ECONOMIC SUBSTANCE AND THE STANDARD OF REVIEW

Christopher M. Pietruszkiewicz[*]

INTRODUCTION

Just as "a melody is more than the notes," tax law is more than the unadorned words of a statute.[1] Judicial safeguards were created and designed as supplemental concepts to disallow certain tax advantages not contemplated by the literal words expressed in a statute. The economic

* Vice Chancellor—Business and Financial Affairs and J.Y. Sanders Professor of Law, LSU Law Center. Former Trial Attorney, United States Department of Justice, Tax Division. The author gratefully acknowledges the invaluable comments and advice of Leandra Lederman and Bryan T. Camp. Finally, a special thank you to my research assistants, Natalie Parria and Carolyn Sinnock, for their research assistance. While I am indebted to all of those individuals I acknowledge above, the responsibility for all errors rests solely with me.

1. Although this eloquent quote analogizing tax law to music comes from the Second Circuit's opinion in *Helvering v. Gregory*, 69 F.2d 809, 811 (2d Cir. 1934), *aff'd*, 293 U.S. 465 (1935), the trial court reached a much different conclusion, stating that "[a] statute so meticulously drafted must be interpreted as a literal expression of the taxing policy, and leaves only the small interstices for judicial consideration." Gregory v. Comm'r, 27 B.T.A 223, 225 (1932), *rev'd sub nom.* Helvering v. Gregory, 69 F.2d 809 (2d Cir. 1934), *aff'd*, 293 U.S. 465 (1935).

substance doctrine, a judicially created concept, is a single-edged sword used by the United States to attack certain transactions that, while "firmly anchored" in the Internal Revenue Code, nonetheless contain tax avoidance features that circumvent the spirit of the congressionally authorized language used in a statute.[2]

Many commentators debate the intricacies of the economic substance doctrine.[3] However, that is not my focus. Instead, this Article considers economic substance in the context of appellate review. Economic substance cases universally involve significant tax liabilities and/or variations on the same type of tax planning or tax sheltering activities. Because of the significant stakes involved in these cases, a decision of the trial court is routinely challenged in an appellate forum.

Thus far, the Supreme Court has avoided the standard of review question in economic substance cases, leaving appellate courts to determine the ultimate winners and losers with the standard of review determining the rules of the game.[4] The rules, however, differ depending on where the game is played, with five circuits favoring a clearly erroneous standard of review,[5] three circuits favoring a de novo standard of review,[6] and four additional circuits equivocating on the appropriate standard of review.[7]

---

2. Joseph Isenbergh, *Musings on Form and Substance in Taxation*, 49 U. CHI. L. REV. 859, 870 (1982) (reviewing BORIS I. BITTKER, FEDERAL TAXATION OF INCOME, ESTATES AND GIFTS (1981)); *see also* Bernard Wolfman, Letter to the Editor, 104 TAX NOTES 445 (2004) ("The [economic substance] doctrine has assured us that neither the government nor practitioners will succeed in their roles if they are excessively literal and mechanical in their reading of the statute, if they fail to read it as part of a statutory scheme through which Congress seeks to accomplish a goal that has breadth and durability."). *But see* Boulware v. United States, 128 S. Ct. 1168, 1176 n.7 (2008) ("We have also recognized that '[t]he legal right of a taxpayer to decrease the amount of what otherwise would be his taxes, or altogether avoid them, by means which the law permits, cannot be doubted.'" (alterations in original) (quoting *Gregory*, 293 U.S. at 469)).

3. *See, e.g.*, Joseph Bankman, *The Economic Substance Doctrine*, 74 S. CAL. L. REV. 5 (2000); Peter C. Canellos, *A Tax Practitioner's Perspective on Substance, Form and Business Purpose in Structuring Business Transactions and in Tax Shelters*, 54 SMU L. REV. 47 (2001); David P. Hariton, *Sorting Out the Tangle of Economic Substance,* 52 TAX LAW. 235 (1999) [hereinafter Hariton, *Sorting Out*]; David P. Hariton, *When and How Should the Economic Substance Doctrine Be Applied?*, 60 TAX L. REV. 29 (2006) [hereinafter Hariton, *When and How*]; David B. McGinty, *Economic Substance, Business Purpose, and Tax Avoidance in Section 351 Contingent Liability Transactions After* Black & Decker*,* Coltec*, and* Hercules, 36 CUMB. L. REV. 1 (2005); Martin J. McMahon, Jr., *Random Thoughts on Applying Judicial Doctrines to Interpret the Internal Revenue Code*, 54 SMU L. REV. 195 (2001); Jeff Rector, Note, *A Review of the Economic Substance Doctrine*, 10 STAN. J.L. BUS. & FIN. 173 (2004).

4. *See, e.g.*, Petition for Writ of Certiorari at *i, Dow Chem. Co. v. United States, No. 06-478 (U.S. Oct. 4, 2006), 2006 WL 2827288, *cert. denied*, Dow Chem. Co. v. United States, 127 S. Ct. 1251 (2007) [hereinafter Petition for Writ of Certiorari, *Dow*]; Petition for Writ of Certiorari at *i, Coltec Indus., Inc. v. United States, No. 06-659 (U.S. Nov. 8, 2006), 2006 WL 3295204, *cert. denied*, Coltec Indus., Inc. v. United States, 127 S. Ct. 1261 (2007).

5. The five circuits favoring the clear error test are the Second, Third, Fourth, Seventh, and D.C. Circuits. *See, e.g.*, Black & Decker Corp. v. United States*,* 436 F.3d 431, 441 (4th Cir. 2006) (analyzing sham transactions as questions of fact (citing Rice's Toyota World, Inc. v. Comm'r, 752 F.2d 89, 92 (4th Cir. 1985) ("Whether . . . a particular transaction is a sham is an issue of fact, and our review of the tax court's subsidiary and ultimate findings on this factual issue is therefore under the clearly erroneous standard."))); Nicole Rose Corp. v. Comm'r, 320 F.3d 282, 284 (2d Cir. 2002)

This Article suggests an appropriate framework for courts to apply when addressing the standard of review in economic substance cases. Part I describes the economic substance doctrine and asserts that the economic substance doctrine arises in cases with inherently factual inquiries involving a determination of particularized facts that are not replicated in subsequent cases.

Part II introduces and analyzes the traditional fact–law distinction that has guided appellate courts in selecting the appropriate standard of review. This part argues that the distinction is ineffective in determining the proper standard of review and suggests that the fact–law distinction is merely a rhetorical pretense used by appellate courts depending on the perceived need for appellate review. Instead, this part proposes that a more or less deferential standard of review should be based on institutional responsibilities of trial and appellate courts. Appellate courts have two primary institutional objectives: to develop the law in a particular area as guidance for future cases and to rectify egregious errors in discrete cases.

("Whether a transaction lacks economic substance is a question of fact that we review under the clearly erroneous standard."); ASA Investerings P'ship v. Comm'r, 201 F.3d 505, 511 (D.C. Cir. 2000) (stating that in sham partnership cases, "mixed questions of law and fact are to be treated like questions of fact"); ACM P'ship v. Comm'r, 157 F.3d 231, 245 (3d Cir. 1998) ("[W]e review [the Tax Court's] factual findings, including its ultimate finding as to the economic substance of a transaction, for clear error."); N. Ind. Pub. Serv. Co. v. Comm'r, 115 F.3d 506, 510 (7th Cir. 1997) ("Our review of decisions regarding the economic substance of transactions for federal income tax purposes is for clear error."); Yosha v. Comm'r, 861 F.2d 494, 499 (7th Cir. 1988) ("The question whether a particular transaction has economic substance, like other questions concerning the application of a legal standard to transactions or events, is governed by the clearly erroneous standard.").

6. The three circuits that favor the de novo test are the Sixth, Tenth, and Federal Circuits. *See, e.g.*, Coltec Indus., Inc. v. United States, 454 F.3d 1340, 1357 (Fed. Cir. 2006) ("The ultimate conclusion as to business purpose is a legal conclusion, which we review without deference."); Dow Chem. Co. v. United States, 435 F.3d 594, 599 (6th Cir. 2006) ("The district court's ultimate conclusion that a transaction is or is not an economic sham is reviewed de novo."); James v. Comm'r, 899 F.2d 905, 909 (10th Cir. 1990) ("[W]e review de novo the ultimate characterization of the transactions as shams.").

7. The four circuits in conflict are the Fifth, Eighth, Ninth, and Eleventh Circuits. *Compare* Estate of Strangi v. Comm'r, 293 F.3d 279, 281 (5th Cir. 2002) (review under "clear error" standard), *and* Lukens v. Comm'r, 945 F.2d 92, 96 (5th Cir. 1991) ("reviewable under the clearly erroneous standard"), *with* Compaq Computer Corp. v. Comm'r, 277 F.3d 778, 780–81 (5th Cir. 2001) ("reviewed de novo"). *Compare* Massengill v. Comm'r, 876 F.2d 616, 619 (8th Cir. 1989) (economic substance inquiry as "essentially factual," and, therefore, subject to clearly erroneous standard of review), *with* IES Indus., Inc. v. United States, 253 F.3d 350, 351 (8th Cir. 2001) (economic substance of transaction is a question of law and subject to de novo review). *Compare* Harbor Bancorp v. Comm'r, 115 F.3d 722, 727 (9th Cir. 1997) ("finding of fact we review for clear error"), Erhard v. Comm'r, 46 F.3d 1470, 1476 (9th Cir. 1995) ("[E]conomic substance is a factual determination that this court reviews for clear error."), *and* Casebeer v. Comm'r, 909 F.2d 1360, 1362 (9th Cir. 1990) ("We review the tax court's ultimate conclusion . . . for clear error."), *with* Sacks v. Comm'r, 69 F.3d 982, 986 (9th Cir. 1995) ("[A]pplication of the legal standards to the facts found [in economic substance cases is] reviewed *de novo*."). *Compare* Karr v. Comm'r, 924 F.2d 1018, 1023 (11th Cir. 1991) ("[T]he Tax Court's finding . . . is normally subject to the clearly erroneous standard of review."), *with* United Parcel Serv. of Am., Inc. v. Comm'r, 254 F.3d 1014, 1017 (11th Cir. 2001) ("The question of the effect of a transaction on tax liability, to the extent it does not concern the accuracy of the tax court's fact-finding, is subject to de novo review."), *and* Kirchman v. Comm'r, 862 F.2d 1486, 1490 (11th Cir. 1989) (standard of review is de novo).

Part III considers the application of an appropriate standard of review in the context of economic substance. The Article concludes that in cases considering nonrecurring facts, institutional considerations suggest that a more deferential standard of review should apply. Because economic substance cases are inherently fact-driven, institutional purposes are best accomplished through the application of a clearly erroneous standard of review.

I. ECONOMIC SUBSTANCE

While there has been a relative resurgence of the economic substance doctrine in recent years based on the proliferation of corporate tax shelters,[8] the economic substance doctrine, in its current incarnation, was born in 1935 in *Gregory v. Helvering*.[9] At its foundation, economic substance is designed to be a fluid concept that eliminates the pretenses of structured transactions and applies to an unlimited range of transactions that allows the Internal Revenue Service to challenge technical tax results based on subjective standards that overlay the objective rules prescribed by the Internal Revenue Code.[10]

The Internal Revenue Service uses economic substance to challenge otherwise valid transactions entered into by private parties that have advantageous tax consequences.[11] Because the taxpayer can select the structure of a transaction, the taxpayer can choose the form that gives rise to tax benefits without considering the interests of the government.[12] Therefore, when parties to a transaction are not adverse and the only nonpartici-

8. From 1995 through October 2006, the United States utilized the economic substance doctrine to challenge taxpayers in 170 decided court cases involving over $4.4 billion in taxable income. Petition for Writ of Certiorari, *Dow*, *supra* note 4, at *12 n.5. Each of these cases, however, involved substantive Code sections as well as raising economic substance issues. Moreover, the 170 cases do not reflect assertions by the United States of the economic substance that were settled or otherwise resolved prior to trial. *Id.* In *Coltec Industries, Inc. v. United States*, the Court of Federal Claims questioned the validity of the economic substance doctrine as violating the separation of powers-based, principally extra-legislative requirements imposed through the economic substance doctrine not enacted by Congress. Coltec Indus., Inc. v. United States, 62 Fed. Cl. 716, 718, 730 (Fed. Cl. 2004). This conclusion was vacated and remanded by the Federal Circuit, reinforcing the viability of the economic substance doctrine. Coltec Indus., Inc. v. United States, 454 F.3d 1340, 1357 (Fed. Cir. 2006).

9. 293 U.S. 465 (1935).

10. *See* Isenbergh, *supra* note 2, at 864–66. The Tax Court in *ACM Partnership* described economic substance in the following terms:

> The tax law . . . requires that the intended transactions have economic substance separate and distinct from economic benefit achieved solely by tax reduction. The doctrine of economic substance becomes applicable, and a judicial remedy is warranted, where a taxpayer seeks to claim tax benefits, unintended by Congress, by means of transactions that serve no economic purpose other than tax savings.

ACM P'ship v. Comm'r, 73 T.C.M. (CCH) 2189, 2215 (T.C. 1997), *aff'd in part and rev'd in part*, 157 F.3d 231 (3d Cir. 1998).

11. *See* Hoffman F. Fuller, *Business Purpose, Sham Transactions and the Relation of Private Law to the Law of Taxation*, 37 TUL. L. REV. 355, 365 (1963).

12. *See* Hariton, *Sorting Out*, *supra* note 3, at 237.

pating adverse party is the tax collector, there is a necessity to protect against manipulating the tax consequences of a transaction.[13]

Economic substance has been described in a variety of ways by a number of different courts and commentators. Despite inconsistent phrasing, all courts agree that the economic substance doctrine has two measuring standards: (1) the subjective intent of the taxpayer in entering into the transaction, and (2) the objective economic impact of the transaction absent the tax implications.[14]

Subjective tax avoidance motives and objective inquiries have been part of our tax landscape as early as 1913, when the first income tax law was passed.[15] While many, if not most, of the provisions in the Internal Revenue Code are objective,[16] subjective features do appear in the Code.[17] Both aspects underlie the economic substance test.

The subjective economic substance test provides that a transaction has economic substance if the transaction is rationally related to a useful nontax business purpose.[18] The objective economic substance test provides that a transaction has economic substance if the transaction results in a meaningful and appreciable enhancement in the net economic position of the taxpayer other than from the reduction of taxes.[19]

---

13. *See* Leandra Lederman, *Statutory Speed Bumps: The Roles Third Parties Play in Tax Compliance*, 60 STAN. L. REV. 695, 727–28 (2007).

