
William Hubbard
AN EMPIRICAL STUDY OF THE EFFECT OF SHADY GROVE V. ALLSTATE ON FORUM SHOPPING IN THE NEW YORK COURTS

William H. J. Hubbard

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

May 2013


William H.J. Hubbard†

Given the considerable prominence of forum-shopping concerns in the jurisprudence and academic literature on the so-called Erie Doctrine, courts and commentators may benefit from data on whether, and to what extent, forum shopping in fact responds to choice-of-law decisions under the Erie Doctrine. Prior to this paper, however, no empirical study quantified the changes in forum shopping behavior caused by a court decision applying the Erie Doctrine. I study changes in filing patterns of cases likely to be affected by the Supreme Court’s recent decision in Shady Grove v. Allstate and find evidence of large shifts in the patterns of original filings and removals in federal courts in New York that are consistent with the predicted forum shopping response to Shady Grove. In addition to providing the first empirical evidence of vertical forum shopping induced by a decision applying the Erie doctrine, this paper seeks to serve as a proof of concept for empirical research in this area. While there are significant obstacles to empirical research on the effects of Erie and its progeny, this paper outlines a methodology that may be feasible for future projects in this area.

† Assistant Professor of Law, University of Chicago Law School. I thank the participants in the *Erie* Roundtable at the American Enterprise Institute on March 12, 2013 for comments. I am grateful for research assistance from M.D. Akinmurele, Alex Cross, Gary DeTurck, Todd Itami, Lea Madry, Eva Mak, Christine Ricardo, Rob Warfield, Robbie Woods, and Charles Zhang. I thank the Coase-Sandor Institute for Law and Economics for providing research support.
INTRODUCTION

Since the Erie case itself,¹ the so-called Erie Doctrine² has been preoccupied with concerns about the “injustice” of vertical forum shopping.³ In Erie, Justice Brandeis began his broadside against the doctrine of Swift v. Tyson⁴ by raising the specter of vertical forum shopping, as embodied in the notorious Black & White Taxicab case.⁵ Hanna v. Plumer famously characterized “discouragement of forum-shopping” as one of “the twin aims of the Erie rule.”⁶ And even though Hanna distinguished cases implicating the Rules of Decision Act (to which Erie applies) from cases implicating the Rules Enabling Act (to which Erie does not apply), subsequent cases involving the Federal Rules of Civil Procedure and state law have continue to struggle with concerns about vertical forum shopping, whether or not they are, strictly speaking, Erie cases rather than Hanna cases.⁷

¹ Erie Railroad Co. v. Tompkins, 304 U.S. 64 (1938).
² I say “so-called” because of the ambiguity surrounding exactly what one is referring to when one intones the words “Erie Doctrine.” For example, as Allan Erbsen explains in an article appearing in this volume, even the ramifications of Erie itself are best understood as a bundle of no less than four distinct doctrines. Allan Erbsen, Erie’s Four Functions and the Fragmentation of Doctrine, ____. For purposes of this paper, I simply mean to refer to the holdings of Erie and the cases that, by their terms, follow it. I include Hanna and its progeny, including Shady Grove, although one might distinguish them as Rules Enabling Act cases rather than Rules of Decision Act cases.
³ 304 U.S. at 76. By “vertical forum shopping,” I mean the selective choice of federal versus state court to gain a strategic advantage in litigation. In contrast, “horizontal forum shopping” would refer to selectively choosing among state courts for the most favorable forum. Although “forum shopping” usually has a negative connotation, this paper takes no position on whether any particular type of forum shopping is desirable or undesirable.
⁴ 41 U.S. 1 (1842).
⁵ As characterized by the Erie Court, in Black & White Taxicab, the plaintiff corporation reincorporated in a new state for the purpose of manufacturing diversity in order to benefit from more favorable federal law that would be available in federal court, thanks to the doctrine of Swift. Erie, 304 U.S. at 73–74 (citing Black & White Taxicab & Transfer Co. v. Brown & Yellow Taxicab & Transfer Co., 276 U.S. 518 (1928)).
Further, numerous scholars have argued that the *Erie* Doctrine involves a trade-off between vertical and horizontal forum shopping: when federal courts employ state rules, they discourage vertical forum shopping but encourage horizontal forum shopping, which takes advantage of courts’ tendency under modern conflicts-of-law rules to employ forum law.\(^8\) Judgments about whether the *Erie* Doctrine represents good policy therefore turn, in part, on the relative benefits and harms from vertical and horizontal forum shopping.\(^9\)

All of this suggests that the contours of the *Erie* Doctrine may (or should) depend on the extent to which forum shopping *in fact* responds to choice-of-law decisions under the *Erie* Doctrine. Yet in the 75 years following the *Erie* decision, there has not been (to my knowledge) a single empirical study quantifying how vertical forum shopping responded to a decision applying the *Erie* Doctrine. This paper presents the first such study. This paper makes use of recently released administrative data on case filings in federal court, supplemented by a unique data set of complaints filed in New York federal court, to quantify the changes in

---


\(^9\) Academic commentary has also argued that for the Rules Enabling Act analysis under *Hanna*, “the risk of vertical forum shopping [may be] so great that substitution of the federal rule for the state does indeed ‘abridge, enlarge or modify any substantive right’ in contravention of the Rules Enabling Act.” Patrick J. Borchers, *The Real Risk of Forum Shopping: A Dissent from* Shady Grove, 44 CREIGHTON L. REV. 29, 33 (2010).
filing and removal patterns among cases whose claims were likely to have been affected by the Supreme Court’s decision in *Shady Grove v. Allstate*. I predict a rise in federal-court filings by plaintiffs and a decline in removals to federal court by defendants following the *Shady Grove* decision. The data largely confirm these predictions; this paper demonstrates the results both graphically and statistically.

