 and 7. Similarly, when we express the amount of confounding in terms of the overall variance explained, we see that the degree of confounding required for the treatment effect to be significant and positive would require the confounding to explain more of the variance than any single observable characteristic.

More generally, we may compare the magnitude of the degree of confounding required for better-ranked counsel to have a positive and significant effect on outcomes to the magnitude

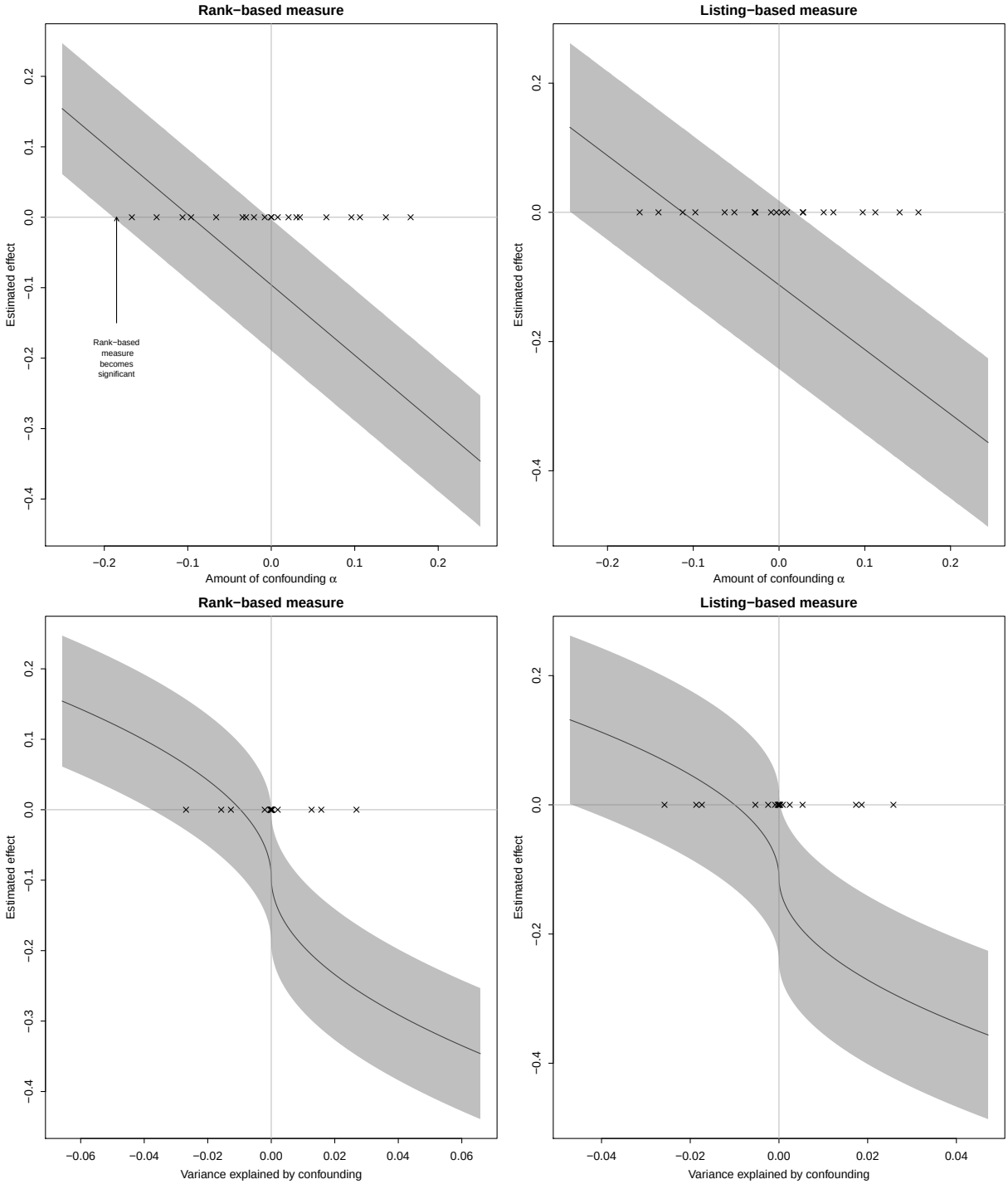

Figure 1. Treatment effect given confounding; negative values mean better-ranked counsel get more difficult cases

of other coefficients in other studies of appellate success. Since most studies of appellate success use logistic regression, we may make those coefficients (which are estimates of log odds ratios) comparable with our ordinary least squares estimates (which are estimates of changes in probabilities) by using the divide-by-four rule.[8]