14. *See* Coltec Indus., Inc. v. United States, 454 F.3d 1340, 1355 (Fed. Cir. 2006); Dow Chem. Co. v. United States, 435 F.3d 594, 599 (6th Cir. 2006); Winn-Dixie Stores, Inc. v. Comm'r, 254 F.3d 1313, 1316 (11th Cir. 2001); United Parcel Serv. of Am., Inc. v. Comm'r, 254 F.3d 1014, 1018 (11th Cir. 2001); Pasternak v. Comm'r, 990 F.2d 893, 898 (6th Cir. 1993); Kirchman v. Comm'r, 862 F.2d 1486, 1492 (11th Cir. 1989); Long Term Capital Holdings v. United States, 330 F. Supp. 2d 122, 171 (D. Conn. 2004).

15. In the Revenue Act of 1913, an accumulated earning tax was enacted that taxed the shareholders on the earnings of any corporation if the accumulation of income in the corporation was for the *purpose* of avoiding the surtax. Revenue Act of 1913, ch. 16, § 2, 38 Stat. 166, *repealed by* Revenue Act of 1916, 39 Stat. 756.

16. *See, e.g.*, I.R.C. § 163 (West Supp. 2008) (allowing the deduction for interest on a qualified residence). Thus, the mortgage interest deduction creates a tax subsidy to encourage home ownership notwithstanding that a taxpayer's motive may be, in large part, tax avoidance. In this sense, a taxpayer can deduct interest payments on a qualified residence owned but not on a residence that is rented. Interestingly, even if a taxpayer confesses his tax avoidance motive in buying a house instead of renting, the mortgage interest deduction is not denied. *See* Alan Gunn, *Tax Avoidance*, 76 MICH. L. REV. 733, 750 (1978).

17. At present, there are forty-three references in the Internal Revenue Code to "principal purpose," suggesting a subjective element incorporated in the particular code section at issue, thereby granting tax-favored status to those business transactions that are not tainted by tax avoidance. *See* I.R.C. §§ 23, 38, 41, 48, 119, 170, 197, 269, 269A, 302, 306, 311, 336, 355, 357, 367, 382, 409, 414, 453, 467, 468B, 501, 514, 614, 643, 751, 864, 877, 953, 954, 1022, 1031, 1272, 1298, 2107, 2501, 4911, 6015, 6050D, 7872, 7874, 9722 (West Supp. 2008). In addition, there are four references in the Internal Revenue Code to "significant purpose." *See* I.R.C. §§ 4944, 6111, 6662, 6662A (West Supp. 2008). Finally, there are nine references in the Internal Revenue Code to "business purpose." *See* I.R.C. §§ 274, 341, 357, 441, 444, 593, 706, 1378, 2032A (West Supp. 2008).

18. *See* Knetsch v. United States, 364 U.S. 361, 365 (1960); ACM P'ship v. Comm'r, 157 F.3d 231, 247 (3d Cir. 1998); *Pasternak*, 990 F.2d at 898; Rice's Toyota World, Inc. v. Comm'r, 752 F.2d 89, 91–92 (4th Cir. 1985).

19. *See* Compaq Computer Corp. v. Comm'r, 277 F.3d 778, 781 (5th Cir. 2001); IES Indus.,

Courts apply the subjective and objective economic substance tests either conjunctively[20] or disjunctively.[21] The conjunctive test requires a taxpayer to satisfy the subjective and objective aspects of the economic substance test—that the taxpayer has a nontax business purpose for the transaction *and* the transaction has objective economic substance. Conversely, the disjunctive test requires a taxpayer to satisfy either the subjective or objective test to obtain the tax benefits of the transaction—the transaction has economic substance if the taxpayer has either a nontax business purpose for the transaction *or* the transaction has objective economic substance.

Congress is moving toward codification of the economic substance doctrine. In its detailed description of codification efforts, the Senate Finance Committee proposes a conjunctive test: (1) the transaction changes in a meaningful way (apart from Federal income tax consequences) the taxpayer's economic position, and (2) the taxpayer has a substantial non-federal tax purpose for entering into such transaction.[22] However, codification of the economic substance doctrine, under this or any other bill introduced on this topic, does not consider, much less address, the standard of review parameters. As a result, even if the economic substance doctrine is codified, the question of the standard of review will remain a relevant consideration.

A. Subjective Economic Substance Test

Subjective economic substance requires that a taxpayer demonstrate a business reason for engaging in a transaction other than the tax savings obtained through the transaction. In other words, a taxpayer must have a business purpose for entering into a transaction.[23]

Inc. v. United States, 253 F.3d 350, 354 (8th Cir. 2001); *ACM P'ship*, 157 F.3d at 247–48; *Rice's Toyota World, Inc.*, 752 F.2d at 94.

20. *See* Coltec Indus., Inc. v. United States 454 F.3d 1340, 1355 (Fed. Cir. 2006); Dow Chem. Co. v. United States, 435 F.3d 594, 599 (6th Cir. 2006); Winn-Dixie Stores, Inc. v. Comm'r, 254 F.3d 1313, 1316 (11th Cir. 2001); United Parcel Serv. of Am., Inc. v. Comm'r, 254 F.3d 1014, 1018 (11th Cir. 2001); *Pasternak*, 990 F.2d at 898; Kirchman v. Comm'r, 862 F.2d 1486, 1492 (11th Cir. 1989); Long Term Capital Holdings v. United States, 330 F. Supp. 2d 122, 171 (D. Conn. 2004).

21. *See IES Indus., Inc.*, 253 F.3d at 358; Horn v. Comm'r, 968 F.2d 1229, 1237–38 (D.C. Cir. 1992); *Rice's Toyota World, Inc.*, 752 F.2d at 91–92; Black & Decker Corp. v. United States, 340 F. Supp. 2d 621, 624 (D. Md. 2004), *rev'd in part*, 436 F.3d 431 (4th Cir. 2006).

22. The Senate Finance Committee has issued a detailed description of the economic substance doctrine amendment added to the Heartland, Habitat, Harvest, and Horticulture Act of 2007, which the Committee approved on October 4, 2007. STAFF OF S. COMM. ON FINANCE, 110TH CONG., ECONOMIC SUBSTANCE DOCTRINE (2007), http://senate.gov/~finance/sitepages/leg/LEG%202007/Leg%20110%20100407agamendment.pdf.

23. The business purpose test "examines whether the taxpayer was induced to commit capital for reasons only relating to tax considerations or whether a nontax motive, or legitimate profit motive, was involved." Shriver v. Comm'r, 899 F.2d 724, 726 (8th Cir. 1990); Rice's Toyota World, Inc., 752 F.2d 89, 92 (4th Cir. 1985).

1. Historical Perspective of Business Purpose

Gregory v. Helvering[24] is universally credited with establishing the business purpose doctrine in 1935. The foundations for *Gregory*, however, were developed by the Fifth Circuit in *Pinellas Ice & Cold Storage Co. v. Commissioner*[25] and the Second Circuit in *Cortland Specialty Co. v. Commissioner*,[26] both decided in 1932 and both involving corporate reorganizations. In *Pinellas Ice*, the Fifth Circuit held that, although a transfer satisfied the literal reorganization provision that admittedly covered the particular transaction, the transaction could not be categorized as a reorganization because of the lack of continuity of the business.[27]

The legacy of *Cortland Specialty Co.* provides additional insight. In this case, Judge Augustus Hand, writing for the Second Circuit, determined that a "[r]eorganization presupposes continuance of business under modified corporate forms."[28] Two years later, in 1934, Judge Learned Hand wrote the opinion for the Second Circuit in *Helvering v. Gregory*, in which he concluded that the reorganization provision at issue presumed that the reason for the reorganization was germane to the enterprise itself and not as a by-product of the minimization or deferral of taxes.[29]

On appeal, the Supreme Court, with Justice Sutherland writing, adopted Judge Learned Hand's line of reasoning, concluding that the proposed tax-free reorganization had no business purpose and the transactions at issue were "a mere device which put on the form of a corporate reorganization as a disguise for concealing its real character."[30] Holding that the transactions fell outside the bounds of the "plain intent of the statute,"[31] the Court disregarded the form of the transactions and determined that the transactions were taxable. In doing so, the Court stated, "To hold otherwise would be to exalt artifice above reality and to deprive the statutory provision in question of all serious purpose."[32]

Gregory, as articulated by Judge Hand and adopted by the Supreme Court, began as a continuity of business test. As the decades passed however, this continuity requirement has morphed into the current business purpose requirement.

---

24. 293 U.S. 465 (1935).
25. 57 F.2d 188 (5th Cir. 1932).
26. 60 F.2d 937 (2d Cir. 1932).
27. *Pinellas Ice*, 57 F.2d at 189.
28. *Cortland Specialty Co.*, 60 F.2d at 940.
29. 69 F.2d 809, 811 (2d Cir. 1934), *aff'd*, 293 U.S. 465 (1935). In addition to serving on the Second Circuit together, Judge Augustus Hand and Judge Learned Hand were cousins and close friends. Judge Augustus Hand also sat on the panel that decided *Helvering v. Gregory*.
30. Gregory v. Helvering, 293 U.S. 465, 469 (1935).
31. *Id.* at 470.
32. *Id.*

Initially, *Gregory* was limited to transactions in which a newly formed corporation was liquidated shortly after its creation and its impact was "confined strictly to the facts there presented."[33] In 1943, however, the Tax Court, relying on *Gregory*, held that the absence of a business purpose prevented a corporation from deducting interest payments on a loan made to shareholders that benefited the shareholders in their individual capacities.[34]

As originally developed by Judge Hand, it appears that a business purpose is inherent in the continuity of business enterprise requirement under the reorganization provisions.[35] Subsequently, it appears that Judge Hand attempted to expand the scope of his original creation, dissenting in *Gilbert v. Commissioner*:

> [T]he literal meaning of the words of a statute is seldom, if ever, the conclusive measure of its scope. Except in rare instances statutes are written in general terms and do not undertake to specify all the occasions that they are meant to cover; and their "interpretation" demands the projection of their expressed purpose, upon occasions, not present in the minds of those who enacted them. . . . If, however, the taxpayer enters into a transaction that does not appreciably affect his beneficial interest except to reduce his tax, the law will disregard it; for we cannot suppose that it was part of the purpose of the [Internal Revenue Code] to provide an escape from the liabilities that it sought to impose.[36]

Applying *Gregory*, a business purpose is the subjective condition necessary to satisfy the requirement of business continuity.[37] Taking *Gregory* in light of its subsequent interpretation, a business purpose is the subjective condition necessary to satisfy the doctrine of economic substance.

2. Modern Application of Business Purpose

The business purpose doctrine is a nonstatutory method used by the government to recast transactions that comply with the literal wording of the Internal Revenue Code but are nonetheless prohibited from receiving the tax advantages resulting from those provisions because such tax advan-

33. Bremer v. White, 10 F. Supp. 9, 12 (D. Mass. 1935) (assuming a continuity of business existed and a reorganization took place without proof that the newly created corporation dissolved).

34. *See* The Humko Co. v. Comm'r, T.C.M. (P-H) 43, 511 (T.C. 1943); *see also* Bazley v. Comm'r, 331 U.S. 737 (1947) (recognizing distinction between corporate business purpose and shareholder business purpose).

35. *See* Fuller, *supra* note 11, at 362.

36. Gilbert v. Comm'r, 248 F.2d 399, 411 (2d Cir. 1957) (Hand, J., dissenting).

37. *See* Fuller, *supra* note 11, at 362.

tages were not intended by the Internal Revenue Code.[38] From its initial narrow interpretation, the scope of *Gregory* and its "business purpose" test has grown such that the use of the term evokes a range of meanings that promotes greater vagueness. It is the very vagueness in the definition of business purpose that makes it a more adaptable definition.[39]

Business purpose is an attempt to reach the correct result based on the specifics of a transaction in light of the specifics of a particular business, considering the motives of the taxpayer and whether the transaction served a useful nontax purpose. The Tax Court, in *ACM Partnership*, described the subjective business purpose doctrine:

> [T]he transaction must be rationally related to a useful nontax purpose that is plausible in light of the taxpayer's conduct and useful in light of the taxpayer's economic situation and intentions. Both the utility of the stated purpose and the rationality of the means chosen to effectuate it must be evaluated in accordance with commercial practices in the relevant industry. A rational relationship between purpose and means ordinarily will not be found unless there was a reasonable expectation that the nontax benefits would be at least commensurate with the transaction costs.[40]

Business purpose also recognizes that taxes play a major part in business decisions and corporate behavior. Business purpose does not exclude taxes from consideration but instead seeks to impose limits on their impact.[41] It does not permit a "high-stakes game in which taxpayers bet transaction costs against their ability to find and exploit anomalies in the Code and regulations."[42]

According to the Internal Revenue Service, a determination of whether a transaction has "business purpose" must consider a number of factors:[43]

> (i) whether a profit was even possible;

38. *See id.* at 365–66.

39. *See* Walter J. Blum, *Motive, Intent, and Purpose in Federal Income Taxation*, 34 U. CHI. L. REV. 485, 495 (1967).

40. ACM P'ship v. Comm'r, 73 T.C.M. (CCH) 2189, 2217 (T.C. 1997) (citing Cherin v. Comm'r, 89 T.C. 986, 993–94 (T.C. 1987)), *aff'd in part and rev'd in part*, 157 F.3d 231 (3d Cir. 1998).

41. *See* Robert Thornton Smith, *Business Purpose: The Assault Upon the Citadel*, 53 TAX LAW. 1, 9 (1999).

42. McMahon, Jr., *supra* note 3, at 196; *see also* Ellen P. Aprill, *Tax Shelters, Tax Law, and Morality: Codifying Judicial Doctrines*, 54 SMU L. REV. 9 (2001); Mark P. Gergen, *The Common Knowledge of Tax Abuse*, 54 SMU L. REV. 131 (2001).

43. The factors are illustrative of how the Internal Revenue Service views the subjective economic substance test. While beyond the scope of this Article, it appears that factors (i) and (iv) relate to the objective determination of economic substance and should not be included in the calculus of whether the business purpose test is satisfied.