This empirical evidence supports what has long been believed on the basis of anecdotal evidence: court decisions applying the *Erie* Doctrine induce changes in choice of forum by both plaintiffs and defendants. Further, the evidence suggests that the changes in forum choice induced by *Shady Grove* were fairly dramatic in terms of magnitude. At least in this one context, it appears that vertical forum shopping is not a *de minimus* concern for judges or policymakers.

In addition to providing the first empirical evidence of vertical forum shopping induced by a decision applying the *Erie* doctrine, this paper seeks to serve as a proof of concept for empirical research in this area. While there are significant obstacles to empirical research on the effects of *Erie* and its progeny, this paper outlines a methodology that may be feasible for future projects in this area.

The remainder of this paper proceeds as follows. Part I briefly reviews the *Shady Grove* decision and its expected effects on vertical forum shopping. Part II reviews empirical research on related questions of choice of law and forum shopping. Part III outlines the data and methodology employed. Part IV presents results.

**I. *SHADY GROVE* AND VERTICAL FORUM SHOPPING**

*Shady Grove* Orthopedic Associates was a medical care provider that submitted insurance claims to Allstate. Allstate paid the claims, but paid them late, and it refused to pay the 2 percent per month interest rate on late benefits payments required by New York Insurance Law § 5106(a) (“Section 5106(a)”). *Shady Grove* then brought suit against Allstate to recover the unpaid statutory interest. It filed the suit in the Eastern District of New York, invoking the diversity jurisdiction of the federal court. The suit was a putative class action, seeking to sue on behalf of everyone to whom Allstate owes statutory interest under Section 5106(a).
Shady Grove’s individual claim was for only a small sum (approximately $500), so the linchpin to its litigation strategy was certification of a class action. The complication here was that New York law prohibits class certification of claims for statutory damages, such as the statutory interest awarded under Section 5106(a).\(^\text{13}\) New York Civil Practice Law § 901(b) (“Section 901(b)”) states, “Unless a statute creating or imposing a penalty, or a minimum measure of recovery specifically authorizes the recovery thereof in a class action, an action to recover a penalty, or minimum measure of recovery created or imposed by statute may not be maintained as a class action.”

This presented an *Erie/Hanna* question: Given that Federal Rule of Civil Procedure 23 provides criteria for the certification of a class action, is a federal court sitting in diversity bound by Section 901(b) in a case seeking class certification of claims for statutory damages under New York law? The district court and the Second Circuit held Section 901(b) applied in a federal diversity suit. The Supreme Court granted certiorari and reversed.

While no opinion commanded a majority of the Court, five justices agreed that Section 901(b) could not apply in federal court. In the wake of *Shady Grove*, lower courts have noted its application to statutory damages regimes under New York law other than Section 5106(a).\(^\text{14}\) Indeed, there are a number of provisions under the New York General Business Law and the New York Labor Law that provide for statutory damages in one form or another.\(^\text{15}\)

\(^{13}\) 130 S. Ct. at 1437.


\(^{15}\) See, e.g., N.Y. Gen. Bus. Law § 349(h) (“[A]ny person who has been injured by reason of any violation of this section may bring . . . an action to recover his actual damages or fifty dollars, whichever is greater, or both such actions. The court may, in its discretion, increase the award of damages to an amount not to exceed three times the actual damages up to one thousand dollars, if the court finds the defendant willfully or knowingly violated this section.”); N.Y. Lab. Law § 198(1-a) (“In any action instituted in the courts upon a wage claim by an employee . . . in which the employee prevails, the court shall allow such employee to recover the full amount of any underpayment, all reasonable attorney’s fees, prejudgment interest as required under the civil practice law and rules, and, unless the employer proves a good faith basis to believe that its underpayment of wages was in compliance with the law, an
Notably, the issue of forum shopping was squarely in the Supreme Court’s sights as it decided *Shady Grove*. The plurality in *Shady Grove* expressly noted what they perceived as the likely effect of the Court’s decision:

We must acknowledge the reality that keeping the federal-court door open to class actions that cannot proceed in state court will produce forum shopping. . . . But divergence from state law, with the attendant consequence of forum shopping, is the inevitable (indeed, one might say the intended) result of a uniform system of federal procedure.  

Scholars, too, have been quick to predict that “the *Shady Grove* decision will encourage federal forum shopping by plaintiffs to avoid the limiting effects of state provisions that prohibit certain types of class actions.” Practitioners and even New York state court judges have concurred in this prediction. Nonetheless, this view is not quite unanimous.

additional amount as liquidated damages equal to one hundred percent of the total amount of the wages found to be due.”).  