Here, I concentrate on the magnitude of dummy coefficients, since these are both commonly used in studies of appellate success and do not require information on the standard deviation of the relevant independent variable to be made comparable. I collected information on 116 coefficients from 11 different studies (Farole Jr, 1999; Haire et al., 1999; Hanretty, 2013; Haynie, 1994; McGuire, 1995, 1998; Sheehan et al., 1992; Songer and Sheehan, 1992; Songer et al., 1999, 2000; Szmer et al., 2007). The required magnitude on the rank-based measure was greater than 82.8% of coefficients in the cited studies, and greater than greater than 91.4% of coefficients on the list-based measure. Thus, the answer to the question 'how much harder must the cases received by better-ranked counsel be if we are to conclude that they do have a positive effect on the outcome?' is either 'about 20% harder' or 'very much harder – at least when set against what we can explain about the likely success of an appeal".

IX. Conclusion

Thus far, I have argued that there is no significant and positive effect of having better-ranked counsel on appellate outcomes for tax cases in England and Wales. I have further argued that in order for better-ranked counsel to have such an effect, better-ranked counsel would have to receive cases that are a degree of magnitude more difficult, and that this is unlikely. Given the nature of lawyer rankings, this sets up a dilemma.

8 The divide-by-four rule gives an upper bound of the change in the predicted probability of an event occurring. See Gelman and Hill (2007, 82) for details.

Suppose better-ranked lawyers do not receive cases that are systematically (and substantially) more difficult than do other lower-ranked lawyers. In this case, we must accept that lawyer rankings do not matter for litigation outcomes, insofar as there is no significant and positive effect shown in either a bivariate or multivariate analysis. Consequently, lawyer rankings are not useful to consumers who wish to improve their chances of securing better outcomes in litigation.

If, however, better-ranked lawyers do receive cases that are systematically (and substantially) more difficult than do other lower-ranked lawyers, then the second horn of the dilemma emerges. For better-ranked lawyers to receive cases that are systematically more difficult, consumers must be able to ascertain the difficulty of a case ex ante, and select a lawyer on that basis. This is an extremely demanding requirement, since most lawyers persistently fail to accurately estimate the ex ante probability of winning (Wagenaar and Keren, 1986; Goodman-Delahunty et al., 2010). Consequently, consumers likely have sufficient knowledge of the law and of the legal world to choose lawyers on the basis of that knowledge alone, without the aid of rankings. If consumers have such knowledge, then this means that, under the characterisation of lawyer rankings as aids for inexpert consumers faced with the problem of selecting a credence good, lawyer rankings are redundant. It may of course be that lawyer rankings cannot be characterised in this way, but must instead be characterised as mere reference works. If so, then this materially affects decisions concerning the wisdom of allowing lawyers to cite such rankings in their advertising (Petroni, 2007).

Note that these findings do not mean that lawyer rankings do not matter for outcomes other than those achieved through litigation. One possibility is that better-ranked lawyers are better at reaching negotiated settlements than worse-ranked lawyers. Strictly speaking, this article only addresses the case where consumers know they wish to litigate, and choose a lawyer on this basis. It is possible that the cases that are eventually litigated by better-ranked lawyers are

not representative, because better-ranked lawyers may advantageously settle easy cases and litigate hard ones. Note, however, that this effect would have to be as large as the previously discussed confounding effect. Note also that when we restrict the analysis to cases which end up in the Court of Appeal or in the House of Lords, and which we might expect to be *ex ante* less likely to settle, perhaps because they involve an important point of principle, the effect of having better counsel is still negative (see appendix Tables 10 and 11).