(ii) whether the taxpayer had a nontax business reason to engage in the transaction;

(iii) whether the taxpayer, or its advisors, considered or investigated the transaction, including market risk;

(iv) whether the taxpayer really committed capital to the transaction;

(v) whether the entities involved in the transaction were entities separate and apart from the taxpayer and engaging in legitimate business before and after the transaction;

(vi) whether all the purported steps were engaged in at arms-length with the parties doing what the parties intended to do; and

(vii) whether the transaction was marketed as a tax shelter in which the purported tax benefits significantly exceeded the taxpayer's actual investment.[44]

In attempting to demonstrate that these factors favoring a business purpose are not satisfied, the Internal Revenue Service may attempt to show the following:

(i) documents or other evidence that the transactions at issue were sold as tax shelters with limited consideration of the underlying economics of the transaction;

(ii) evidence that the taxpayer, or its advisors, did not investigate the market risk prior to entering into the transaction;

(iii) evidence that the independent parts making up the transaction were not entered into at arm's length; and

(iv) evidence that a prudent investor would have or could have accomplished similar objectives using much simpler or more direct methods.[45]

44. Donald L. Korb, Remarks at the 2005 University of Southern California Tax Institute: The Economic Substance Doctrine in the Current Tax Shelter Environment 9–10 (Jan. 25, 2005), *available at* http://www.irs.gov/pub/irs-utl/economic_substance_(1_25_05).pdf.

I do not argue that these factors are necessarily the appropriate factors, nor that these factors constitute an exclusive list. For my purposes, these factors demonstrate potential considerations that may be relevant to a determination of business purpose.

45. *Id.* at 14.

These types of considerations inevitably involve written documentation concerning the transaction. Simply because the benefits are achieved does not vitiate the subjective element of business purpose. While the Internal Revenue Service may consider the relationship between the tax benefits and the amount of taxable income to be reduced,[46] more important is whether there is also a business purpose to the transaction, demonstrated through competent evidence that nontax considerations were evident in considering the transaction.

B. Objective Economic Substance Test

The objective economic substance test requires a taxpayer to demonstrate that its economic position is enhanced by engaging in the transaction.[47] However, clairvoyance is not required to satisfy the objective element. The transaction is not required to actually produce a profit to satisfy this test; instead, it is the *potential* for profit that is measured.[48] Profit potential applies a reasonable person standard—in this case a reasonable businessman standard—by determining whether a potential for profit exists such that a reasonable businessman would invest in the venture based on standards applicable to the particular industry.[49] As a result, the transaction must be potentially profitable to satisfy this inquiry.

The objective economic substance test does not determine liability based on a subjective thought process, but it also does not grant favorable

46. *See id.* at 13.

47. *See, e.g.*, Rice's Toyota World, Inc. v. Comm'r 752 F.2d 89, 94 (4th Cir. 1985) (explaining that the economic substance inquiry requires an objective determination of whether a reasonable possibility of profit from the transaction existed apart from tax benefits); *see also* Compaq Computer Corp. v. Comm'r, 277 F.3d 778, 781 (5th Cir. 2001); IES Indus., Inc. v. United States, 253 F.3d 350, 354 (8th Cir. 2001).

48. *See* Knetsch v. United States, 364 U.S. 361, 364–65 (1960); Goldstein v. Comm'r, 364 F.2d 734, 739 (2d Cir. 1966); Abramson v. Comm'r, 86 T.C. 360, 371 (T.C. 1986). Some courts have applied the economic substance doctrine to deny the tax benefits claimed if economic risks and profit potential, while existent, were nonetheless insignificant compared to the tax benefits. *See, e.g.*, *Goldstein*, 364 F.2d at 739–40 (disallowing deduction even though the taxpayer had a possibility of small gain or loss by owning Treasury bills); Sheldon v. Comm'r, 94 T.C. 738, 768 (T.C. 1990) ("[P]otential for 'gain' . . . is infinitesimally nominal and vastly insignificant when considered in comparison with the claimed deductions."); *see also* Blum, *supra* note 39, at 499 ("[T]here may currently be no tax rules which specifically provide that classification of an action is to be postponed until the outcome of the action is known.").

49. *See* Cherin v. Comm'r, 89 T.C. 986, 994 (T.C. 1987). A Monte–Carlo analysis is a statistical technique that identifies probabilities associated with particular outcomes and is helpful in decision-making processes. *See* Steve Pomerantz, *Securities Insight for Attorneys Monte-Carlo Analysis: A Tool for Evaluating Investment Returns*, NYSBA COMM'L & FED. LITIG. SECTION NEWSLETTER, Fall/Winter 2005, at 3. This model has been applied in a number of tax cases to determine acceptable industry standards. *See* Boca Investerings P'ship v. United States, 167 F. Supp. 2d 298, 357 (D.D.C. 2001), *rev'd*, 314 F.3d 625 (D.C. Cir. 2003); S. Cal. Fed. Sav. & Loan Ass'n v. United States, 57 Fed. Cl. 598, 636–37 (Fed. Cl. 2003), *aff'd in part, rev'd in part*, 422 F.3d 1319 (Fed. Cir. 2005); Bank One Corp. v. Comm'r, 120 T.C. 174, 231–32 (T.C. 2003), *aff'd in part, vacated in part sub nom.* JP Morgan Chase & Co. v. Comm'r, 458 F.3d 564 (7th Cir. 2006).

tax consequences based on the mere absence of thought.[50] Consequently, self-induced ignorance does not create a tax advantage.

Demonstrating that the economic position of a taxpayer is enhanced by entering into a transaction can avoid the need to inquire into the subjective intentions of a taxpayer because the answer to that inquiry can be determined on the basis of the facts of the transaction itself.[51] Necessarily, business purpose is absent from such a determination because objective economic substance lacks an inquiry into motives of a transaction but instead views the transaction objectively. This, of course, does not preclude credibility determinations by a trier of fact concerning whether certain facts should be considered as part of the transaction that is to be measured objectively. For example, in answering the question of whether there is an appreciable enhancement in the net economic position of the taxpayer, a factual inquiry into whether there was a commitment to invest future cash in a corporate-owned life insurance (COLI) plan does not alter the objectiveness of the economic substance test.[52]

The underlying concepts of business purpose and profit potential define the measure by which trial courts initially resolve whether a transaction has economic substance in addition to satisfying the literal terms of the Internal Revenue Code.[53] Thus, to qualify for the benefits permitted under the Code, the twin prongs of the economic substance doctrine raise an additional impediment that a taxpayer must navigate to otherwise obtain the tax benefits expressly permitted by the tax code.[54] The standard of review on appeal will determine how an appellate court approaches a determination by a trial court that considers the economic substance doctrine.

II. STANDARDS OF REVIEW

Standards of review define the scope of power between judicial actors each functioning within a statutory or rule-based scheme and carrying out

50. *See* Comm'r v. Wilson, 353 F.2d 184, 187 (9th Cir. 1965) (explaining that "so much in the way of liability for taxes can hardly be allowed to depend solely upon what goes on in someone's mind"); *see also* Blum, *supra* note 39, at 519–20.

51. Subjective inquires would still exist if the conjunctive test applied because both the subjective and objective economic substance tests would be applied. If, however, a disjunctive test was applied, the satisfaction of the objective test necessarily eliminates an inquiry into the subjective intentions of the taxpayer. *But see* Petition for Writ of Certiorari, *Dow*, *supra* note 4, at *27–28; John B. Magee & Gerald Goldman, *Uncut Gems: Judicial Review in Economic Substance Appeals*, 116 TAX NOTES 481, 482 n.9 (2007) ("Even under the objective prong of the two-prong test, the taxpayer's subjective business plans logically must be considered simply to establish the terms of the transaction whose objective economic substance is to be assessed.").

52. *See* Petition for Writ of Certiorari, *Dow*, *supra* note 4, at *2.

53. *See Knetsch*, 364 U.S. at 367.

54. *See* Lerman v. Comm'r, 939 F.2d 44, 52 (3d Cir. 1991) ("[I]t is settled federal tax law that for transactions to be recognized for tax purposes they must have economic substance. Therefore, economic substance is a *prerequisite* to the application of any Code provisions allowing deductions . . . .").

specified responsibilities under that system.[55] Because the selection of a standard of review often dictates the ultimate resolution of a case, the framework for selecting the applicable standard is significant.[56] A standard of review reflects the degree to which the original decision maker must be wrong for a reviewer to reverse the original decision.[57] Traditionally, judicial review utilizes three standards—de novo, clearly erroneous, and abuse of discretion.[58]

De novo review applies to questions of law.[59] It is the least restrictive standard of review as it calls for no degree of deference and permits a reviewing court to determine the correct resolution of an issue on its own accord.[60] In essence, de novo review provides no deference but is a judicial determination of an issue entirely independent of the prior resolution.[61] Under a de novo standard, a reviewing court is "willing to reverse [a prior] conclusion of law solely on the basis that it believes that conclusion to be incorrect."[62]

The clearly erroneous standard of review applies to questions of fact and gives a significant amount of deference to the trier of fact.[63] Under the clearly erroneous standard of review, "[a] finding is 'clearly erroneous' when although there is evidence to support it, the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed."[64] In other words, a reviewing court will not substi-

55. *See* Edward H. Cooper, *Civil Rule 52(a): Rationing and Rationalizing the Resources of Appellate Review*, 63 NOTRE DAME L. REV. 645, 645 (1988) ("Rule 52(a) serves a vital institutional role in allocating the responsibility and the power of decision between district courts and the courts of appeals."); *see also* Ronald R. Hofer, *Standards of Review—Looking Beyond the Labels*, 74 MARQ. L. REV. 231, 232 (1991) (describing standards of review as the level of deference that appellate courts provide to trials courts); Alvin B. Rubin, *The Admiralty Case on Appeal in the Fifth Circuit*, 43 LA. L. REV. 869, 873 (1983) ("[S]tandards of review . . . indicate the decibel level at which the appellate advocate must play to catch the judicial ear.").

56. A number of appellate courts have reversed determinations by the trial court using a de novo standard of review. Conversely, courts that applied the clearly erroneous test affirmed the decision of the trial court. *See* Magee & Goldman, *supra* note 51, at 482 n.15.

57. *See* Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115, 1126 (1987).

58. *See* Elder v. Holloway, 510 U.S. 510, 516 (1994) (de novo); United States v. U.S. Gypsum Co., 333 U.S. 364, 395 (1948) (clearly erroneous); Gen. Elec. Co. v. Joiner, 522 U.S. 136, 141–42 (1997) (abuse of discretion).

59. *See, e.g.*, *Elder*, 510 U.S. at 516.

60. *See* Salve Regina Coll. v. Russell, 499 U.S. 225, 238 (1991).

61. RICHARD J. PIERCE, JR., SIDNEY A. SHAPIRO & PAUL R. VERKUIL, ADMINISTRATIVE LAW AND PROCESS 370 (1985).

62. *Id.*

63. *See U.S. Gypsum Co.*, 333 U.S. at 395; *see also* Anderson v. City of Bessemer City, 470 U.S. 564, 573–74 (1985); 9C CHARLES ALAN WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE AND PROCEDURE § 2585, at 368–74 (3d ed. 2008); *id.* § 2588, at 443–44; Irving Wilner, *Civil Appeals: Are They Useful in the Administration of Justice?*, 56 GEO. L.J. 417, 430 (1968) ("[T]he conventional fact-law distinction . . . is the purported foundation of appellate review, the archpremise which has woven review into the fabric of the administration of justice since the fifteenth century.").

64. United States v. U.S. Gypsum Co., 333 U.S. 364, 395 (1948); *see also Anderson*, 470 U.S. at 573 (explaining that the clearly erroneous standard "plainly does not entitle a reviewing court to

tute its judgment for the judgment of the trier of fact notwithstanding that had the appellate court been sitting as fact finder, it might have reached a different finding.[65] Any determination of the trial court under this standard carries with it significant weight.[66] Consequently, the role of the appellate court under a clearly erroneous standard of review is "not to decide factual issues *de novo*."[67]

The third standard of review, abuse of discretion, provides the highest degree of deference to a determination by the trial court. Under this standard, an abuse of discretion occurs when an adjudicator fails to exercise sound, reasonable, and legal decision-making skills.[68] This standard applies to the discretionary functions of a trial court and has been applied by appellate courts reviewing decisions such as the adequacy or excessiveness of jury verdicts, the exclusion of scientific evidence, evidentiary rulings, estoppel, sanctions, and attorneys' fees.[69] Consequently, the abuse of discretion standard of review is not a standard of review that appellate courts would articulate in reviewing a trial court decision regarding the economic substance doctrine.

A. Historical Perspectives

The scope of review by appellate courts is not a new question. Beginning with the Constitution over 200 years ago, the question first surfaced

reverse the finding of the trier of fact simply because it is convinced that it would have decided the case differently").

65. *See Anderson*, 470 U.S. at 574 ("[If the district court's finding] is plausible in light of the record viewed in its entirety, the [reviewing court] may not reverse it even though convinced that had it been sitting as the trier of fact, it would have weighed the evidence differently.").

66. *See U.S. Gypsum Co.*, 333 U.S. at 395; *see also Anderson*, 470 U.S. at 573–74; NAACP v. Fordice, 252 F.3d 361, 365 (5th Cir. 2001); Nolan v. Golden Rule Ins. Co., 171 F.3d 990, 992 (5th Cir. 1999); Miller v. Mercy Hosp., Inc., 720 F.2d 356, 360–61 (4th Cir. 1983). Utilizing such a standard, an appellate court must accept the trier of fact's determination of facts as correct unless clear evidence suggests an alternative. *See U.S. Gypsum Co.*, 333 U.S. at 395.

67. Zenith Radio Corp. v. Hazeltine Research, Inc., 395 U.S. 100, 123 (1969) (explaining that where there are two permissible views of the evidence, the fact finder's choice between them cannot be clearly erroneous); *see also* United States v. Yellow Cab Co., 338 U.S. 338, 342 (1949) ("Such a choice between two permissible views of the weight of evidence is not 'clearly erroneous.'"); Inwood Labs., Inc. v. Ives Labs., Inc., 456 U.S. 844, 855 (1982). In more colorful language, the Seventh Circuit noted that the clearly erroneous standard permitted reversal only if the decision "strike[s] [the court] as wrong with the force of a five-week-old, unrefrigerated dead fish." Parts & Elec. Motors, Inc. v. Sterling Elec., Inc., 866 F.2d 228, 233 (7th Cir. 1988).

68. *See* Gen. Elec. Co. v. Joiner, 522 U.S. 136, 140–44 (1997); Wilton v. Seven Falls Co., 515 U.S. 277, 289 (1995).