16 See *Shady Grove*, 130 S. Ct. at 1447–48 (Scalia, J., plurality opinion).  
18 See Aaron D. Van Oort & Eileen M. Hunter, *Shady Grove v. Allstate: A Case Study in Formalism Versus Pragmatism*, 11 ENGAGE 105, 109 (Sept. 2010) (“As the dissent emphasizes, the plurality’s formalist approach—and the concurrence’s measured formalist approach as applied in this case—will increase forum-shopping . . . .”).  
19 See Thomas A. Dickerson, John M. Leventhal, Cheryl E. Chambers, *New York State Consumer Protection Law and Class Actions in 2010*, N.Y. ST. B.J. 38, 41 (May 2011) (“Clearly, there will be an increase in federal class actions and defendants may be less anxious to remove such cases to federal court under the Class Action Fairness Act.”). The Hon. Thomas A. Dickerson, Hon. John M. Leventhal, and Hon. Cheryl E. Chambers are Associate Justices of the New York Appellate Division, Second Department.  
20 See W. Daniel “Dee” Miles, III, et al., *Shady Grove v. Allstate: An Erie Sequel(?) and Its Effects Moving Forward*, SS015 ALI-ABA 183, 190 (2010) (arguing that *Shady Grove* will not necessarily increase class action filings and
II. EMPIRICAL RESEARCH ON FORUM SHOPPING

To my knowledge, this is the first empirical study of the effect of the *Erie* Doctrine on vertical forum shopping. There have been, however, various efforts at the theoretical and empirical study of choice of law and forum shopping more generally. Most closely related is the handful of studies that have attempted to identify empirical patterns in vertical forum shopping, though not in the *Erie* context. Specifically, studies have focused on the strategic use of filing in state court by plaintiffs and of removal by defense attorneys.21

A related literature examines empirical evidence of horizontal forum shopping,22 explores the causes of horizontal forum shopping,23 noting that “[a contrary outcome in *Shady Grove*] would have done little to stop class-action lawsuits from occurring in total . . . .”.


studies the use of choice-of-forum clauses, and looks for empirical evidence of possible effects of horizontal forum shopping. Other work provides theoretical models of strategic forum shopping decisions.

An important methodological difference between this study and most of this literature is that I use a discrete change in a legal rule to identify the causal relationship between the legal rule and litigant behavior. Almost none of the studies cited above were designed to do this. While this approach is not always feasible, it has the advantage of allowing one to compare the behavior of litigants in the same types of cases and in the same courts but under two different legal rules. To this extent, the approach taken by this study controls for the characteristics of a given set of cases and court. With this in mind, I turn now to a description of my datasets and methodology.

III. METHODOLOGY AND DATA

This study seeks to shed light on the larger question of the relationships between the Erie Doctrine and forum shopping. This larger question, though, frames an entire research agenda, to which this paper can only make an initial contribution. The precise question this paper asks is a narrow one: What effect did the Supreme Court’s deci-

27 An example of a study that takes something similar to this approach is Atkinson, Marco, and Turner’s examination of filing patterns before and after the change in the appellate structure of the federal district courts brought on by the creation of the Federal Circuit. Scott E. Atkinson, Alan C. Marco, John L. Turner, The Economics of a Centralized Judiciary: Uniformity, Forum Shopping, and the Federal Circuit, 52 J. LAW & ECON. 411 (2009). Taha’s paper on forum shopping in response to judges’ political orientations also uses an empirical strategy that exploits within-district variation over time. Ahmed E. Taha, Judge Shopping: Testing Whether Judges’ Political Orientations Affect Case Filings, 78 U. CIN. L. REV. 1007 (2010). There may be some concern about endogeneity of the variation in that context, however.
sion in *Shady Grove* have on the rates at which plaintiffs filed in federal court, and defendants removed to federal court, putative class actions seeking statutory damages under New York law?

*Shady Grove*, as an exposition of the current state of the *Erie* Doctrine, is hardly transparent. The Court offers three separate opinions, none of which command a majority of justices, and each of which presents a different vision of how to go about deciding the dispute. But for purposes of the empirical question this paper poses, however, the decision is crystal clear: on March 31, 2010, the Supreme Court held that Rule 23 applied, and Section 901(b) did not apply, to diversity cases raising claims for statutory damages under New York law in federal court. Just as crucially, this decision reversed the contrary judgment of the Second Circuit Court of Appeals.

As a consequence, on March 31, 2010, there was a sharp break in the applicable choice-of-law rule for diversity cases brought under New York law seeking statutory damages and class certification. Before March 31, 2010, Section 901(b) applied in federal court. After March 31, 2010, it did not. I utilize this clear break in the application of the *Erie* Doctrine in New York federal courts to identify how patterns of forum shopping respond to that application of the *Erie* Doctrine. My methodology, in essence, is to examine the patterns of putative, diversity-jurisdiction class action filings in New York federal courts. I compare the rates at which plaintiff file in, and defendants remove to, federal court before and after March 31, 2010.

To the extent that *Shady Grove* has affected vertical forum shopping, one should expect not only to see changes in filing rates, but

---

28 Between May 4, 2009, when the Supreme Court granted certiorari in *Shady Grove*, and March 31, 2010, the future status of Section 901(b) in federal court was more uncertain, due to the pending, rather than final, status of the *Shady Grove* litigation. See *Shady Grove Orthopedic Assocs., P.A. v. Allstate Ins. Co.*, 129 S. Ct. 2160 (2009) (granting certiorari). Nonetheless, so long as the decision of the Supreme Court was not a foregone conclusion, March 31, 2010 represents a sharp break in judges' and practitioners' understandings of whether Section 901(b) would apply in federal court. In this respect, the close vote in the *Shady Grove* decision supports the inference that the outcome of the case was uncertain prior to March 31, 2010. Attorneys and judges would have had difficulty predicting the outcome of *Shady Grove* and adjusting their behavior in anticipation of the decision prior to March 31, 2010. Of course, the closeness of the Supreme Court decision is not necessary for such an inference (nor might it be sufficient), and I have argued elsewhere that in some circumstances even a fairly lopsided Supreme Court decision can come as a shock and surprise to both the bar and the bench. See William H.J. Hubbard, *Testing for Change in Procedural Standards, with Application to Bell Atlantic v. Twombly*, 42 J Legal Stud. (forthcoming 2013).
changes in a predictable direction. In this regard, I assume that in most cases involving statutory damages claims, plaintiffs prefer class treatment and defendants do not.29

First, Shady Grove should make plaintiffs’ attorneys more willing to file their cases in federal court rather than state court. Original filings in federal court—i.e., cases that are initially filed in federal court rather than removed to federal court—should rise after the Shady Grove decision.