These findings does not necessarily mean that lawyer rankings do not matter for other aspects of quality. In the discussion of lawyer quality above, I mentioned that there were five recognized procedures by which quality might be assessed, but implicitly assumed that quality was a unidimensional latent trait, such that measures based on outcomes ought also to track measures based on other procedures for assessment. This assumption may be erroneous: it may be that lawyers who secure better litigation outcomes for clients are different from lawyers who assiduously cater to clients' demands. Yet if this is true, it seems to undermine the usefulness of rankings, since they do not offer the hope of better outcomes, merely the hope of greater satisfaction.

What these findings do mean is that we should avoid using lawyer rankings as any indication of quality (though they may still serve as indicators of specialization: Haire et al. (1999)), and that we should be generally skeptical of any claim made on behalf of producers of lawyer rankings that these rankings can aid consumers as guides to quality, since quality in this sense does not seem to be connected to outcomes.

References

Abel, R. L. (2003). *English lawyers between market and state: The politics of professionalism.* Oxford University Press, Oxford.

Abrams, D. S. and Yoon, A. H. (2007). The luck of the draw: Using random case assignment to investigate attorney ability. *University of Chicago Law Review*, 74(4):1145–1177.

Anderson, J. and Heaton, P. (2011). How much difference does the lawyer make? the effect of defense counsel on murder case outcomes. *RAND Working Papers*, (WR-870-NIJ).

Blackwell, M. (2013). A selection bias approach to sensitivity analysis for causal effects. *Political Analysis.* In press.

Carlson, R. J. (1976). Measuring the quality of legal services: an idea whose time has not come. *Law and Society*, 11(2):287–317.

Carter, T. (2007). The ratings game. *American Bar Association Journal*, 93:27.

Clermont, K. and Eisenberg, T. (2002). Plaintiphobia in the appellate courts: Civil rights really do differ from negotiable instruments? *University of Illinois Law Review*, 2002:947–977.

Eisenberg, T. and Heise, M. (2009). Plaintiphobia in state courts? an empirical study of state court trials on appeal. *Journal of Legal Studies*, 38(1):121–155.

Espeland, W. N. and Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113(1):1–40.

Farole Jr, D. J. (1999). Reexamining litigant success in state supreme courts. *Law & Society Review*, 33(4):1043–1058.

Gelman, A. and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge.

Gerring, J. (2007). Is there a (viable) crucial-case method? *Comparative Political Studies*, 40(3):231–253.

Gill, R. D., Lazos, S. R., and Waters, M. M. (2011). Are judicial performance evaluations fair to women and minorities? A cautionary tale from Clark County, Nevada. *Law & Society Review*, 45(3):731–759.

Goodman-Delahunty, J., Granhag, P. A., Hartwig, M., and Loftus, E. F. (2010). Insightful or wishful: Lawyers' ability to predict case outcomes. *Psychology, Public Policy, and Law*, 16(2):133–157.

Grewal, R., Dearden, J. A., and Llilien, G. L. (2008). The university rankings game. *The American Statistician*, 62(3).

Haire, S. B., Lindquist, S. A., and Hartley, R. (1999). Attorney expertise, litigant success, and judicial decisionmaking in the US Courts of Appeals. *Law and Society Review*, 33(3):667–685.

Hanretty, C. (2013). Haves and have-nots before the Law Lords. *Political Studies*. In press.

Haynie, S. L. (1994). Resource inequalities and litigation outcomes in the Philippine supreme court. *The Journal of Politics*, 56(03):752–772.

Haynie, S. L. and Sill, K. L. (2007). Experienced advocates and litigation outcomes: Repeat players in the South African Supreme Court of Appeal. *Political Research Quarterly*, 60(3):443–453.

Huang, K.-C., Chen, K.-P., and Lin, C.-C. (2010). Does the type of criminal defense counsel affect case outcomes?: A natural experiment in Taiwan. *International Review of Law and Economics*, 30(2):113–127.