69. *See* Richard H. W. Maloy, *"Standards of Review" – Just a Tip of the Icicle*, 77 U. DET. MERCY L. REV. 603, 630 nn.238–39, 631 nn.240, 242 & 244–45 (2000) (citing Gasperini v. Ctr. For Humanities, Inc., 518 U.S. 415, 438 (1996) (inadequacy or excessiveness of jury verdicts); *Joiner*, 522 U.S. at 141–43 (exclusion of scientific evidence); Koster v. Trans World Airlines, Inc., 181 F.3d 24, 32 (1st Cir. 1999) (evidentiary rulings); Klein v. Stahl GMBH & Co. Maschinefabrik, 185 F.3d 98, 108 (3d Cir. 1999) (judicial estoppel); Cooter & Gell v. Hartmarx Corp., 496 U.S. 384, 405 (1990) (sanctions imposed); Spegon v. Catholic Bishop, 175 F.3d 544, 550 (7th Cir. 1999) (award of attorneys' fees as compensation)).

regarding a provision that provides that the "supreme Court shall have appellate Jurisdiction, both as to Law and Fact, with such Exceptions, and under such Regulations as the Congress shall make."[70]

Historically, suits initiated under the common law were heard by juries and suits in equity were heard by judges sitting without a jury.[71] In the equity context, because cases were presented through depositions and interrogatories completed by the parties outside of the presence of the trial court, an appellate court was just as capable of making decisions on factual issues as was the trial court.[72] As a result, appellate courts possessed unlimited power to review the entire record of the trial court and considered factual and legal questions independently of a determination by a trial judge.[73]

In 1912, trial courts sitting in equity began to hear oral testimony pursuant to the Federal Equity Rules, and appellate courts, recognizing that the previous de novo method of review placed them at a disadvantage compared to the trial judges who heard the testimony, abandoned the de novo standard in favor of a standard giving great weight to the findings of trial judges.[74] This self-imposed limitation was based principally on the superior ability of trial judges to make credibility determinations. Because the trial judge heard the testimony that served as the basis for findings of fact, those findings were not disturbed unless such findings were clearly erroneous.[75]

---

70. U.S. CONST. art. III, § 2, cl. 2. The scope of review of jury determinations has historically been of equal interest. The Sixth Amendment provides that "[i]n all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed," U.S. CONST. amend. VI, while the Seventh Amendment provides that "[i]n Suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved, and no fact tried by a jury, shall be otherwise re-examined in any Court of the United States, than according to the rules of the common law." U.S. CONST. amend. VII. In this context, a jury makes credibility determinations, determines the weight applied to evidence presented, and draws inferences from the evidence presented. *See* Robert L. Stern, *Review of Findings of Administrators, Judges and Juries: A Comparative Analysis*, 58 HARV. L. REV. 70, 73–74 (1944). This construction permits a body of persons representing the community, without particular legal expertise, to make factual determinations. *See id.* at 81; *see also* Francis A. Bohlen, *Mixed Questions of Law and Fact*, 72 U. PA. L. REV. 111, 116 (1924) (contrasting judicial rulings and jury determinations). The scope of this Article, however, is restricted to facts determined by a trial judge.

71. Stern, *supra* note 70, at 79.

72. *See* Charles E. Clark & Ferdinand F. Stone, *Review of Findings of Fact*, 4 U. CHI. L. REV. 190, 204 (1937); *see also* Stern, *supra* note 70, at 79.

73. *See* 9 JAMES WM. MOORE, MOORE'S FEDERAL PRACTICE ¶ 52 app. 100 (3d ed. 2008); *see also* District of Columbia v. Pace, 320 U.S. 698, 701–02 (1944); Morley Constr. Co. v. Md. Cas. Co., 300 U.S. 185, 191 (1937); Keller v. Potomac Elec. Power Co., 261 U.S. 428, 444 (1923); La Abra Silver Mining Co. v. United States, 175 U.S. 423, 464–65 (1899).

74. FED. R. EQ. 46, 226 U.S. 649 (1912) (repealed 1938); *see also* Clark & Stone, *supra* note 72, at 203–04; Butte & Superior Copper Co. v. Clark-Montana Realty Co., 249 U.S. 12, 30 (1919) (recognizing that findings of fact by a trial judge were presumptively correct and were not disturbed on appeal unless clearly wrong); Baker v. Schofield, 243 U.S. 114, 118 (1917) ("[W]hen two courts have reached the same conclusion on a question of fact, their finding will not be disturbed unless it is clear that their conclusion was erroneous.").

75. *See* Stern, *supra* note 70, at 113.

Thus, the predecessor to the "clearly erroneous" rule that was later adopted in the Federal Rules of Civil Procedure applied to findings based on conflicting testimony or facts derived or inferred from such testimony. In cases of facts derived or inferred from uncontradicted evidence, documentary evidence, or deposition testimony, appellate courts did not consider themselves bound by trial court determinations because they saw themselves as just as qualified to make such determinations.

The distinctions between law and equity were confused at best, and in 1934, a new of set of rules attempted to eliminate such distinctions. Initially, the draft version of the Federal Rules of Civil Procedure provided that facts determined by a judge (as opposed to a jury) would "'have the same effect as that heretofore given to findings in suits in equity.'"[76] The draft rule prompted a discussion that highlighted the distinction between the heightened deference standard that was applied to findings by a jury at law and the lower standard of review given in equity practice.[77] The rule ultimately was drafted using the clearly erroneous language, which applied to appellate review of findings at equity based on oral testimony.[78]

The Advisory Committee note accompanying Rule 52 provided that the clearly erroneous standard of review should apply whether the finding of fact was determined based on conflicting testimony, or "deduced or inferred from uncontradicted testimony."[79] According to William D. Mitchell, former Attorney General and the Chairman of the Advisory Committee appointed by the Supreme Court to draft the Federal Rules of Civil Procedure, the rule that

> appl[ies] to all cases tried by a judge without a jury is practically the modern equity rule; it is not the ancient equity rule which allowed trial *de novo* on the facts in the appellate court—it is a limited provision—the appellate court may not set aside the findings unless they are clearly erroneous or against the clear weight of evidence.[80]

As a result of this rule, an appellate court must examine the entire record and determine for itself the appropriate findings of fact, subject to the major restriction that the appellate court must accept as practically conclusive findings of fact by a trial judge that are based on the credibility

76. 9C WRIGHT & MILLER, *supra* note 63, § 2571, at 225 n.16 (quoting Draft Rule 68).
77. *See* Cooper, *supra* note 55, at 648.
78. *Id.*
79. FED. R. CIV. P. 52 advisory committee's notes (1937).
80. William D. Mitchell, Address at the Symposium at New York City (Oct. 17–19, 1938), *in* FEDERAL RULES OF CIVIL PROCEDURE: PROCEEDINGS OF THE INSTITUTE AT WASHINGTON, D.C., OCTOBER 6, 7, 8, 1938, AND OF THE SYMPOSIUM AT NEW YORK CITY, OCTOBER 17, 18, 19, 1938, at 287 (Edward H. Hammond ed., 1939).

of witnesses.[81] Should the appellate court disagree with the findings of the trial court, it will reverse only if it is convinced that the findings of the trial court are "unquestionably wrong."[82]

Thus, prior to the enactment of the Federal Rules of Civil Procedure, particular emphasis was given to findings based on the testimony of witnesses. In cases in which the evidence was contradictory, appellate courts did not disturb the findings of the trial judge.[83] By contrast, if the findings were either undisputed or were based on a factual determination consisting solely of documentary evidence, the appellate court sat in a comparable position to the trial court.[84] The Federal Rules provided that the findings of fact by a trial judge are to be accepted unless clearly erroneous.[85] It was not explicit, however, that this standard should be applied to documentary evidence.[86] Consequently, the only major modification to Rule 52(a) since its adoption in 1938 was adopted in 1985, adding language that the standard of review—clearly erroneous—applied to findings of fact "whether based on oral or documentary evidence."[87] With this change, it became clear that the clearly erroneous standard of review applied not only to credibility determinations but also to facts derived from oral and documentary evidence, an obvious signal that institutional responsibilities of trial courts and appellate courts should be respected.

B. Fact–Law Distinction

Traditionally, the level of appellate review hinged on the distinction between law and fact. Application of this distinction to appellate review is a rather simplistic exercise—appellate courts are free to review legal conclusions de novo, and factual findings are allowed a level of deference

81. Stern, *supra* note 70, at 89.
82. *See id.*
83. *See id.* at 112.
84. *See id.* at 112–13.
85. FED. R. CIV. P. 52(a)(6).
86. *See* United States v. Gen. Motors Corp., 384 U.S. 127, 141 n.16 (1966), which seemed to raise a question concerning the applicability of Rule 52 to factual findings based on documentary evidence:

> [T]he trial court's customary opportunity to evaluate the demeanor and thus the credibility of the witnesses, which is the rationale behind Rule 52(a) . . . plays only a restricted role here. This was essentially a "paper case." It did not unfold by the testimony of "live" witnesses. Of the 38 witnesses who gave testimony, only three appeared in person. The testimony of the other 35 witnesses was submitted either by affidavit, by deposition, or in the form of an agreed-upon narrative of testimony given in the earlier criminal proceeding before another judge. A vast number of documents were also introduced, and bear on the question for decision.

87. *See* FED. R. CIV. P. 52 advisory committee's notes (1985). In 2007, Rule 52(a) was changed for stylistic purposes. The rule is now set forth in Rule 52(a)(6) and provides: "Findings of fact, whether based on oral or other evidence, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial court's opportunity to judge the witnesses' credibility." FED. R. CIV. P. 52(a)(6).

such that those findings, whether based on oral or documentary evidence, shall not be set aside unless clearly erroneous.[88] While seemingly straightforward, the distinction between a factual finding and a legal conclusion is often murky.[89] The systematic difficulty is in describing what constitutes fact, what constitutes law, and what constitutes both.

Just as defining the level of deference under the clearly erroneous standard to an objective certainty is impractical,[90] defining the distinction between fact and law is just as unworkable.[91] While courts often create sound bites and elaborate musings on the definition of each, the distinction cannot be articulated to an objective certainty.[92] Professor Clark noted this dilemma:

> There seems to be no hard and fast distinction between questions of law and questions of fact; . . . and since there are so many cases on the border line between ultimate facts and conclusions of

88. *See, e.g.*, Elder v. Holloway, 510 U.S. 510, 516 (1994); United States v. U.S. Gypsum Co., 333 U.S. 364, 395 (1948).

89. *See* Nathan Isaacs, *The Law and the Facts*, 22 COLUM. L. REV. 1, 1 (1922) (defining the distinction between law and fact in terms of "delusive simplicity"); *see also* Carissa Byrne Hessick & F. Andrew Hessick, *Appellate Review of Sentencing Decisions*, 60 ALA. L. REV. 1, 14–18, 25 (2008) (describing the difficult distinction in the context of criminal sentencing review).

90. *See* Christopher M. Pietruszkiewicz, *Conflating Standards of Review in the Tax Court: A Lesson in Ambiguity*, 44 HOUS. L. REV. 1337, 1370–71 (2008) ("When Congress does not utilize terminology with which courts interact frequently, thereby providing conflicting messages, consistent application of a standard is not possible. While it is impossible to pinpoint the definition of preponderance of the evidence, clear and convincing evidence, and beyond a reasonable doubt, courts are nonetheless able to apply these standards because they are universal and commonly articulated."); *id.* at 1363 ("Due to the imprecision of language and the unrealistic view that infinite degrees of deference can exist, standards of review should be limited to a familiar set of gradations."); *see also* Clermont, *supra* note 57, at 1148.

91. *See* Dobson v. Comm'r, 320 U.S. 489, 500–01 (1943); JOHN DICKINSON, ADMINISTRATIVE JUSTICE AND THE SUPREMACY OF LAW IN THE UNITED STATES 55 (1927); CARL MCFARLAND, JUDICIAL CONTROL OF THE FEDERAL TRADE COMMISSION AND THE INTERSTATE COMMERCE COMMISSION 1920–1930, at 25 (1933); ATT'Y GEN.'S COMM. ON ADMIN. PROCEDURE, ADMINISTRATIVE PROCEDURE IN GOVERNMENT AGENCIES, S. DOC. NO. 77-8, at 88 (1st Sess. 1941); Ray A. Brown, *Fact and Law in Judicial Review*, 56 HARV. L. REV. 899 (1943); Robert M. Cooper, *Administrative Justice and the Role of Discretion*, 47 YALE L.J. 577, 588–90 (1938); John Dickinson, *Judicial Review of Administrative Determinations, a Summary and Evaluation*, 25 MINN. L. REV. 588 (1941); Isaacs, *supra* note 89, at 1; James M. Landis, *Administrative Policies and the Courts*, 47 YALE L.J. 519, 531–32 (1938); Randolph Paul, Dobson v. Commissioner*: The Strange Ways of Law and Fact*, 57 HARV. L. REV. 753, 803 (1944).

92. *See* Cooper, *supra* note 55, at 645; *see also* Tupman v. Haberkern, 280 P. 970, 973 (Cal. 1929) (stating that questions of fact are decided by the trial court and that questions of law are decided by the appellate court). Rule 52 of the Federal Rules of Civil Procedure also makes this distinction, providing that "[f]indings of fact, whether based on oral or other evidence, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial court's opportunity to judge the witnesses' credibility." FED. R. CIV. P. 52(a)(6). By negative implication, Rule 52(a) promotes independent review of legal questions. *See* Pullman–Standard v. Swint, 456 U.S. 273, 287–88 (1982).

> facts, *i.e.*, law, the way is open to a court with understanding to accomplish substantial justice, whatever the formula.[93]

The use of fact and law distinctions allow appellate courts to cast questions as either of fact or of law depending on whether the appellate court ultimately favors reversing, or affirming, the trial court.[94] As a result, appellate courts can produce a "correct result" by utilizing the fact–law distinction and the corresponding standard of review to achieve its design.[95] In other words, based on the less than clearly defined line between law and fact, an appellate court can easily review and decide a factual issue by simply recasting it as a question of law and applying a de novo standard of review instead of the more stringent clearly erroneous standard.[96] As Professor Wright laments:

> The principal means by which appellate courts have obtained such complete control of litigation has been the transmutation of specific circumstances into questions of law. . . . [And] unless the appellate judge handling the case is a dullard, some doctrine is always at hand to achieve the ends of justice, as they appear to the appellate court.[97]

---

93. Clark & Stone, *supra* note 72, at 211 n.93; *see also* Louis L. Jaffe, *Judicial Review: Question of Law*, 69 HARV. L. REV. 239, 239–40 (1955) (defining the fact–law distinction as imperceptibly blending into each other).

94. *See, e.g.*, Paul L. Caron, *Tax Myopia Meets Tax Hyperopia: The Unproven Case of Increased Judicial Deference to Revenue Rulings*, 57 OHIO ST. L.J. 637, 644–45 (1996); Russell L. Weaver, *A Foolish Consistency Is the Hobgoblin of Little Minds*, 44 BAYLOR L. REV. 529, 553 (1992); *see also* Quintin Johnstone, *An Evaluation of the Rules of Statutory Interpretation*, 3 U. KAN. L. REV. 1, 5 (1954) (courts "pull respectable-sounding rules to justify any possible result"); Christopher M. Pietruszkiewicz, *Discarded Deference: Judicial Independence in Informal Agency Guidance*, 74 TENN. L. REV. 1, 8 (2006).