Second, to the extent that putative class actions are still filed in state court, defendant in state court will be less willing than before to remove these cases to federal court. This change in attitude should show up in the data as a decline in the number of cases that enter the federal court system by way of removal from state courts.

To test these hypotheses I bring to bear two related data sets. My primary data set is composed of administrative data on cases filed in federal court compiled by the Administrative Office of the U.S. Courts (“AO”) and made available to the public on a restricted use basis.30 This data (“AO Data”) contains basic information, such as filing date, jurisdictional basis, nature of suit category, and district of filing, for every case filed in federal court.

From the AO Data, I draw a data set of cases most likely to involve the same types of claims as Shady Grove—statutory damages claims under New York law invoking diversity jurisdiction. I will refer to this data set as the Administrative Dataset. It is impossible, though, to determine from the AO Data whether or not New York law applies in a particular case, let alone whether the plaintiff is seeking statutory damages or whether Section 901(b) might be implicated. Thus, my goal in creating the Administrative Dataset was identify a set of cases most similar to the Shady Grove case itself, and thus plausibly more

29 In Shady Grove the plaintiffs invoked federal jurisdiction under the Class Action Fairness Act of 2005 (“CAFA”), see 28 U.S.C. § 1332(d), arguing that their putative class action was a diversity suit in which the damages sought exceeded $5 million. See Shady Grove Orthopedic Assocs., P.A. v. Allstate Ins. Co., 466 F. Supp. 2d 467, 469 (E.D. N.Y. 2006). While CAFA is largely structured to ensure greater access to federal court for class-action defendants, it clearly favors plaintiffs when Section 901(b) would otherwise apply in state court.

30 See Federal Judicial Center, Federal Court Cases: Integrated Database Series. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. For codebooks and information on this database series, see http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/00072.
likely to involve statutory damages claims under New York law. To do this, I focused on cases with the following characteristics:\(^{31}\)

1. were filed in federal court in New York;\(^{32}\)
2. brought by a represented party;\(^{33}\)
3. invoked either original diversity jurisdiction or removal diversity jurisdiction;\(^{34}\)
4. were not reopenings of earlier-filed cases or appeals from administrative proceedings;\(^{35}\)
5. were coded as cases involving either contract law generally, insurance, or fraud;\(^{36}\) and
6. were filed in the period from November 19, 2008 through September 30, 2010. This represents the time from the Second Circuit decision in *Shady Grove*\(^{37}\) through six months after the *Shady Grove* decision, which is the latest date for which complete federal court filing data is available.\(^{38}\)

\(^{31}\) For the purpose of facilitating replication of this study, footnotes 32–41 refer to variable names in the AO Data and the numerical codes for the indicated values of those variables. Further details on the AO data and methods for processing this data are provided in William H.J. Hubbard, *Testing for Change in Procedural Standards, with Application to Bell Atlantic v. Twombly*, 42 *J Legal Stud.* (forthcoming 2013). The working paper version is available online on SSRN or at http://home.uchicago.edu/~whhubbar/.

\(^{32}\) This corresponds to district codes 06 through 09 for the four districts in New York.

\(^{33}\) I excluded observations coded as pro se or in forma pauperis in the variables prose and ifp, respectively.

\(^{34}\) This corresponds to jurisdiction code 4 (“diversity of citizenship”).

\(^{35}\) This corresponds to origin codes 1 (“original proceeding”) and 2 (“removed from state court”).

\(^{36}\) This corresponds to natureofsuit codes 110 (“insurance”), 190 (“other contract”), and 370 (“other fraud”). I also included in the scope of the database codes 371 (“truth in lending”), 480 (“consumer credit”), 690 (“other forfeiture and penalty suits”), and 890 (“other statutory actions”), but no observations fell within these categories, presumably because most cases in these categories are brought under federal question jurisdiction and invoke federal statutes.


\(^{38}\) Filing date information was derived from the variables fileyear, filemonth, and fileday. While a longer time period after *Shady Grove* would be ideal, the AO Data currently available provides complete data on filed cases only through September 30, 2010.
I focus on insurance, contract, and fraud cases because these appear to be the most likely to involve claims similar to those in *Shady Grove*. *Shady Grove* itself involved a claim arising out of an alleged breach of contract and New York insurance law, and consumer fraud is an area in which statutory causes of action often provide for statutory damages.39

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations</td>
<td>573</td>
<td>799</td>
<td>711</td>
<td>662</td>
</tr>
<tr>
<td>Class Allegation</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Original Jurisdiction</td>
<td>0.85</td>
<td>0.84</td>
<td>0.80</td>
<td>0.85</td>
</tr>
<tr>
<td>Insurance</td>
<td>0.16</td>
<td>0.19</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>Other Contract</td>
<td>0.79</td>
<td>0.74</td>
<td>0.72</td>
<td>0.74</td>
</tr>
<tr>
<td>Fraud</td>
<td>0.05</td>
<td>0.07</td>
<td>0.05</td>
<td>0.07</td>
</tr>
</tbody>
</table>

I further divide the Administrative Dataset into two groups of cases: a “treatment” group of cases coded as involving class action allegations and a “control” group of cases sharing all of the characteristics of the treatment group other than putative class action status.41 Because Section 901(b) and Rule 23 affect only putative class actions, the treatment group may be affected by *Shady Grove*, while the control group will not be.42 Summary statistics for the Administrative Dataset appear in Table 1.