Johnson, T. R., Wahlbeck, P. J., and Spriggs, J. F. (2006). The influence of oral arguments on the US Supreme Court. *American Political Science Review*, 100(1):99–113.

Macdonald, K. M. (1995). *The sociology of the professions*. Sage.

McCormick, P. (1993). Party capability theory and appellate success in the Supreme Court of Canada, 1949–1992. *Canadian Journal of Political Science*, 26(03):523–540.

McGuire, K. T. (1995). Repeat players in the Supreme Court: The role of experienced lawyers in litigation success. *The Journal of Politics*, 57(01):187–196.

McGuire, K. T. (1998). Explaining executive success in the US Supreme Court. *Political Research Quarterly*, 51(2):505–526.

Moorhead, R., Sherr, A., and Paterson, A. (1994). Judging on results? Outcome measures: quality, strategy, and the search for objectivity. *International Journal of the Legal Profession*, 1(2):191–210.

Moorhead, R., Sherr, A., and Paterson, A. (2003). What clients know: client perspectives and legal competence. *International Journal of the Legal Profession*, 10(1):5–35.

Morriss, A. P. and Henderson, W. D. (2008). Measuring outcomes: Post-graduation measures of success in the US News & World Report Law School Rankings. *Ind. LJ*, 83(3):791–834.

Petroni, C. (2007). Third-party ratings as modern reputational information: how rules of professional conduct could better serve lower-income legal consumers. *University of Pennsylvania Law Review*, 156(1):197–227.

Rosenthal, D. E. (1974). *Lawyer and Client: who's in charge?* Russell Sage, New York.

Rosenthal, D. E. (1976). Evaluating the competence of lawyers. *Law and Society Review*, 11(2):257–285.

Sauder, M. and Lancaster, R. (2006). Do rankings matter? The effects of US News & World Report rankings on the admissions process of law schools. *Law & Society Review*, 40(1):105–134.

Schultz, M., Mouritsen, J., and Gabrielsen, G. (2001). Sticky reputation: Analyzing a ranking system. *Corporate Reputation Review*, 4(1):24–41.

Sen, M. (2014). How judicial qualification ratings may disadvantage minority and female candidates. *Journal of Law and Courts*, 2(1):33–65.

Sheehan, R. S., Mishler, W., and Songer, D. R. (1992). Ideology, status, and the differential success of direct parties before the Supreme Court. *American Political Science Review*, 86(02):464–471.

Shinall, J. B. (2010). Slipping away from justice: The effect of attorney skill on trial outcomes. *Vanderbilt Law Review*, 63(1):267–306.

Smyth, R. (2000). The 'haves' and the 'have nots': An empirical study of the rational actor and party capability hypotheses in the High Court 1948–99. *Australian Journal of Political Science*, 35(2):255–274.

Songer, D., Kuersten, A., and Kaheny, E. (2000). Why the haves don't always come out ahead: Repeat players meet amici curiae for the disadvantaged. *Political Research Quarterly*, 53(3):537–556.

Songer, D. R. and Sheehan, R. S. (1992). Who wins on appeal? upperdogs and underdogs in the United States Courts of Appeals. *American Journal of Political Science*, 36(1):235–258.

Songer, D. R., Sheehan, R. S., and Haire, S. B. (1999). Do the haves come out ahead over time-applying Galanter's framework to decisions of the United States Courts of Appeals, 1925-1988. *Law & Society Review*, 33(4):811–832.

Szmer, J., Johnson, S. W., and Sarver, T. A. (2007). Does the lawyer matter? influencing outcomes on the Supreme Court of Canada. *Law & Society Review*, 41(2):279–304.

Wagenaar, W. A. and Keren, G. B. (1986). Does the expert know? the reliability of predictions and confidence ratings of experts. In *Intelligent decision support in process environments*, pages 87–103. Springer.

Zacharias, F. C. (2008). Effects of reputation on the legal profession. *Washington & Lee Law Review*, 65(1):173–212.