95. Caron, *supra* note 94, at 644–45 ("The precise verbal formulation used by a court is mere window-dressing that does not have any effect on the ultimate resolution of the case."); Richard J. Pierce, Jr., *Legislative Reform of Judicial Review of Agency Actions*, 44 DUKE L.J. 1110, 1110 (1995) ("[A] court often can write an opinion that reverses a major agency action as easily as it can write an opinion that upholds the same action."); *see also* Pietruszkiewicz, *supra* note 94, at 8; Patricia M. Wald, *The Rhetoric of Results and the Results of Rhetoric: Judicial Writings*, 62 U. CHI. L. REV. 1371, 1391–92 (1995) (contrasting the tones of two opinions written in the same circuit, one of which upheld a NLRB determination and another overturning a NLRB determination).

96. *See* David Frisch, *Contractual Choice of Law and the Prudential Foundations of Appellate Review,* 56 VAND. L. REV. 57, 92 (2003); Paul, *supra* note 91, at 811–12 ("'[W]hen the courts are unwilling to review, they are tempted to explain by the easy device of calling the question one of fact; and when otherwise disposed, they say that it is a question of law.'" (internal quotation marks omitted) (quoting Dickinson, *supra* note 91, at 55)); Charles Alan Wright, *The Doubtful Omniscience of Appellate Courts*, 41 MINN. L. REV. 751, 751 (1957) (arguing that appellate courts can control outcome thereby diminishing significance of trial courts).

97. Wright, *supra* note 96, at 751; *see also* Evan Tsen Lee, *Principled Decision Making and the Proper Role of Federal Appellate Courts: The Mixed Questions Conflict*, 64 S. CAL. L. REV. 235, 236 (1991) ("[T]he labels 'law' and 'fact' often amount to little more than divisions of decision making authority between judges and juries or between appellate courts and trial courts."); Stephen A. Wein-

This is not a new problem, and the adoption of the Federal Rules of Civil Procedure had no impact on resolving the distinction, instead creating the hybrid form of a "mixed question."[98] In fact, proper application of such mixed questions was debated as early as 1786.[99]

Courts, however, have not grasped the distinction—or more appropriately, the lack of distinction—between law and fact and continue to maintain the misguided illusion that the standard of review is dictated by whether the question presented is a determination of law or a determination of fact.[100] This perception continues notwithstanding that the Supreme Court in 1944 alluded to the lack of a clear dividing line, questioning whether review by an appellate court related to a question of fact or a question of law and stating that the distinction is "never self-executing."[101] This trend continued in *Pullman–Standard v. Swint*, where the Court described the distinction between fact and law as "vexing,"[102] and in *Bose Corp. v. Consumers Union of United States, Inc.*, suggesting that the fact–law distinction "varies according to the nature of the substantive law at issue."[103]

Professor Cooper sums up the application of a fact–law determination, noting that "[t]he fundamental secret is out, and notoriously so. Characterization of an issue of law application as fact or law for purposes of identifying a formalized standard of review depends on the perceived need for review, not on the actual status of the issue."[104] This follows a similar conclusion in 1927 that "any factual state or relation which the courts conclude to regard as sufficiently important to be made decisive for all subsequent cases of similar character becomes thereby a matter of law for formulation by the court."[105]

---

er, *The Civil Nonjury Trial and the Law–Fact Distinction*, 55 CAL. L. REV. 1020, 1022 (1967) ("Since law application cannot be meaningfully described as either lawmaking or factfinding, such terminology is not a useful analytical tool in answering the question confronting the court.").

98. *Compare* Walling v. Plymouth Mfg. Corp., 139 F.2d 178 (7th Cir. 1943) (bound by "clearly erroneous" standard in reviewing determination of a trial judge of facts to statute), *and* Gary Theatre Co. v. Columbia Pictures Corp., 120 F.2d 891 (7th Cir. 1941), *with* Murray v. Noblesville Milling Co., 131 F.2d 470 (7th Cir. 1942) (not bound by "clearly erroneous" standard in reviewing determination of a trial judge of facts to statute), *and* United States v. Anderson, 108 F.2d 475 (7th Cir. 1939).

99. *See* JAMES BRADLEY THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW 224 (Boston, Little, Brown, and Co. 1898) (describing Lord Mansfield's use of the term "mixed question of law and fact"); Stern, *supra* 70, at 110.

100. *See* Paul D. Carrington, *The Power of District Judges and the Responsibility of Courts of Appeals*, 3 GA. L. REV. 507, 518 (1969) ("[F]indings of fact may be defined as the class of decisions we choose to leave to the trier of fact subject only to limited review, while conclusions of law are the class of decisions which reviewers chose to make for themselves without deference to the judgment of the trial forum.").

101. Baumgartner v. United States, 322 U.S. 665, 671 (1944).

102. 456 U.S. 273, 288 (1982).

103. 466 U.S. 485, 501 n.17 (1984).

104. Cooper, *supra* note 55, at 660.

105. DICKINSON, *supra* note 91, at 312; s*ee also* Brown, *supra* note 91, at 904–05, 911; Paul,

C. Division of Responsibility in Fact–Law Analysis

If the fact–law distinction is merely a guise that permits an appellate court to create a deference standard that suits its view of a particular case, a more appropriate framework of appellate review should consist of an allocation of institutional responsibility between trial courts and appellate courts. Legal Realists and Legal Proceduralists share the jurisprudential theme of discretion in decision making. Realists recognize discretion as a key component of the decisional process and provide a "rational basis for . . . justifying decisions."[106] Necessarily, Realists reject the notion of legal formalism in favor of judges having the ability to decide cases in contradictory ways and subsequently finding adequate grounds for reaching such a result after the fact.[107] Conversely, Proceduralists focus attention on the institution and structure of decision making and believe that the function of discretion is to allocate responsibility among those decision makers.[108] While Realists focus on the individual thought process of reaching a result,[109] Proceduralists focus on the relationship between the decision maker and the institution in an effort to determine the correct answer in law.[110]

Because of institutional factors and role of the appellate courts within the judicial hierarchy, the fact–law distinction necessarily involves discretion. The question then is how much discretion is appropriate for each decision maker in the judicial hierarchy.[111] The answer cannot be based on a mere reflection of the current judicial practice, defined by the imprecise rules regarding what is fact and what is law that have been debated for

---

supra note 91, at 829–30.

106. Charles M. Yablon, *Justifying the Judge's Hunch: An Essay on Discretion*, 41 HASTINGS L.J. 231, 233 (1990). For scholarship discussing Legal Realism in relation to a variety of theories, see generally LAURA KALMAN, LEGAL REALISM AT YALE: 1927–1960 (1986); WILFRID E. RUMBLE, JR., AMERICAN LEGAL REALISM: SKEPTICISM, REFORM, AND THE JUDICIAL PROCESS (1968); Andrew Altman, *Legal Realism, Critical Legal Studies, and Dworkin*, 15 PHIL. & PUB. AFF. 205 (1986); John Henry Schlegel, *American Legal Realism and Empirical Social Science: From the Yale Experience*, 28 BUFF. L. REV. 459 (1979); Zipporah Batshaw Wiseman, *The Limits of Vision: Karl Llewellyn and the Merchant Rules*, 100 HARV. L. REV. 465 (1987); John Henry Schlegel, *The Ten Thousand Dollar Question*, 41 STAN. L. REV. 435 (1989) (reviewing LAURA KALMAN, LEGAL REALISM AT YALE: 1927–1960 (1986)).

107. Yablon, *supra* note 106, at 235–36. For research regarding the link between pragmatism and Realists, see generally Morris R. Cohen, *The Basis of Contract*, 46 HARV. L. REV. 553 (1933); Morris R. Cohen, *Property and Sovereignty*, 13 CORNELL L. REV. 8 (1927); John Dewey, *The Historic Background of Corporate Legal Personality*, 35 YALE L.J. 655 (1926); John Dewey, *Logical Method and Law*, 10 CORNELL L. REV. 17 (1924). I do not wish to enter the debate of Realism but provide a brief, generalized view of Realism to contrast that of the Proceduralist form of decision making.

108. *See* Yablon, *supra* note 106, at 244.

109. *See id.* at 231–33.

110. *See id.* at 244.

111. Intuitively, appellate judges should be more comfortable with determinations that suggest that they have considered other inferences and have rejected those inferences based on a well-grounded rationale. *See generally* RICHARD A. POSNER, OVERCOMING LAW 109–44 (1995) (considering "What Do Judges Maximize?"); Richard S. Higgins & Paul H. Rubin, *Judicial Discretion*, 9 J. LEGAL STUD. 129 (1980).

centuries. More relevant to the role of appellate review is the function of federal courts as institutional actors.

D. Role of Appellate Review

Appellate courts have two primary purposes, and the standard of review to be applied by appellate courts should relate to those two purposes. Appellate courts should serve to develop the law in a particular area as guidance for future cases and to rectify egregious errors in particular cases.[112]

It is too restrictive to complain that the increasing volume of cases on appeal should dictate the level of review by an appellate court.[113] One possible solution to the volume problem would be to increase the capacity of the appellate courts to "handle" all appeals. This result is not feasible or desirable. Instead, appellate courts should focus their energy on getting the decision "right" in the context of those cases in which the error below rises to an egregious level.[114] A deferential standard of review—clearly

112. *See* PAUL D. CARRINGTON, DANIEL J. MEADOR & MAURICE ROSENBERG, JUSTICE ON APPEAL 2–3 (1976); *see also* ROBERT J. MARTINEAU, MODERN APPELLATE PRACTICE: FEDERAL AND STATE CIVIL APPEALS § 1.7, at 19 (1983) ("Although some commentators identify other functions, these other purposes are usually aspects of either error correction or law development . . . ."). I recognize, of course, the difference between a jury determination and findings of fact as articulated by a trial judge. Judges can make or be coerced to make detailed findings that may be reviewed by an appellate court which cannot be duplicated in a jury trial. *See* Cooper, *supra* note 55, at 650.

113. *See* Cooper, *supra* note 55, at 650 ("It would be remarkable if the actual working standard of review were not affected by the shifting functional ability of the courts of appeals to devote attention to the wisdom of specific findings in particular cases."). As the Ninth Circuit noted:

> It can hardly be disputed that application of a non-deferential standard of review requires a greater investment of appellate resources th[an] does application of the clearly erroneous standard. Appellate courts could do their work more quickly if they applied the clearly erroneous standard in most circumstances, because the courts then need only determine if the lower court's decision is a reasonable one, not substitute their own judgment for that of the trial judge.

United States v. McConney, 728 F.2d 1195, 1201 n.7 (9th Cir. 1984), *overruled by* Estate of Merchant v. Comm'r, 947 F.2d 1390 (9th Cir. 1991).

A related question, which is not susceptible to measurement, is the level of judicial review afforded to district judges based on past history or reputation. District judges, and appellate judges who review their decisions, including findings of fact, do not act without influence from previous cases. In this sense, appellate review can dictate a specified level of review, but a history of an appellate court with the findings of a particular district judge may influence—subconsciously or otherwise—the level of inquiry notwithstanding the stated level of discretion. *See* FLEMING JAMES, JR. & GEOFFREY C. HAZARD, JR., CIVIL PROCEDURE § 12.8, at 668 (3d ed. 1985) ("[A]n appellate court's inclination to accept a trial judge's findings depends . . . on the court's unstated degree of confidence in the trial judge's fairmindedness."); Cooper, *supra* note 55, at 655–56; Martin B. Louis, *Allocating Adjudicative Decision Making Authority Between the Trial and Appellate Levels: A Unified View of the Scope of Review, the Judge/Jury Question, and Procedural Discretion*, 64 N.C. L. REV. 993, 1015–16, 1016 n.160 (1986) (finding that, off the record, "[f]ormer appellate law clerks . . . regularly attest that both important cases and the decisions of certain notorious trial judges are scrutinized more carefully than others.").

114. *See* Anderson v. City of Bessemer City, 470 U.S. 564, 575 (1985) ("Documents or objective evidence may contradict the witness' story; or the story itself may be so internally inconsistent or implausible on its face that a reasonable factfinder would not credit it. Where such factors are present,

erroneous—accomplishes this result. The standard of review, therefore, describes more than the level of deference: it defines the level of responsibility of appellate courts.[115]

The effect of this division of labor, which uses responsibility as a guidepost, accomplishes both purposes. First, appellate courts can focus their capacities on developing law as opposed to focusing on factually intensive, case-specific questions with little value beyond the case at issue.[116] Second, a clearly erroneous standard of review allows appellate courts to monitor trial courts for major errors even if a particular case does not relate to the first objective of developing the law in a particular area.[117] Deference is appropriate in cases based primarily on "'multifarious, fleeting, special, narrow facts that utterly resist generalization'"[118] and where the "investment of appellate energy will . . . fail to produce the normal law-clarifying benefits that come from an appellate decision on a question of law."[119] Simply then, trial courts sort through evidence, written or oral, credibility-based or not, and make determinations of fact, and appellate courts take those facts as determined by a trial court and develop the law in a particular area.[120] Judge Posner ascribes to this view:

> [T]he main reason for appellate deference to the findings of fact made by the trial court is not the appellate court's lack of access to the materials for decision but that its main responsibility is to maintain the uniformity and coherence of the law, a responsibility

---

the court of appeals may well find clear error even in a finding purportedly based on a credibility determination."); United States v. U.S. Gypsum Co., 333 U.S. 364, 396 (1948) ("Where such testimony is in conflict with contemporaneous documents we can give it little weight, particularly when the crucial issues involve mixed questions of law and fact. Despite the opportunity of the trial court to appraise the credibility of the witnesses, we cannot under the circumstances of this case rule otherwise than that [this finding] is clearly erroneous.").

115. *See* Cooper, *supra* note 55, at 651 ("[I]t may be wise to serve the interests of all litigants by adopting standards of review that help sift out all but the more extreme claims of error."). This division of responsibility may also rest on the belief that trial court judges and appellate court judges develop skills essential to each function and, over time, increase competency in the area with which they become most familiar. *See* Frisch, *supra* note 96, at 77.

116. Cooper, *supra* note 55, at 652; *see also* Thompson v. Keohane, 516 U.S. 99, 114–15 (1995) (stating that "the trial court makes an individual-specific decision, one unlikely to have precedential value").