39 For citations to relevant New York law, see notes 11–15 and accompanying text.
40 Values in all rows other than “Number of Observations” represent shares.
41 Putative class action status was derived from the classaction variable.
42 Note that the predictions above assume that the underlying rate at which plaintiffs' attorneys have the opportunity to file cases does not change dra-
My second dataset supplements the first. One weakness of the AO Data, as noted above, is that it does not contain sufficient information to determine whether Section 901(b) is actually implicated in a given case. Rather, the Administrative Dataset relies only on proxies for cases most similar to *Shady Grove*. In addition, past research has found that the *classaction* variable was less reliably coded than other variables. For these reasons, I created a second dataset of information drawn from individual, human review of a sample of complaints (and notices of removal) from cases in the Administrative Dataset. This data constitutes what I will refer to as the Complaints Subset.

With the assistance of a team of research assistants, I conducted automated word searches and individualized, manual review of representative samples of complaints and notices of removal from cases in the Administrative Dataset to identify complaints in which (1) the plaintiff was *clearly* alleging that class certification was appropriate and/or (2) the plaintiff was *clearly* making a claim for statutory damages under New York law.

\[\text{matically at around the time of the} \quad \text{Shady Grove} \quad \text{decision. If, however, there simply were not any statutory damages claims to bring in the months following Shady Grove, one would not see an increase in federal-court filings even if plaintiffs’ attorneys became more willing to file in federal court. To address this concern, I use a difference-in-differences (“diff-in-diff”) empirical strategy. This approach relies on a second, “control” group of cases similar to the treatment group, such that trends over time in the filings rates of the two groups are likely to be similar. Rather than only looking at the change in filing rates of treatment group cases after *Shady Grove* (the “difference”), I also compare the change in treatment group cases after *Shady Grove* with the change in control group cases (the “difference-in-differences”).}\]

\[43 \text{ It is important to note here that this measurement error in the Administrative Dataset leads to untreated observations being coded as belonging to the treatment group. Consequently, the bias introduced by the measurement error is attenuation bias, } \text{i.e., bias toward a finding of no effect. Below, I report statistically significant effects on forum shopping consistent with the predictions above. Given the likely attenuation bias in the Administrative Dataset, the actual effects on vertical forum shopping may be even } \text{larger than estimated.}\]

\[44 \text{ Thomas E. Willging, et al., } \text{Empirical Study of Class Actions in Four Federal District Courts: Final Report to the Advisory Committee on Civil Rules} \quad 198–199 (Federal Judicial Center 1996) \text{ (examining class action codes for period 1989–1994).}\]
## Table 2: Summary Statistics, Complaints Subset\(^{45}\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Observations</td>
<td>94</td>
<td>21</td>
</tr>
<tr>
<td>Class Allegation</td>
<td>0.32</td>
<td>0.19</td>
</tr>
<tr>
<td>Original Jurisdiction</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td>Statutory Damages</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>Insurance</td>
<td>0.24</td>
<td>0.48</td>
</tr>
<tr>
<td>Other Contract</td>
<td>0.71</td>
<td>0.52</td>
</tr>
<tr>
<td>Fraud</td>
<td>0.04</td>
<td>0.00</td>
</tr>
</tbody>
</table>

For this reason, the Complaints Subset represents a subsample of the Administrative Dataset for which I have high confidence that the treatment-group cases involved statutory damages and class-action allegations. This attempts to counterbalance the risk that the “treatment” group in the Administrative Data is over-inclusive, in that it contains cases not involving class action allegations or statutory damages. It creates the converse risk, however, of a sample that includes far fewer cases than actually were affected by *Shady Grove*.\(^{46}\) The small sample size of the Complaints Subset means that this data has little statistical power. Thus, I rely on it to supplement the Administrative Dataset but do not subject it to regression analysis. Notably,

---

\(^{45}\) Values in all rows other than “Number of Observations” represent shares. Note that this dataset is heavily skewed toward removed cases relative to the Administrative dataset as a whole. This reflects deliberate effort to oversample removed cases due to the relative infrequency of removed class action cases involving statutory damages. Despite this oversampling, we were unable to find any removed, putative class action involving statutory damages claims after *Shady Grove*.

\(^{46}\) Note that under federal pleading rules, there is no requirement that the original complaint allege that the plaintiff will seek class certification, nor that the plaintiff clearly distinguish claims for statutory damages from claims for actual, nominal, or punitive damages.
all of the observed effects in the Complaints Subset match the direction of the effects in the Administrative Dataset. Summary statistics for the Complaints Subset appear in Table 2.

IV. RESULTS

I begin with an informal, graphical presentation of results.

**Figure 1: All Cases, Filings per Month, Administrative Dataset**

I first establish a baseline for what federal-court filing rates look like for cases within the scope of the Administrative Dataset including those cases in the “control” group. Given that the vast majority of cases are not class actions, we should expect the overall rate of case filings not to be affected by the *Shady Grove* decision on March 31, 2010. This is exactly what I find. Figure 1 plots the per-month number of insurance, contract, and fraud cases based on diversity jurisdiction filed in or removed to the federal courts of New York during the two-year period from November 19, 2008 through September 30, 2010. While there is a slight downward trend over time, this trend is steady throughout the time period and appears no different after *Shady Grove* than before.
If I look only at cases coded as putative class actions, however, a very different picture emerges. Beginning in April 2010, there is an immediate spike in class-action filings. This sudden and dramatic rise in class-action filings stands in contrast to the continued and gradual decline in overall filings in the Administrative Dataset.