A. Robustness to alternative specifications of the treatment

In the body of the article, I presented models based on two specifications of the treatment – a specification according to which the treatment was based on whether or not the average rank, for that year, of the appellant team was higher than the average rank, for that year, of the respondent team (the rank-based measure); and a specification according to which the treatment was based on whether or not a higher proportion of the appellant team was listed in the rankings for that year, compared to the respondent team (the listing-based measure).

However, there are other ways to match and aggregate rankings. One option concerns the choice of year. I have matched each lawyer argued a case in calendar year x with their ranking for the edition of Chambers published in year x, such that the ranking marginally precedes the performance (though likely does not precede the receipt of instructions). However, we might equally well match each lawyer with their ranking:

- in the following year (to check whether ranking information was even capable of retrodicting behaviour)
- in the previous year (to check whether ranking information was strongly predictive)
- over their career (to check whether averaging removes random year-to-year noise which prevents a significant effect being found)

Equally, we may also consider different choices concerning the aggregation of rankings. Here, I have taken the average. It would also be possible to take the maximum, on the (plausible) assumptions that the lead barrister in each team lays out the legal strategy; that legal strategy rather than on-the-day advocacy is the primary determinant of outcomes; and that rankings offer information on who in each team is likely to be the lead barrister.

Figure 2 therefore plots the value of the coefficient on the treatment for each of the different specifications of the treatment: using listings (left hand column) or ranks (right hand column);