117. Some have argued that a clearly erroneous standard of review may discourage a party that did not prevail at the trial level from taking an appeal with little chance of success. *See Anderson*, 470 U.S. at 575 ("[T]he parties to a case on appeal have already been forced to concentrate their energies and resources on persuading the trial judge that their account of the facts is the correct one; requiring them to persuade three more judges at the appellate level is requiring too much.").

118. Pierce v. Underwood, 487 U.S. 552, 561–62 (1988) (quoting Maurice Rosenberg, *Judicial Discretion of the Trial Court, Viewed from Above*, 22 SYRACUSE L. REV. 635, 662 (1971)); *see also* Magee & Goldman, supra note 51, at 483–84.

119. *Pierce*, 487 U.S. at 561; *see also* Magee & Goldman, *supra* note 51, at 483–84.

120. *See* Mucha v. King, 792 F.2d 602, 605–06 (7th Cir. 1986).

> not engaged in if the only question is the legal significance of a particular and nonrecurring set of historical events.[121]

The appellate function is not to determine whether the findings of fact as determined by the trial court are correct, but instead it is to determine whether the findings of fact as determined by the trial court are clearly wrong.[122] Under this rationale, nearly correct factual determinations are close enough that "making a finer determination" of factual issues on appeal "is either not possible or not worth the time and effort," considering the appellate courts' dual role of error correction and development of the law.[123] In making factual determinations and drawing inferences from those determinations, the clearly erroneous standard suggests that, given the theoretical nature of fact finding, no amount of additional consideration at an appellate level is likely to produce, in theoretical terms, a more correct result.[124]

Professor Maurice Rosenberg best describes the phenomenon as the difference between primary discretion and secondary discretion.[125] Although referring to legal decision making and how discretion affects those choices, the theory is equally applicable to the standard of review artificially created by legal rules and common law jurisprudence.

Primary discretion is grounded in the ability to act independently when rules do not exist to guide the resolution.[126] More interesting for these purposes is secondary discretion, which concerns the relationship among judges in the hierarchical judicial system.[127] These relationships define how much an appellate court is willing to restrain its own views in light of the responsibility of the trial court.[128] In other words, a trial court can be wrong to a certain extent without an appellate court finding "enough" er-

---

121. *Id.*; s*ee also* Lee, *supra* note 97, at 240 ("The nub of Judge Posner's rationale is that if an appellate court is not going to create useful precedent with its decision [or to promote uniformity], then no rationale justifies discarding the district court's work by non-deferential review."); Wright, *supra* note 96, at 779 ("From the earliest times appellate courts have been empowered to reverse for errors of law, to announce the rules which are to be applied, and to ensure uniformity in the rules applied by various inferior tribunals.").

122. Cooper, *supra* note 55, at 657; *see also* United States v. Aluminum Co. of Am., 148 F.2d 416, 433 (2d Cir. 1945) (Hand, J.) (Reluctance to reverse the findings of a trial judge "is true to a considerable degree even when the judge has not seen the witnesses. His duty is to sift the evidence, to put it into logical sequence and to make the proper inferences from it; and in the case of a record of over 40,000 pages like that before us, it is physically impossible for an appellate court to function at all without ascribing some prima facie validity to his conclusions.").

123. Yablon, *supra* note 106, at 269–70; s*ee also* Lee, *supra* note 97, at 236.

124. *See* Yablon, *supra* note 106, at 269–70.

125. Maurice Rosenberg, *Judicial Discretion of the Trial Court, Viewed From Above*, 22 SYRACUSE L. REV. 635, 637 (1971).

126. *Id.*

127. *Id.*

128. *See id.*

ror to reverse its determination. Secondary discretion defines how wrong a trial court must be for an appellate court to substitute its views.[129]

If secondary discretion is to apply in a standard of review context, de novo review is not an appropriate standard. De novo review is appropriate for rules (i.e., law determinations). If the responsibility of an appellate court is to develop the law, then review of action by a trial court that influences the development is appropriate.[130] Appellate courts ensure that lower courts follow the law, promote efficiency and predictability, and ultimately provide a legitimacy that is grounded in the rule of law.[131] Legal determinations are justified by a trial court by demonstrating the application of the factual issues to the appropriate controlling rule.

Appellate courts, by the nature of their responsibility within the judicial hierarchy and its attending institutional structure, are deemed to know more about the law and the appropriate rule that controls the legal issues.[132] Thus, de novo review in this context fits within the system of judicial institutional goals. Considering the entire judicial system as integrated, appellate resources should be devoted to matters that are most important to the proper functioning of the judicial system—promoting uniformity and predictability in the law.[133]

Appellate courts, through their law-developing function, act both affirmatively and negatively in directing the trial court on the legal decision-making process. Of course, affirmative action consists of announcing a legal position or precedent to be followed by a trial court. Just as importantly, however, appellate courts restrict the choices of trial courts by ruling out particular options or raising concerns regarding particular avenues of decision making, suggesting to trial courts that particular options will be rejected on review.[134] The law-development function promotes out-

129. *See id.*

130. CARRINGTON, MEADOR & ROSENBERG, *supra* note 112, at 2 ("[T]he review for correctness serves to reinforce the dignity, authority, and acceptability of the trial, and to control the adverse effects of any personal shortcomings of the basic decisionmakers.").

131. *See id.* at 147 (stating that uniformity is "one of the imperatives of appellate justice"); Gerald Gunther, *Congressional Power to Curtail Federal Court Jurisdiction: An Opinionated Guide to the Ongoing Debate*, 36 STAN. L. REV. 895, 911 (1984) ("Although the uniformity-assuring function of the Court does not strike me as a constitutionally mandated one, as a matter of policy, our system—any system—would be poorer and less coherent in the absence of a single, ultimately authoritative court at the apex of the judicial hierarchy."). For research regarding the role of predictability in judicial decisions, see Frisch, *supra* note 96, at 78 n.92 (quoting David Lyons, *Formal Justice and Judicial Precedent*, 38 VAND. L. REV. 495, 496 (1985) ("The reason most often given for the practice of precedent is that it increases the predictability of judicial decisions."); Earl Maltz, *The Nature of Precedent*, 66 N.C. L. REV. 367, 368 (1988) ("The most commonly heard justification for the doctrine of stare decisis rests on the need for certainty in the law.")).

132. *See* Yablon, *supra* note 106, at 273.

133. *See* Frisch, *supra* note 96, at 79 ("[V]aluable appellate resources are conserved for those issues that appellate courts are best situated to decide." (citing United States v. McConney, 728 F.2d 1195, 1201 (9th Cir. 1984), *overruled by* Estate of Merchant v. Comm'r, 947 F.2d 1390 (9th Cir. 1991))).

134. *See* Frisch, *supra* note 96, at 85; *see also* MARTINEAU, *supra* note 112, § 1.9, at 20.

comes that are based within the range of acceptable legal results and advances predictability as a welcome consequence.

Such a view comports with the notion that language is not a sufficient basis for communicating a precise rule.[135] Because language—and therefore rules of law—cannot be defined with particularized clarity, appellate courts must constantly refine their guidance to trial courts on the basis of the law and how that law is applied to a specific set of facts as determined by a trial court.[136] Appellate courts, through their responsibility to develop the law, should more easily reverse those decisions by the trial court that have an impact on the law's development. This is not simply done by the creation of more elaborate rules, but by the courts' willingness to reverse those determinations in which that action is warranted.

Trial judges find facts in the context of all facts considered in arriving at a factual determination. It is because of superior institutional competence that an appellate court does not reverse a trial court even if it would have reached a contrary conclusion on the facts presented or based on inferences that can be drawn from those facts. Upholding the factual determinations of the trial judge does not demonstrate that the trial judge was correct in those determinations, but it does reflect the institutional competence of the trial court. As long as the findings do not rise to the necessary level of wrong (i.e., clearly erroneous), the factual findings are not subject to reversal.[137]

Others have argued that parties who know that a reversal on appeal is unlikely will devote more energy to the initial trial.[138] While seemingly plausible, there is a significant difference in "making a record" for appeal and devoting the energies necessary to be successful at the trial stage. It is unlikely that a party would rely on an appellate court to remedy a perceived injustice at the trial stage when the party itself did not advance an effort to successfully convince a trier of fact of its basis for prevailing at trial.[139] Although the likelihood of success on appeal under a heightened standard of review is reduced, a de novo standard should not alter the approach to the initial trial. In other words, it is implausible that a litigant

135. *See* Pietruszkiewicz, *supra* note 90, at 1362 ("Language is not a sufficient basis for defining a deference standard."); *see also* Clermont, *supra* note 57, at 1148.

136. *See* Cooper, *supra* note 55, at 662.

137. *See* Yablon, *supra* note 106, at 267–68.

138. *See* Cooper, *supra* note 55, at 652; *see also* Anderson v. City of Bessemer City, 470 U.S. 564, 575 (1985) (determining that de novo review of factual determinations would unduly burden the party who "ha[s] already been forced to concentrate [its] energies and resources on persuading the trial judge that [its] account of the facts is the correct one").

139. *See* Wright, *supra* note 96, at 780 ("It is hard to believe that there has been any great public dissatisfaction with the restricted appellate review which was traditional in this country."). Professor Wright quoted Chief Justice Ellsworth: "'[S]urely, it cannot be deemed a denial of justice, that a man shall not be permitted to try his case two or three times over.'" *Id.* at 780–81 (quoting Wiscart v. D'Auchy, 3 U.S. (3 Dall.) 321, 329 (1796)).

would simply try a case for appeal in circumstances where additional evidence cannot be introduced for the first time at the appellate level.[140]

Perhaps more believable, however, is the argument that the findings of fact of a panel of three appellate judges are more accurate than the findings of one judge at the district court level.[141] As the Supreme Court suggests, "Duplication of the trial judge's efforts in the court of appeals would very likely contribute only negligibly to the accuracy of fact determination at a huge cost in diversion of judicial resources."[142] More significantly, appellate review of factual findings made by a trial court do not present the appellate courts with an opportunity to meaningfully develop the law and are case-specific determinations lacking any appreciable impact on predictability and uniformity.[143] In short, the division of responsibility should be the determining factor of the fact–law distinction.

III. APPLYING STANDARDS OF REVIEW IN THE CONTEXT OF ECONOMIC SUBSTANCE

Courts can make principled distinctions as to whether the economic substance test should be conjunctive or disjunctive. The question here, however, is the standard of review to be applied to economic substance as articulated by an appellate court charged with reviewing a determination by a trial court. The division of responsibilities argument provides an adequate basis for creating a clearly erroneous standard of review under both the subjective and the objective economic substance tests.

A. The Rule 52(a) Phenomenon

The objective and subjective economic substance tests differ in one major respect. Objective economic substance focuses on the reasonable businessman model, determining if a potential for profit in a particular industry exists by applying a reasonable businessman standard.[144] On the other hand, under the subjective economic substance test the question is whether the taxpayer believed that he was acting for nontax business reasons, notwithstanding whether a reasonable businessman standard would

140. A heightened standard, however, reduces the chances that the losing party at trial will be ultimately successful on appeal and, theoretically, would also reduce the number of appeals.

141. *See* Cooper, *supra* note 55, at 653 ("[I]t is easier to criticize a symphony than to write one, and much easier for one person to write a symphony than for a panel of three."); *see also* Judges of the Federal Courts, 901 F.2d vii–xxx (1990) (excluding the Federal Circuit, eighty-one judges of 206 federal appellate judges are former district judges).

142. *Anderson*, 470 U.S. at 574–75.

143. *See* Frisch, *supra* note 96, at 79; David P. Leonard, *The Correctness Function of Appellate Decision-Making: Judicial Obligation in an Era of Fragmentation*, 17 LOY. L.A. L. REV. 299, 301 (1984) (describing trial courts as being "at the front lines of fact-finding" and "therefore, do not exist for the purpose of making law").

144. *See* Cherin v. Comm'r, 89 T.C. 986, 994 (T.C. 1987).

otherwise be met.[145] In considering the standard of review in economic substance cases however, there is insufficient reason to treat these tests differently.

The traditional analysis of the fact–law distinction is outdated—judges hear witnesses and, to the extent that a determination is based on testimony, appellate courts should defer to the experience of the trier of fact in light of its unique ability to weigh credibility issues. In fact, Rule 52(a) expressly considers the ability of a trial court to judge the demeanor of a witness—"the reviewing court must give due regard to the trial court's opportunity to judge the witnesses' credibility."[146] The "due regard" language, combined with the precursor to this clause—"[f]indings of fact . . . must not be set aside unless clearly erroneous"[147]—serves as a demanding hurdle to reverse a finding of fact on appeal based on the testimony of a witness. Thus, unless a witness is patently unbelievable, it is inappropriate for an appellate court to reverse a finding of fact when the testimony of two or more witnesses is plausible. To resolve such questions in this manner, however, is likely too simplistic.

This is particularly so in light of the language *intentionally* omitted above from Rule 52(a). Rule 52(a)(6) provides in full that:

> Findings of fact, *whether based on oral or other evidence*, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial court's opportunity to judge the witnesses' credibility.[148]

The italicized language, added in 1985, makes clear that it is not simply the demeanor and credibility of witnesses that justifies the level of deference accorded a trier of fact.[149] Instead, this language suggests that it is the specific responsibilities of each court in the judicial hierarchy that justify the result. In this sense, a Proceduralist interpretation of discretion would apply based principally on the adoption of institutional, as opposed to ideological, factors that contribute to the division of responsibility.[150]

---

145. *See* Blum, *supra* note 39, at 524 n.106.
146. FED. R. CIV. P. 52(a)(6).
147. *Id.*
148. *Id.* (emphasis added).
149. Upon adoption of the 1985 amendments to Rule 52, the advisory committee noted: "To permit courts of appeals to share more actively in the fact-finding function would tend to undermine the legitimacy of the district courts in the eyes of litigants, multiply appeals by encouraging appellate retrial of some factual issues, and needlessly reallocate judicial authority." FED. R. CIV. P. 52 advisory committee's notes (1985).
150. United States v. McConney, 728 F.2d 1195, 1201 (9th Cir. 1984), *overruled by* Estate of Merchant v. Comm'r, 947 F.2d 1390 (9th Cir. 1991) (noting that the application of Rule 52(a)'s clearly erroneous standard "emphasizes . . . the trial judge's opportunity to judge the accuracy of witnesses' recollections and make credibility determinations").