**Figure 2: Class-Action-Coded Cases, Filings per Month, Administrative Dataset**

Breaking down the Administrative Dataset based on case origin—original filing versus removal—is even more illuminating. Among all cases, we should expect to see no overall effect of *Shady Grove* on original filings or removals. This is in fact what I find, as Figure 3 reports. The thick, solid line represents original filings, while the dashed line represents removals. The thin lines indicate the average numbers of filings in each category over the pre- and post-*Shady Grove* periods. As the thin lines indicate, the average for original filings is somewhat lower, reflecting the gradual downward trend in filings. The average for removals is about the same before and after, although slightly lower in the later period.

When focusing on class-action cases, we should expect to see original filings rise after *Shady Grove*, but removals to fall, as defendants no longer see federal court as a friendlier forum for defending a class
action. Figure 4 presents the data for putative class actions in the Administrative Dataset, and the patterns match the predictions exactly. There is a sharp jump in original filings beginning April 2010. And although few putative class actions were ever removed in this dataset, the already low rate of removals goes to zero after Shady Grove. There was literally not one class action removed after March 31, 2010.

![Figure 3: All Cases, Filings per Month, Original versus Removal Jurisdiction](image)

Regression analysis largely confirms that significant changes that occurred in the pattern of filings and removals in the wake of Shady Grove. To estimate the effect of Shady Grove on original filings of putative class actions, I first estimate the following model:

\[
Filings = \beta \text{Shady} + \gamma \text{Trend} + \delta + \epsilon
\]

where Shady is an indicator equal to one for all months after March 2010, Trend is a monthly linear trend variable, and \(\delta\) is a vector of fixed effects for judicial district. The outcome variable is the number of putative class actions filed in a district in a month. Column 1 of Table 3 presents the ordinary least squares (OLS) regression estimate results for Equation 1. The estimated effect of Shady Grove statistically significant at the 1 percent level and is extremely large—the average
number of putative-class-action filings per month per district in this sample is approximately 1.02, so a coefficient of 1.21 implies that *Shady Grove* led to more than a doubling of the rate of original filings of putative class actions.

**Figure 4: Class-Action Coded Cases, Filings Per Month, Original versus Removal Jurisdiction**

Although OLS represents the simplest and most familiar estimation model, it is worth considering an alternate regression method. The outcome I am measuring, filings per month, is a “count”—a non-negative, whole number. Further, there are a large number of “zeros” in the data—district-month observations in which no putative class actions were filed. For this reason, an estimation method suited to counts data with large numbers of zeros, such as the negative binomial regression or the zero-inflated Poisson regression, may be appropriate.\(^{47}\) Column 2 in Table 3 reports results using a negative binomial

\(^{47}\) The problem with OLS in this context is that it assumes a linear relationship between the regressors and the dependent variable. For any non-zero coefficient vector, this implies that there are values of the regressors such that the dependent variable is negative, which is impossible for counts data. In many contexts, this is a theoretical rather than a practical concern, but
model. I report marginal effects, rather than estimation coefficients, so that the negative binomial results are comparable to the OLS results. The results are robust, insofar as the estimates in Column 2 are nearly identical to those in Column 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>(1) OLS</th>
<th>(2) NB</th>
<th>(3) OLS</th>
<th>(4) NB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>All NY</td>
<td>All NY</td>
<td>SD NY</td>
<td>SD NY</td>
</tr>
<tr>
<td>Shady</td>
<td>1.352 (0.420)**</td>
<td>1.358 (0.437)**</td>
<td>5.219 (1.201)**</td>
<td>5.343 (1.714)**</td>
</tr>
<tr>
<td>Time Trend</td>
<td>–0.035 (0.029)</td>
<td>–0.046 (0.030)</td>
<td>–0.152 (0.084)</td>
<td>–0.198 (0.118)</td>
</tr>
<tr>
<td>District FE</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Constant</td>
<td>0.438 (0.261)</td>
<td>1.482 (0.747)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>88</td>
<td>88</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

A related concern with the estimates in Column 1 is the fact that for most district-month observations for districts other than the Southern District of New York, the number of filings is zero. Thus, the Southern District may be the only district with sufficiently regular class-action practice to support statistical inferences about class actions. For this reason, Columns 3 and 4 repeat the analysis from the given that most counts in this data were at or close to zero, this may not be the case here.

48 Results using a Poisson or zero-inflated Poisson model were quite similar for all specifications reported herein, although the zero-inflated Poisson failed to converge for some specifications.

49 Note: * and ** denote statistical significance at the 5 percent and 1 percent levels, respectively. OLS indicates ordinary least squares regression and NB indicates negative binomial regression. For negative binomial regressions, marginal effects are reported.

50 Only four district-month observations outside the Southern District have non-zero values for class action filings.
first two columns, but restrict the data to observations in the Southern District. Once again, the effects are highly significant. For the Southern District, *Shady Grove* appears to have tripled the rate of filings—a huge effect, but not surprising given Figure 4.

A third and final concern with relying on the straightforward OLS results in Table 3 is the fact that it may be desirable to compare the changes for putative class actions in the Administrative Dataset with a “control” group in order to control for any changes in filing patterns over time that are caused by unobserved factors. To account for this possibility, I employ a difference-in-difference strategy, which I estimate as follows:

\[
(2) \quad \text{Filings} = \beta \, \text{Shady} + \theta \, \text{Class} + \rho \, (\text{Shady} \times \text{Class}) + \gamma \, \text{Trend} + \delta + \varepsilon
\]

where, in addition to the variables from Equation 1, *Class* is an indicator equal to one for putative class actions and *Shady×Class* is an interaction term equal to one for putative class actions filed after *Shady Grove*. I now use the entire Administrative Dataset, rather than only the cases coded as involving class allegations. Thus, there are two observations per district per month: one for putative class action filings in that district in that month, and one for all other filings. This allows me to use the non-class-action cases as a control group against which to compare the change in filing rates after *Shady Grove*.