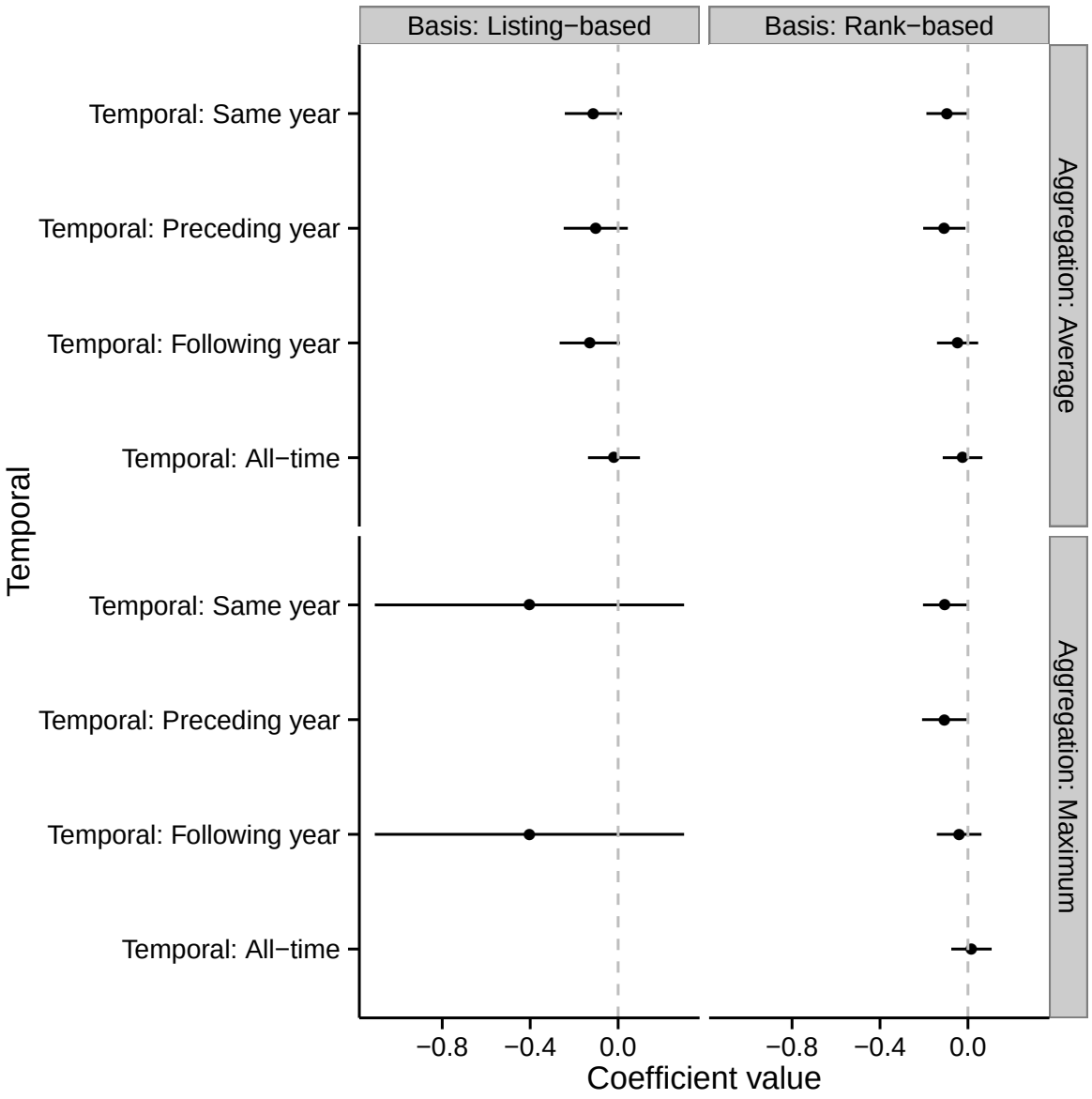

Figure 2. Coefficients by treatment specification

using means (top row) or maxima (bottom row), and using different matches between case years and listing-years. In two cases, the coefficient values are ill-defined, because the method of aggregation and the temporal scope mean we have nothing to distinguish between the appellant and respondent teams.

The figure shows that in only one specification is the coefficient associated with having better counsel positive, and in no specification is the coefficient positive and significant. In 4 specifications the coefficient is negative and significant at the 5% level. Consequently, even allowing for different understandings of 'better-ranked' counsel, there is no positive and significant effect to be found. If anything, we find significant negative effects.

B. Robustness to alternative operationalizations of complexity

In the body of the article I used the log of the number of words in the case as a measure of complexity. Here, in tables 8 and 9, I show the results of the outcome and selection models using a different measure of complexity – namely the log of one plus the number of cases cited in the judgement. These models are a slightly poorer fit to the data. The results are similar, although the effect of having better-ranked lawyers on the listing measure ceases to be significant. The results of the sensitivity analyses, however, are very similar indeed.

C. Restriction to Court of Appeal and Lords

Here, in tables 10 and 11, I estimate the same models as before, but restricting my analysis to cases argued in the Court of Appeal or in the House of Lords. Once again, the effect of having better counsel, on either the listings- or rank-based measure, is negative, but is not significant. For the rank-based measure, the negative effect of having better counsel is smaller; for the listings-based measure, the negative effect of having better counsel is larger.

Table 8. Regression models, ranking-based measure

	Dependent variable:	
	Appellant wins	Treated
	OLS	*logistic*
	(1)	(2)
Better-ranked lawyers	−0.081*	
	(0.047)	
Appellant: Government	0.144***	−0.058
	(0.055)	(0.244)
Appellant: Trust	−0.034	0.193
	(0.121)	(0.523)
Appellant: Public limited company	0.025	0.746**
	(0.088)	(0.374)
Appellant: Individual or sole trader	−0.155**	−0.250
	(0.064)	(0.290)
No. appellant lawyers - no. respondent lawyers	0.0004	
	(0.022)	
log(Complexity (alt.))	0.044*	0.228**
	(0.023)	(0.105)
High Court v. Court of Appeal	0.009	0.091
	(0.050)	(0.225)
HoL/UKSC/PC v. Court of Appeal	0.092	0.084
	(0.095)	(0.422)
Post 2008	−0.036	0.606*
	(0.080)	(0.341)
Constant	0.391***	−0.710***
	(0.058)	(0.251)
Observations	462	462
R^2	0.070	
Adjusted R^2	0.049	
Log Likelihood		−293.160
Akaike Inf. Crit.		604.319
Residual Std. Error	0.477 (df = 451)	
F Statistic	3.386*** (df = 10; 451)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 9. Regression models, listing-based measure