Cases, and particularly tax cases, often do not implicate conflicting testimony—or testimony at all—but are instead submitted on a paper record.[151] In such a case, an appellate court has the same information as the trial court, but Rule 52(a) dictates that findings based on a written record—documentary evidence—shall nonetheless remain undisturbed unless clearly erroneous.[152]

While Rule 52(a) suggests that the clearly erroneous standard applies to all factual evidence,[153] the Supreme Court has recognized that, compared to a paper record, greater deference should be given to findings of fact based on the credibility of witnesses.[154] In addition, the Supreme Court's inquiry is markedly more pointed if a conflict exists between documentary evidence and oral testimony.[155] This distinction does not modify the clearly erroneous standard for documentary evidence under Rule 52(a),

151. For example, Tax Court Rule 122(a) provides that:

> Any case not requiring a trial for the submission of evidence (as, for example, where sufficient facts have been admitted, stipulated, established by deposition, or included in the record in some other way) may be submitted at any time after joinder of issue . . . by motion of the parties filed with the Court.

TAX CT. R. 122(a).

152. Prior to the 1985 amendment of Rule 52, courts had widely varying views as to whether documentary evidence was entitled to the same deferential standard of review as evidence gathered based on the credibility of witnesses. First, "[s]ome courts of appeal have stated that when a trial court's findings do not rest on demeanor evidence and evaluation of a witness' credibility, there is no reason to defer to the trial court's findings and the appellate court more readily can find them to be clearly erroneous." FED. R. CIV. P. 52 advisory committee's notes (1985) (citing Marcum v. United States, 621 F.2d 142, 144–45 (5th Cir. 1980)). Another group of courts of appeals concluded "that appellate review may be had without application of the 'clearly erroneous' test since the appellate court is in as good a position as the trial court to review a purely documentary record." *Id.* (citing Atari, Inc. v. N. Am. Philips Consumer Elecs. Corp., 672 F.2d 607, 614 (7th Cir. 1982); Lydle v. United States, 635 F.2d 763, 765 n.1 (6th Cir. 1981); Swanson v. Baker Indus., Inc., 615 F.2d 479, 483 (8th Cir. 1980); Taylor v. Lombard, 606 F.2d 371, 372 (2d Cir. 1979); Jack Kahn Music Co. v. Baldwin Piano & Organ Co., 604 F.2d 755, 758 (2d Cir. 1979); John R. Thompson Co. v. United States, 477 F.2d 164, 167 (7th Cir. 1973)). Finally, "[a] third group has adopted the view that the 'clearly erroneous' rule applies in all nonjury cases even when findings are based solely on documentary evidence or on inferences from undisputed facts." *Id.* (citing Maxwell v. Sumner, 673 F.2d 1031, 1036 (9th Cir. 1982); United States v. Texas Educ. Agency, 647 F.2d 504, 506–07 (5th Cir. 1981); Constructora Maza, Inc. v. Banco de Ponce, 616 F.2d 573, 576 (1st Cir. 1980); *In re* Sierra Trading Corp., 482 F.2d 333, 336–37 (10th Cir. 1973); Case v. Morrisette, 475 F.2d 1300, 1306–07 (D.C. Cir. 1973)).

153. *See* FED. R. CIV. P. 52(a)(6); *see also* McFarland v. T.E. Mercer Trucking Co., 781 F.2d 1146, 1148 (5th Cir. 1986) (finding of a trial judge based on depositions is entitled to the same deference as a finding "based on oral, in-court testimony").

154. *See* Anderson v. City of Bessemer City, 470 U.S. 564, 575 (1985) ("When findings are based on determinations regarding the credibility of witnesses, Rule 52(a) demands even greater deference to the trial court's findings; for only the trial judge can be aware of the variations in demeanor and tone of voice that bear so heavily on the listener's understanding of and belief in what is said.").

155. *See* United States v. U.S. Gypsum Co., 333 U.S. 364, 396 (1948) ("Where such testimony is in conflict with contemporaneous documents we can give it little weight, particularly when the crucial issues involve mixed questions of law and fact. Despite the opportunity of the trial court to appraise the credibility of the witnesses, we cannot . . . rule otherwise than that [the finding] is clearly erroneous.").

nor does it suggest that a de novo standard of review should apply in this context.[156]

As Professor Wright suggests, Rule 52 does not say that "[f]indings of fact shall not be set aside unless clearly erroneous *if* the trial court has had an opportunity to judge of the credibility of the witnesses."[157] It does say that "[f]indings of fact . . . shall not be set aside unless clearly erroneous, *and* due regard shall be given to the opportunity of the trial court to judge of the credibility of the witnesses."[158]

The rationale of Rule 52(a) is inexorably intertwined with the division of responsibility among judicial actors. There is no principled distinction between a trial judge and an appellate panel of judges in deciding factual questions based on documentary evidence, except for the appreciation of the role of a trial court and the role of an appellate court in the efficient adjudication of cases.

Moreover, there is no policy reason why—particularly in document-intensive cases—an appellate court should give the findings of a trial judge greater weight than the judge's legal conclusions. If compared to a jury system, it would seem that appellate judges, which are at least three in number, would better reflect a representative sample of the community than a single trial judge.

The clearly erroneous standard cannot then be entirely a reflection of whether a trial court or an appellate court is in a better position, but instead is inescapably linked to the function and responsibility of each court within the judicial system. This desirability of a division of responsibility is not grounded, as in the case of a jury, in the Constitution, but in judicial administration. There is nothing that requires how findings of fact by a court must be treated by an appellate court. The application of the clearly erroneous rule is grounded in its creation by the Supreme Court.

Congress empowered the Supreme Court to establish rules of procedure, and it has chosen, in the interest of a division of responsibility, to create a clearly erroneous standard in Rule 52(a) by which to separate

156. *See Anderson*, 470 U.S. at 575. While greater deference applies to testimony based on the credibility of witnesses, the clearly erroneous standard of review nonetheless applies to documentary evidence.

157. Wright, *supra* note 96, at 769–70. Furthermore, Professor Wright states: "That Rule 52 required application of the 'clearly erroneous' test to all findings, regardless of the nature of the evidence, should thus have been apparent to anyone who understands the difference between a hypothetical and a conjunctive proposition." *Id*. at 770.

158. FED. R. CIV. P. 52(a)(6) (emphasis added). Rule 52 was amended in 2007 to provide that "[f]indings of fact, whether based on oral or other evidence, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial court's opportunity to judge the witnesses' credibility." *Id*. According to the comments accompanying the changes, "Rule 52 has been amended as part of the general restyling of the Civil Rules to make them more easily understood and to make style and terminology consistent throughout the rules. These changes are intended to be stylistic only." FED. R. CIV. P. 52 advisory committee's notes (2007).

lines of authority within the courts.[159] The subjective and objective economic substance tests, just like Rule 52(a), should follow that division of judicial responsibility, requiring appellate courts to defer to trial courts on matters that do not advance the law in a particular area.

It is tempting, of course, for appellate courts to consider individualized objectives in ensuring that an appropriate result occurs in a particular case. In light of the institutional factors guiding the responsibilities of the courts, however, the function of appellate courts is to provide guidance for future cases and promote predictable resolution. Individualized justice does not promote this result.

B. Institutional Considerations

Under both the subjective and objective economic substance tests, decisions of trial courts, while likely to be significant in terms of dollars at stake, are unlikely to have any appreciable impact on the development of the law in the economic substance area. The subjective economic substance test incorporates the concept of business purpose and, as a result, relies on the particularized exigencies of a specific business. Necessarily, a business purpose for one business may or may not be a business purpose for another business, even if both businesses are similar in nature. Because this is a subjective inquiry, a trial court decision considering whether one business satisfies the subjective business purpose doctrine is irrelevant to the analysis of the business purpose of another. The result in one case, therefore, fails to advance the law in another.[160]

Under the objective standard, there are no particularized business exigencies similar to those that arise under the subjective economic substance standard. The objective economic substance test itself relies on external factors unrelated to the inherent uniqueness of a particular business.[161] However, under the objective business purpose doctrine, a determination of whether there is a meaningful increase in the net economic position of the taxpayer other than the reduction of taxes fails to advance the law with respect to other cases. Just as under the subjective economic substance test, the result of this mathematical calculation has little, if any, bearing on results that occur in other cases. The clearly erroneous standard of review

159. *See* 28 U.S.C. § 2072(a) (2000) ("The Supreme Court shall have the power to prescribe general rules of practice and procedure and rules of evidence for cases in the United States district courts (including proceedings before magistrate judges thereof) and courts of appeals."); *see also* Stern, *supra* note 70, at 84.

160. *See, e.g.*, Comm'r v. Duberstein, 363 U.S. 278, 290–91 (1960) (question of whether a transfer constitutes a gift is a question for a trial court and appellate review is quite limited).

161. *But see* Petition for Writ of Certiorari, *Dow*, *supra* note 4, at *27–*28; Magee & Goldman, *supra* note 51, at 482 n.9 ("Even under the objective prong of the two-prong test, the taxpayer's subjective business plans logically must be considered simply to establish the terms of the transaction whose objective economic substance is to be assessed.").

is not a rubber stamp simply because a decision in an economic substance case does not advance the law in that area. Appellate courts may still rectify egregious errors by a trial court in specific cases and adhere to the principles of institutional efficiency.

C. Tax Avoidance vs. Tax Minimization

Code sections often intertwine objective rules and subjective elements. These subjective elements are primarily reflected in the use of terms such as "principal purpose" for a transaction, the requirement of the absence of a tax avoidance motive as part of the transaction, or specific references to the "significant purpose" or "business purpose" of the transaction. These anti-avoidance principles appear to rest exclusively on the state of mind of the taxpayer, suggesting that if tax avoidance is a motive, the tax advantages of a transaction will be—or should be—disallowed.

Unquestionably, however, taxpayers can arrange their affairs in order to minimize taxes. The famous, oft-quoted endorsement of Judge Learned Hand that "[a]ny one may so arrange his affairs that his taxes shall be as low as possible; he is not bound to choose that pattern which will best pay the Treasury; there is not even a patriotic duty to increase one's taxes"[162] is universally known in the tax community. Thirteen years later, Judge Hand sounded the same chord:

> Over and over again courts have said that there is nothing sinister in so arranging one's affairs as to keep taxes as low as possible. Everybody does so, rich or poor; and all do right, for nobody owes any public duty to pay more than the law demands: taxes are enforced exactions, not voluntary contributions. To demand more in the name of morals is mere cant.[163]

In 2008, the Supreme Court reiterated this concept: "We have also recognized that '[t]he legal right of a taxpayer to decrease the amount of what otherwise would be his taxes, or altogether avoid them, by means which the law permits, cannot be doubted.'"[164] Because of the lack of involvement of the sovereign in the transaction, there is, of course, a temptation to believe that tax avoidance is so pervasive that in looking to the state of mind of the taxpayer (i.e., the business purpose of the transaction), courts require overwhelming evidence to suggest that business deci-

162. Helvering v. Gregory, 69 F.2d 809, 810 (2d Cir. 1934), *aff'd*, 293 U.S. 465 (1935).
163. Comm'r v. Newman, 159 F.2d 848, 850–51 (2d Cir. 1947) (Hand, J., dissenting).
164. Boulware v. United States, 128 S. Ct. 1168, 1176 n.7 (2008) (quoting Gregory v. Helvering, 293 U.S. 465, 469 (1935)).

sions were based on aberrations or eccentricities of the taxpayer in issue.[165] Courts must guard against such a requirement.

If saving taxes is an acceptable form of human behavior, it could hardly be prohibited. If tax minimization is acceptable but tax avoidance is not tolerable, there must be a manner in which distinctions can be drawn.[166] Tax avoidance and tax minimization both involve subjective thresholds. The problem, however, is that, if both involve a state of mind, it is objectively impossible to distinguish a desire to reduce taxes from a desire to avoid taxes.[167]

It is not adequate to define tax avoidance as a heightened level of tax minimization. Instead, the difference must balance the interests being considered. Under this rationale, tax avoidance must contemplate circumstances in which the nontax goals of a transaction are insufficient in weight when compared to the goal of tax minimization.[168]

In a case in which the issue is whether a taxpayer is engaged in a tax avoidance transaction instead of a tax minimization arrangement, a taxpayer must demonstrate the existence and the significance of a nontax objective—the business purpose of a transaction.[169] If able to demonstrate a nontax goal that is comparatively more compelling than the tax objective, then tax reduction as opposed to tax avoidance provides a sufficient basis to provide a taxpayer with the tax benefits of the transaction. If the taxpayer is unable to elevate the nontax goal over the tax objective, then tax avoidance objectives should disallow the tax benefit of the transaction.

It is no easy task to describe—much less define or prove—state of mind. As a result, the subjective "state of mind" inquiry focuses on whether the nontax goals were plausibly compelling enough to justify the tax advantages that resulted from the nontax goals.[170]

By its very nature, the subjective economic substance test must consider all of the facts of the transaction that generated the tax benefit to determine whether the transaction also maintained an appreciable business purpose. As a result, specifying a test applicable to all taxpayers based on business purpose is inherently difficult, if not impossible. Subjectivity

165. *See* Blum, *supra* note 39, at 498.

166. *See* ACM P'ship v. Comm'r, 157 F.3d 231, 248 n.31 (3d Cir. 1998) ("*Gregory* and its progeny 'do not allow the Commissioner to disregard economic transactions . . . which result in actual, non-tax-related changes in economic position' regardless of 'tax-avoidance motive.'" (quoting N. Ind. Pub. Serv. Co. v. Comm'r, 115 F.3d 506, 512 (7th Cir. 1997))).

167. *See* Blum, *supra* note 39, at 515.

168. *See id.* at 516; *cf.* Gunn, *supra* note 16, at 738 ("If the tax success of a transaction depends on a business purpose, and if 'business purpose' means what it seems to mean—a purpose other than that of reducing taxes—the principle that tax-avoidance purpose does not count against the taxpayer has been seriously undermined.").

169. *But see* Gunn, *supra* note 16, at 740 (arguing that "a strong tax motive is fatal" despite the assurance by the Supreme Court in *Gregory v. Helvering* that a strong tax motive is not fatal).

170. Blum, *supra* note 39, at 523.

necessitates an individualized review of the particular circumstances of the taxpayer engaged in the transaction.