Column 1 of Table 4 presents OLS regression estimates for Equation 2. Note that for the diff-in-diff specification, the coefficient of interest is the coefficient for *Shady×Class*, not *Shady*. The estimated effect of *Shady Grove* is large, but not statistically significant. Column 2 presents the results of the negative binomial regression, and Columns 3 and 4 report the results for the Southern District only. For these latter three columns, the estimates effects of *Shady Grove* are highly statistically significant (although implausibly large).

In reporting regression estimates for the Administrative Dataset, I have focused exclusively on estimates of the effect of *Shady Grove* on original filings. Given the relatively small numbers of removed class actions in the sample, I am reluctant to draw inferences from specifications that rely on removal numbers to estimate the effect of *Shady Grove*. Thus, the results for the Administrative Dataset that I report examine the effect of *Shady Grove* on original filings only. Unreported results of regressions on removals (available from author) broadly confirm the predictions for removal rates, although many of the estimates are not statistically significant.
TABLE 4: DIFF-IN-DIFF REGRESSION RESULTS, ORIGINAL JURISDICTION CASES, ADMINISTRATIVE DATASET51

<table>
<thead>
<tr>
<th>Model</th>
<th>(1) OLS</th>
<th>(2) NB</th>
<th>(3) OLS</th>
<th>(4) NB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>All NY</td>
<td>All NY</td>
<td>SD NY</td>
<td>SD NY</td>
</tr>
<tr>
<td>Shady</td>
<td>-0.728</td>
<td>1.294</td>
<td>0.363</td>
<td>6.587</td>
</tr>
<tr>
<td></td>
<td>(4.141)</td>
<td>(1.098)</td>
<td>(5.315)</td>
<td>(3.495)</td>
</tr>
<tr>
<td>Class</td>
<td>-25.80</td>
<td>-47.60</td>
<td>-66.63</td>
<td>-112.9</td>
</tr>
<tr>
<td></td>
<td>(2.320)**</td>
<td>(2.294)**</td>
<td>(2.977)**</td>
<td>(6.290)**</td>
</tr>
<tr>
<td>Shady×</td>
<td>5.672</td>
<td>13.81</td>
<td>17.63</td>
<td>37.33</td>
</tr>
<tr>
<td>Class</td>
<td>(4.442)</td>
<td>(2.967)**</td>
<td>(5.701)**</td>
<td>(8.196)**</td>
</tr>
<tr>
<td>Time Trend</td>
<td>-0.362</td>
<td>-0.339</td>
<td>-1.313</td>
<td>-1.275</td>
</tr>
<tr>
<td></td>
<td>(0.245)</td>
<td>(0.075)**</td>
<td>(0.315)**</td>
<td>(0.232)**</td>
</tr>
<tr>
<td>District FE</td>
<td>Y</td>
<td>Y</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Constant</td>
<td>23.79</td>
<td></td>
<td>59.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.072)**</td>
<td></td>
<td>(3.164)**</td>
<td></td>
</tr>
</tbody>
</table>

| N   | 176 | 176 | 44 | 44 |

I now turn to the Complaints Subset. The number of cases in the Complaints Subset clearly involving both statutory damages and class action allegations was quite small: fourteen to be exact. But because the Complaints Subset contains information on both class-action allegations and statutory damages claims, it permits two approaches to the difference-in-differences methodology.

First, one can look at putative class actions, and compare those with and without statutory damages claims. The prediction for original filings would be that statutory damages class actions would rise relative to all class actions after Shady Grove. The reverse would be true for removals. Rows (1) and (3) in Table 5 do this for original filings and removals, respectively.

51 Note: * and ** denote statistical significance at the 5 percent and 1 percent levels, respectively. OLS indicates ordinary least squares regression and NB indicates negative binomial regression. For negative binomial regressions, marginal effects are reported.
Second, one can look at cases with statutory damages claims, and compare those with and without class action allegations. The prediction for original filings would be that statutory damages class actions would rise as a share of all statutory damages cases after *Shady Grove*. The reverse prediction would apply to removed cases. Rows (2) and (4) in Table 5 do this for original filings and removals, respectively.

**TABLE 5: DIFF-IN-DIFF STATISTICS, COMPLAINTS SUBSET**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1)</strong> Original Jurisdiction Cases with Class Allegations</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Share with Statutory Damages Claim</td>
<td>64%</td>
<td>75%</td>
</tr>
<tr>
<td><strong>(2)</strong> Original Jurisdiction Cases Claiming Statutory Damages</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Share with Class Allegations</td>
<td>90%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>(3)</strong> Removed Cases with Class Allegations</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Share with Statutory Damages Claim</td>
<td>33%</td>
<td>--</td>
</tr>
<tr>
<td><strong>(4)</strong> Removed Cases Claiming Statutory Damages</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Share with Class Allegations</td>
<td>20%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Rows (1), (2), and (4) show shifts after *Shady Grove* consistent with the predictions. Row (3) does not allow a comparison of data to prediction, but the reason is telling: as noted above, within the scope of the cases in the Administrative Dataset, no class actions have been re-
moved since Shady Grove. Overall, the patterns of case filings within the Complaints Subset reinforce the findings based on the Administrative Dataset.\(^{52}\)

**CONCLUSION**

Courts and commentators have long assumed that vertical forum shopping results when federal courts apply the *Erie* Doctrine in ways that preference federal rules. Prior to this paper, however, no empirical study had quantified the changes in forum shopping behavior caused by a court decision applying the *Erie* Doctrine. I study changes in filing patterns of cases likely to be affected by the *Shady Grove* decision and find evidence of large shifts in the patterns of original filings and removals in federal courts in New York.