	Dependent variable:	
	Appellant wins	Treated
	OLS	*logistic*
	(1)	(2)
Better-ranked lawyers	−0.099	
	(0.066)	
Appellant: Government	0.147***	0.223
	(0.055)	(0.275)
Appellant: Trust	−0.032	0.015
	(0.121)	(0.612)
Appellant: Public limited company	0.016	−0.239
	(0.088)	(0.468)
Appellant: Individual or sole trader	−0.151**	−0.035
	(0.064)	(0.331)
No. appellant lawyers - no. respondent lawyers	−0.024	
	(0.028)	
log(Complexity (alt.))	0.044*	0.234*
	(0.023)	(0.120)
High Court v. Court of Appeal	0.005	−0.167
	(0.050)	(0.249)
HoL/UKSC/PC v. Court of Appeal	0.085	−0.233
	(0.096)	(0.483)
Post 2008	−0.051	0.224
	(0.080)	(0.377)
Constant	0.388***	−1.217***
	(0.058)	(0.281)
Observations	462	462
R^2	0.068	
Adjusted R^2	0.048	
Log Likelihood		−242.886
Akaike Inf. Crit.		503.773
Residual Std. Error	0.477 (df = 451)	
F Statistic	3.305*** (df = 10; 451)	
Note:		*p<0.1; **p<0.05; ***p<0.01