The notion that appellate courts are superior to trial courts and should guide the degree of deference to trial courts is antiquated.[171] While this notion may have been true long ago, such a notion does not fit with the reality of modern complexities presented at the trial level.[172] Economic substance mandates an inquiry of all relevant facts and circumstances of a transaction. Thus, trials involving economic substance tend to last for weeks rather than hours.[173] For example, in *Dow Chemical Co. v. United States*, the Internal Revenue Service asserted a $22.21 million assessment challenging Dow's corporate-owned whole life insurance (COLI) plans as lacking economic substance.[174]

After a two-month bench trial, 26 witnesses, and over 1,500 exhibits, the district court issued a 139-page opinion and a 15-page post-judgment decision[175] rejecting the economic substance argument advanced by the

171. *See* Frisch, *supra* note 96, at 72–73. In the 1920s, the number of cases heard by appellate courts permitted a more searching inquiry and appellate courts could devote sufficient time to making a decision. *See* Arthur D. Hellman, *Central Staff in Appellate Courts: The Experience of the Ninth Circuit*, 68 CAL. L. REV. 937, 938 (1980). More time could be devoted to briefs, transcripts, oral argument, and ultimately, decision writing. *See* FELIX FRANKFURTER & JAMES M. LANDIS, THE BUSINESS OF THE SUPREME COURT, at vii–viii (1927); John Bilyeu Oakley & Robert S. Thompson, *Law Clerks in Judges' Eyes: Tradition and Innovation in the Use of Legal Staff by American Judges*, 67 CAL. L. REV. 1286 (1979); *see also* WARREN E. BURGER, YEAR-END REPORT ON THE JUDICIARY 2 (1981) (quoting Solicitor General, Robert Bork, as stating that appellate courts are moving "towards an assembly line model"); CARRINGTON, MEADOR & ROSENBERG, *supra* note 112, at 7 ("Changes which are characteristic of any shift from individually crafted works to mass production methods can be seen to be occurring in appellate processes and institutions."); Robert H. Bork, Solicitor General of the U.S., Dealing with the Overload in Article III Courts, Addresses Delivered at the National Conference on the Causes of Popular Dissatisfaction with the Administration of Justice (Apr. 7–9, 1976), *in* 70 F.R.D. 231, 233 (1976) (referring to appellate courts as "processing institutions"); Winslow Christian, *Appellate Bloat Threatens Our Courts*, JUDGES' J., Spring 1980, at 27, 27 ("A law teacher, a legislator, or even a newly appointed appellate judge may hold quaint ideas about what goes on inside an appellate court: The judges are supposed to give thoughtful attention to full-scale oral argument; they supposedly consider the cases thoroughly in conference; and then a judge (assisted by an admiring young law clerk in the role of apprentice) studies the record, collects the authorities, and goes through several drafts before presenting his colleagues with an opinion embodying the collegial conclusion of the judges. It is a charming and reassuring picture. But that picture is contrary to fact in every appellate court that I know about.").

172. *See* Frisch, *supra* note 96, at 73.

173. *See* Bose Corp. v. Consumers Union of the U.S., Inc., 466 U.S. 485, 500 (1984) (a level of deference "tends to increase when trial judges have lived with the controversy for weeks or months instead of just a few hours"); Container Corp. of Am. v. Franchise Tax Bd., 463 U.S. 159, 176 (1983) ("The factual records in [unity business] cases . . . tend to be long and complex . . . . It will do the cause of legal certainty little good if this Court turns every colorable claim that a state court erred in a particular application of those principles into a *de novo* adjudication, whose unintended nuances would then spawn further litigation and an avalanche of critical comment. Rather, our task must be to determine whether the . . . court applied the correct standards to the case; and if it did, whether its judgment was within the realm of permissible judgment." (internal quotation marks omitted) (citing ASARCO, Inc. v. Idaho State Tax Comm'n, 458 U.S. 307, 326 n.22–23 (1982))).

174. *See* 435 F.3d 594, 596–98 (6th Cir. 2006).

175. *See* Petition for Writ of Certiorari, *Dow*, *supra* note 4, at *1–2.

United States and allowing the deductions claimed by Dow.[176] Similarly, *Coltec Industries, Inc. v. United States*[177] involved a 10-day trial and 29 witnesses.[178] Professor Cooper argues that this deference should "reflect . . . the familiarity of prolonged exposure [that] can enhance the empathetic and intuitive aspects of decision."[179]

Because the test measures the business purpose of the transaction, a number of considerations may be relevant. Necessarily, a business purpose for the transaction cannot be invented by the taxpayer after challenge by the Internal Revenue Service, but instead must relate to the planning of the transaction before it was executed. That is, even if an objective business purpose justification for the transaction exists, the subjective aspect of the test must require that its contemplation preceded undertaking the transaction.

Business purpose necessarily involves subjectivity because no two applications can be the same; no taxpayer can be in precisely the same position as another taxpayer. Accordingly, the business purpose of one entity should not be imputed to another merely because a transaction was deemed to fail or satisfy the business purpose test for the former entity. In fact, by tailoring a set of facts to a particular business, the business-purpose component of economic substance can alter what would appear to be the correct result in a particular case.[180]

Business purpose can "override both the formal mechanics and purported effect of the transaction as structured by the parties,"[181] but must be justified based on the circumstances of each particular case. As the Tax Court recognized in one 1959 case, "We are convinced . . . from our study of *all the facts and circumstances* that none of the alleged [corporate] advantages . . . constituted any actual business purpose *in the instant case*."[182]

---

176. *See id.* The district court determined that "Dow has established by a preponderance of the evidence that . . . its . . . COLI plans . . . had substantial effects on the beneficial interest of the taxpayer apart from the income tax deductions." Dow Chem. Co. v. United States, 250 F. Supp. 2d 748, 811 (E.D. Mich. 2003), *modified by* 278 F. Supp. 2d 844 (E.D. Mich. 2003), *rev'd*, 435 F.3d 594 (6th Cir. 2006).

177. 62 Fed. Cl. 716 (Fed. Cl. 2004), *vacated*, 454 F.3d 1340 (Fed. Cir. 2006).

178. *See id.* at 718, 730 (ten-day trial involving twenty-nine witnesses and focusing, in the government's words, on "a complex series of prearranged steps").

179. Cooper, *supra* note 55, at 654. *See, e.g.*, Bose Corp. v. Consumers Union of U.S., Inc., 466 U.S. 485, 500–01 (1984).

180. *See* Gergen, *supra* note 42, at 140 ("What is abusive is to do what was done in *ACM*—to enter a partnership that has little or no economic substance with the goal from the start of creating an artificial loss. The message is that you can pick up tax gold if you find it in the street while going about your business, but you cannot go hunting for it.").

181. Isenbergh, *supra* note 2, at 873.

182. Aldon Homes, Inc. v. Comm'r, 33 T.C. 582, 597–98 (T.C. 1959) (emphasis added); *see also* Thompson v. Comm'r, 631 F.2d 642, 646 (9th Cir. 1980) (stating that economic substance "requires the Tax Court to focus on the facts and circumstances of particular transactions and resolve whether, as a practical matter, those transactions have any economic impact outside the creation of tax deduc-

The Supreme Court in *Ornelas v. United States*[183] provides some context. While addressing probable cause to search a vehicle, the Court noted that

> *de novo* review tends to unify precedent and will come closer to providing law enforcement officers with a defined "set of rules which, in most instances, makes it possible to reach a correct determination beforehand as to whether an invasion of privacy is justified in the interest of law enforcement."[184]

The Court continued, "It is true that because the mosaic which is analyzed for a reasonable-suspicion or probable-cause inquiry is multi-faceted, 'one determination will seldom be a useful "precedent" for another.'"[185]

Although this opinion is founded in criminal law principles, the "reasonable suspicion" or "probable cause" analysis is equally applicable to economic substance because of the intentional vagueness associated with each test. A definitive, precise, and objective definition of business purpose is implausible.[186] The business purpose doctrine makes sense of the practical position that "[e]ven the smartest drafters of legislation and regulation cannot be expected to anticipate every device."[187] Business purpose does not describe a baseball box score and similarly cannot describe a double play as "6–4–3" or "Tinker to Evers to Chance."[188] This is precisely because business purpose is not an objective measurement and cannot be listed in a box score. Moreover, it is impossible to foresee—much less write—rules that produce an intended result in a complex financial

---

tions").

183. 517 U.S. 690 (1996).

184. *Id.* at 697–98 (quoting New York v. Belton, 453 U.S. 454, 458 (1981)); *see also* Thompson v. Keohane, 516 U.S. 99, 100 (1995) (stating that "the law declaration aspect of independent review potentially may guide police, unify precedent, and stabilize the law," which may "serve legitimate law enforcement interests").

185. *Ornelas*, 517 U.S. at 698 (quoting Illinois v. Gates, 462 U.S. 213, 238 n.11 (1983)).

186. *See* Blum, *supra* note 39, at 543.

187. ASA Investerings P'ship v. Comm'r, 201 F.3d 505, 513 (D.C. Cir. 2000).

188. *Baseball's Sad Lexicon*, also referred to as *Tinker to Evers to Chance*, refers to a baseball poem by Franklin Pierce Adams written in 1910 referring to a double play "turned" by Chicago Cubs shortstop, Joe Tinker, second baseman Johnny Evers, and first baseman Frank Chance against the New York Giants. In full, the poem reads—

> These are the saddest of possible words:
> "Tinker to Evers to Chance."
> Trio of bear cubs, and fleeter than birds,
> Tinker and Evers and Chance.
> Ruthlessly pricking our gonfalon bubble,
> Making a Giant hit into a double —
> Words that are heavy with nothing but trouble:
> "Tinker to Evers to Chance."

FRANKLIN P. ADAMS, *Baseball's Sad Lexicon*, *in* IN OTHER WORDS 62, 62 (1912).

world in which taxpayers utilize sophisticated tax lawyers to construct new transactions and permutations of existing transactions.[189]

In such a case, review should be deferential simply because of the improbability that there will exist two identical cases. In other words, this is a classic example of a type of case that calls for limited appellate inquiry.[190] The responsibility of the appellate court for maintaining or promoting uniformity is not triggered, and therefore, de novo review is unwarranted.[191] The Seventh Circuit has stated that "[c]onsiderations which favor a de novo standard, such as the desire to create a uniform rule, are not present [under the subjective construct] since no single rule could embrace the varied fact patterns which may arise . . . ."[192]

As a result, the determination of economic substance as it relates to the subjective test is a factual determination "focus[ed] on the facts of circumstances of [a] particular transaction" to determine whether it has economic impact beyond the tax deductions created by the transaction.[193] However, the mere creation of a complex transaction that provides a tax savings or tax advantage that otherwise has economic substance should not undermine that economic substance.[194] Under this formulation, judges have the discretion to view the entire transaction to determine business objectives and distinguish between genuine business transactions and artificial tax shelters, permitting tax planning but disallowing tax avoidance.

CONCLUSION

"One man's tax shelter being everyone else's budget deficit, all taxpayers ultimately have an interest in the reasonable interpretation of the tax laws."[195] The government[196] has taken the position in some cases that the standard of review is a de novo standard while, in others, it has claimed that a clearly erroneous standard of review applies.[197] The varia-

189. *See* Hariton, *When and How*, *supra* note 3, at 33.
190. *See* Magee & Goldman, *supra* note 51, at 484.
191. *See* Mucha v. King, 792 F.2d 602, 605–06 (7th Cir. 1986).
192. Rexnord, Inc. v. United States, 940 F.2d 1094, 1097 (7th Cir. 1991).
193. *See* Thompson v. Comm'r, 631 F.2d 642, 646 (9th Cir. 1980).
194. *See* Gergen, *supra* note 42, at 144.
195. Isenbergh, *supra* note 2, at 863 n.15.
196. The use of the term "government" is intentionally a generic one. In Tax Court, the named party is the Commissioner of Internal Revenue while the United States is the named party in the Court of Federal Claims and in federal district court. In cases originating in Tax Court, cases are litigated by the Internal Revenue Service, Office of the Chief Counsel. In the Court of Federal Claims and federal district courts, the client is the Internal Revenue Service but cases are litigated by the Department of Justice, Tax Division. All circuit court cases, whether originating in the Tax Court, Court of Federal Claims, of federal district court, are litigated by the Tax Division of the Department of Justice. In cases before the United States Supreme Court, the Department of Justice's Office of the Solicitor General represents the government.
197. *Compare* Am. Elec. Power Co. v. United States, 326 F.3d 737, 741 (6th Cir. 2003) (asserting a clearly erroneous standard of review after decision of the trial court in favor of the government),

tion in position concerning the standard of review leads to a logical conclusion that the government is trying to achieve favorable results on a case-by-case basis.[198] This is not an appropriate approach by the government.

The courts, on the other hand, have been inconsistent in their application of a standard of review in economic substance cases, creating a differing standard based on the circuit that has jurisdiction over an almost certain appeal. Instead, tax policy and the rule of law should dictate consistency of position notwithstanding the result reached by a court below, and courts should adopt a framework that recognizes economic substance cases as inherently involving non-recurring facts that institutionally suggest the adoption of a clearly erroneous standard of review.

The economic substance doctrine provides an appropriate balance between the objective results contemplated by the Internal Revenue Code and the subjective constraints necessary to achieve its purposes.[199] Economic substance cases are decided based on the particularized facts and circumstances associated with each case. Because of this type of inquiry, appellate decisions offer little guidance for the development of the law. They are case-specific determinations lacking an impact on predictability and uniformity in subsequent considerations of economic substance.

In the context of economic substance cases, appellate review should be guided by the division of responsibilities in the judicial system. Appellate courts have two primary objectives: to develop the law in a particular area as guidance for future cases and to rectify egregious errors in a particular case. The distribution of discretion along the judicial hierarchy should relate to those two purposes. In considering subjective and objective economic substance, these purposes are best accomplished through the application of a clearly erroneous standard of review.

Rexnord, Inc., 940 F.2d at 1096 (same), Cashman v. United States, No. 90-15806, 1991 WL 67902, at *4 (9th Cir. May 2, 1991) (same), *and* Newman v. Comm'r, 902 F.2d 159, 162 (2d Cir. 1990) (same), *with* Brief for Appellant at 33, Dow Chem. Co. v. United States, 435 F.3d 594 (6th Cir. 2006) (No. 03-2360), 2004 WL 5329408 (arguing that "[t]he conclusion that Dow's COLI plans were not shams is reviewed *de novo*" after decision of the trial court in favor of taxpayer (citing *Am. Elec. Power*, 326 F.3d at 742)).

198. *See* Richard M. Lipton, *What Will Be the Long-Term Impact of the Sixth Circuit's Divided Decision in* Dow Chemical*?*, 104 J. TAX'N 332, 337 (2006) ("The expansive interpretation of *Knetsch* . . . is likely to continue the impression that the government is more interested in favorable results than the rule of law, which can only lead to long-term adverse consequences.").

While unrelated to the carousel of positions by the Internal Revenue Service on the standard of review issue, Professor Ginsburg's point seems appropriate: "[E]very stick crafted to beat on the head of a taxpayer will metamorphose sooner or later into a large green snake and bite the commissioner on the hind part." Martin D. Ginsburg, *Making Tax Law through the Judicial Process*, A.B.A. J., Mar. 1984, at 74, 76.

199. *See* Hariton, *Sorting Out*, supra note 3, at 241.