While the existence of vertical forum shopping has scarcely been doubted, its extent has not been systematically studied, and evidence regarding the magnitude of vertical forum shopping activity can inform the debate about the merits of the many facets of the *Erie* Doctrine. I conclude by noting three key limitations of this study and opportunities for further inquiry.

First, I have examined only the effect of a single decision on vertical forum shopping. Whether future decisions, Federal Rules, or statutes will have similar effects remains an open (and potentially very important) question. It is worth noting that in some ways, *Shady Grove* may represent the “worst case” scenario for vertical forum shopping, in that it presents a situation in which forum choice is maximally sensitive to the vertical choice-of-law rule. Section 901(b) is unusual in that it uniformly benefits plaintiffs and its application or non-applications changes the stakes of a lawsuit by orders of magnitude. Other rules implicating the *Erie* Doctrine may not have so stark consequences.\(^{53}\)

\(^{52}\) A final piece of evidence emerges from the Complaints Subset, although it is merely anecdotal in character. Prior to *Shady Grove*, a few class action complaints *explicitly disclaimed* that they were seeking statutory damages under New York law. See No. 09-cv-03219 (S.D. N.Y. filed Apr. 2, 2009) (“[N]o penalties, liquidated damages or punitive damages, whether statutory or otherwise, are sought by plaintiff for himself or on behalf of the class in this action, and any such relief is expressly waived.”); No. 09-cv-00759 (S.D. N.Y. filed Jan. 29, 2009) (“Plaintiff does not seek liquidated damages under the NYLL on behalf of the Rule 23 class.”). Needless to say, no such language has appeared since *Shady Grove*.

\(^{53}\) Consider, for example, the New York rule governing review of jury awards that was the subject of *Gasperini v. Center for Humanities, Inc.*, 518 U.S. 415
Second, this study used only federal court data. Data from the state courts would be a useful check on inferences drawn from patterns in federal court data. The relative inaccessibility of most state-court data, however, remains an impediment to ambitious projects in this area.

Third, and most importantly, a full consideration of forum shopping also requires quantifying the scope and scale of the response of horizontal forum shopping to decisions applying the *Erie* Doctrine. Indeed, a central criticism of the *Erie* Doctrine has been that it seeks to discourage vertical forum shopping, but it ignores its effects on horizontal forum shopping.\(^{54}\) Examining patterns of forum shopping across state courts, however, requires the collection of (relative to federal court data) less accessible state court data, and careful consideration of the fact that horizontal forum shopping requires the consideration of as many as fifty-one alternative forums; vertical forum shopping requires considering only two. This is a serious challenge for future research.

\(^{1996}\). New York Civil Practice Law and Rules Section 5501(c) gives courts the power to “determine that an award is excessive or inadequate”—a standard that could help the defendant or the plaintiff, depending on the case.

\(^{54}\) See note 8.
Readers with comments should address them to:

Professor William H. J. Hubbard  
University of Chicago Law School  
1111 East 60th Street  
Chicago, IL  60637  
whhubbar@uchicago.edu
For a listing of papers 1–600 please go to Working Papers at http://www.law.uchicago.edu/Lawecon/index.html

602. Saul Levmore, Harmonization, Preferences, and the Calculus of Consent in Commercial and Other Law, June 2012
603. David S. Evans, Excessive Litigation by Business Users of Free Platform Services, June 2012
604. Ariel Porat, Mistake under the Common European Sales Law, June 2012
608. Lior Jacob Strahilevitz, Absolute Preferences and Relative Preferences in Property Law, July 2012
611. Joseph Isenbergh, Cliff Schmiff, August 2012
613. M. Todd Henderson, Voice versus Exit in Health Care Policy, October 2012
615. William H. J. Hubbard, Another Look at the Eurobarometer Surveys, October 2012
616. Lee Anne Fennell, Resource Access Costs, October 2012
617. Ariel Porat, Negligence Liability for Non-Negligent Behavior, November 2012
618. William A. Birdthistle and M. Todd Henderson, Becoming the Fifth Branch, November 2012
620. Rosa M. Abrantes-Metz and David S. Evans, Replacing the LIBOR with a Transparent and Reliable Index of Interbank Borrowing: Comments on the Wheatley Review of LIBOR Initial Discussion Paper, November 2012
621. Reid Thompson and David Weisbach, Attributes of Ownership, November 2012
626. David S. Evans, Economics of Vertical Restraints for Multi-Sided Platforms, January 2013
627. David S. Evans, Attention to Rivalry among Online Platforms and Its Implications for Antitrust Analysis, January 2013
632. Adam B. Cox and Thomas J. Miles, Policing Immigration, February 2013
633. Anup Malani and Jonathan S. Masur, Raising the Stakes in Patent Cases, February 2013
637. Lior Jacob Strahilevitz, Toward a Positive Theory of Privacy Law, March 2013
639. Lisa Bernstein, Merchant Law in a Modern Economy, April 2013
640. Omri Ben-Shahar, Regulation through Boilerplate: An Apologia, April 2013
641. Anthony J. Casey and Andres Sawicki, Copyright in Teams, May 2013