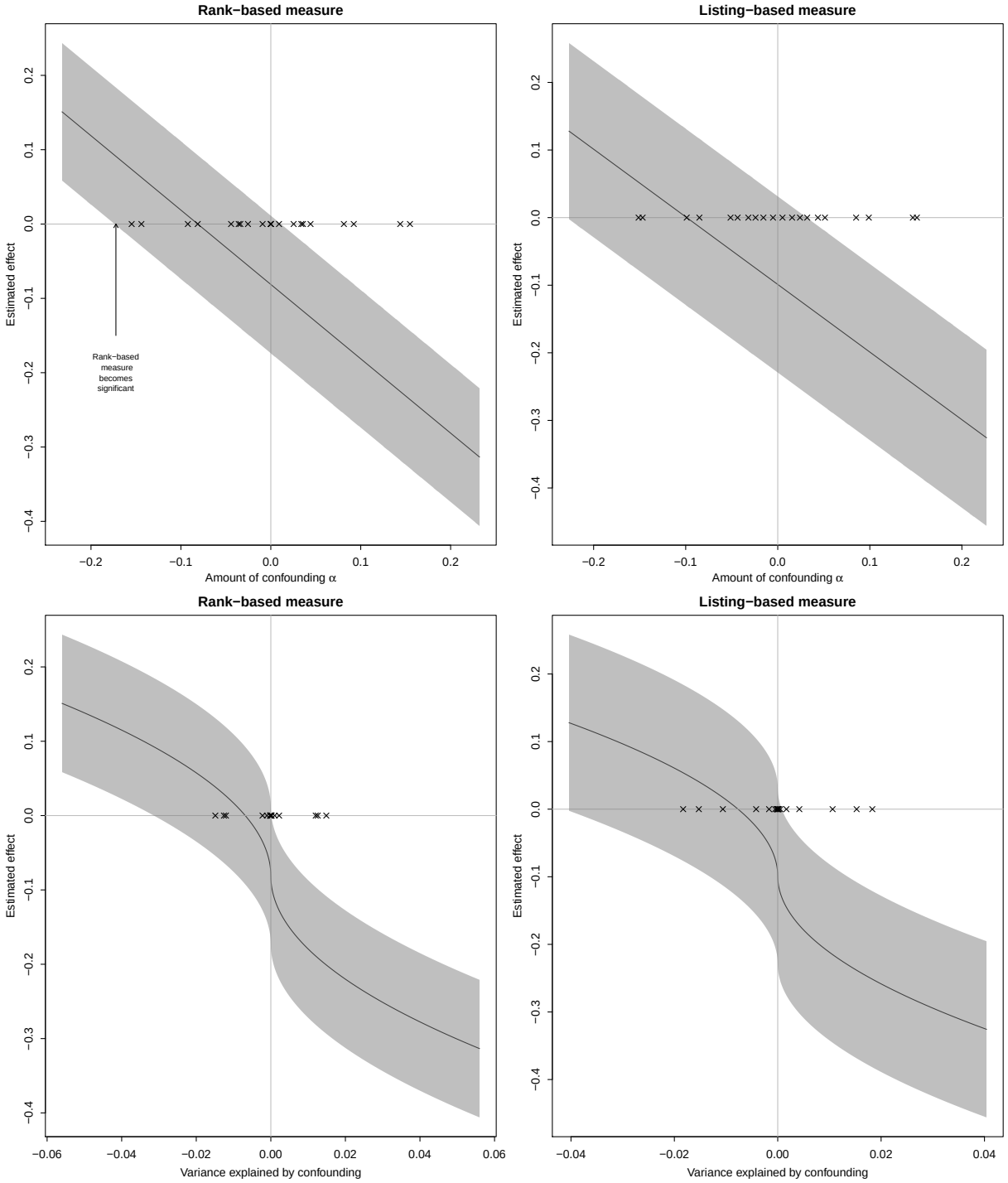

Figure 3. Treatment effect given confounding; negative values mean better-ranked counsel get more difficult cases

Table 10. Regression models, rank-based measure, Court of Appeal and Lords only

	Dependent variable:	
	Appellant wins	Treated
	OLS	*logistic*
	(1)	(2)
Better-ranked lawyers	−0.058	
	(0.079)	
Appellant: Government	−0.016	0.561
	(0.093)	(0.403)
Appellant: Trust	0.099	0.234
	(0.183)	(0.793)
Appellant: Public limited company	−0.143	0.730
	(0.159)	(0.669)
Appellant: Individual or sole trader	−0.168	−0.209
	(0.108)	(0.491)
No. appellant lawyers - no. respondent lawyers	0.0001	
	(0.038)	
log(Complexity)	0.066*	0.127
	(0.039)	(0.170)
HoL/UKSC/PC v. Court of Appeal	0.134	0.134
	(0.101)	(0.432)
Post 2008	−0.047	0.975**
	(0.117)	(0.486)
Constant	0.453***	−0.972***
	(0.080)	(0.346)
Observations	176	176
R^2	0.058	
Adjusted R^2	0.007	
Log Likelihood		−110.820
Akaike Inf. Crit.		237.639
Residual Std. Error	0.490 (df = 166)	
F Statistic	1.137 (df = 9; 166)	
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 11. Regression models, listing-based measure, Court of Appeal and Lords only

	Dependent variable:	
	Appellant wins	Treated
	OLS	*logistic*
	(1)	(2)
Better-ranked lawyers	−0.141	
	(0.113)	
Appellant: Government	−0.018	0.127
	(0.092)	(0.443)
Appellant: Trust	0.099	−0.975
	(0.183)	(1.129)
Appellant: Public limited company	−0.152	−0.075
	(0.158)	(0.758)
Appellant: Individual or sole trader	−0.171	0.164
	(0.108)	(0.505)
No. appellant lawyers - no. respondent lawyers	−0.039	
	(0.050)	
log(Complexity)	0.068*	0.289
	(0.039)	(0.189)
HoL/UKSC/PC v. Court of Appeal	0.125	−0.113
	(0.101)	(0.497)
Post 2008	−0.061	0.455
	(0.115)	(0.498)
Constant	0.473***	−1.214***
	(0.082)	(0.371)
Observations	176	176
R^2	0.064	
Adjusted R^2	0.013	
Log Likelihood		−96.101
Akaike Inf. Crit.		208.201
Residual Std. Error	0.489 (df = 166)	
F Statistic	1.256 (df = 9; 166)	
Note:	*p<0.1; **p<0.05; ***p<0.01	