Denver Law Review

Volume 72 | Issue 1 Article 9

January 2021

Vol. 72, no. 1: Full Issue

Denver University Law Review

DENVER UNIVERSITY LAW REVIEW

VOLUME 72

1994-1995

Published by the
University of Denver
College of Law

DENVER UNIVERSITY LAW REVIEW

1994 Volume 72 Issue 1

TABLE OF CONTENTS

ARTICLES

A Pragmatic Strategy for the Scope of Sales Law, the Statute of Frauds, and the Global Currency Bazaar — *Raj Bhala* — 1

Hitting 'Em Where It Hurts: Using Title IX Litigation to Bring Gender Equity to Athletics — *Melody Harris* — 57

Yardstick Damages in Lost Profit Cases: An Econometric Approach — *Roger D. Blair & Amanda Kay Esquibel* — 113

Freedom and Culture: Why We Should Not Buy Commercial Speech — *R. George Wright* — 137

COMMENTS

J.E.B. v. Alabama ex rel. T.B.: Discrimination by any Other Name . . . — *Pamela R. Garfield* — 169

Central Bank of Denver v. First Interstate Bank: Pruning the Judicial Oak by Severing the Aiding and Abetting Branch — *Paul Dmitri Zier* — 191

ERRATUM

Several years ago I published *Holding Women's Psyches Hostage: An Interpretive Analogy on the Thomas/Hill Hearings*, 69 DENV. U. L. REV. 171 (1992). This article began with a short narrative describing the experiences and death of a sexually harassed woman named Betty. The article notes that Betty's death occurred in conjunction with and was partially motivated by the Thomas/Hill hearings. In writing the narrative, I relied on several knowledgeable sources and Betty's unemployment benefits case file. Nevertheless, a year later I received information indicating Betty's death occurred before, rather than concurrent with, the Thomas/Hill hearings. I write to acknowledge my mistake and to express regret for the error.

Penelope E. Bryan

A PRAGMATIC STRATEGY FOR THE SCOPE OF SALES LAW, THE STATUTE OF FRAUDS, AND THE GLOBAL CURRENCY BAZAAR

RAJ BHALA*

TABLE OF CONTENTS

I. Sales Law and the Currency Bazaar 2
II. The Hypothetical Spot Transaction 6
A. The Basic Terms 7
B. Negotiating the Deal 8
C. The Dispute 10
III. Scope—Should Foreign Exchange Be a "Good"? 11
A. The Carelessly Inclusive Approach 11
1. The Five Careless Courts 12
2. Statutory Ambiguity 14
B. The Aggressively Exclusive Approach 18
1. The Stalemate of Pre-U.C.C. Cases 19
2. A Formalistic Distinction 20
3. The Drafters' Intention 21
4. Underlying Principles 21
C. The Pragmatic Strategy 23
IV. Enforceability—Does the Statute of Frauds Serve Market Needs? 25
A. The Clash of Cultures 27
B. Case One: Tape Recordings as "Writings" 29
1. The Tangibility Paradigm 29
2. Possible Legislative Amendments 29
3. Judicial Re-interpretation 31
4. Certainty 34
5. Fraud Prevention or Fraud Promotion? 35
6. Results-Oriented Jurisprudence 37
C. Case Two: The Costs and Benefits of Confirmations 38
1. The Tangibility Paradigm Again 38
2. Delays and Costs 41
3. Practical Irrelevance 43

* Assistant Professor of Law, Marshall-Wythe School of Law, College of William and Mary. A.B., 1984, Duke; M.Sc., 1985, London School of Economics; M.Sc., 1986, Oxford University; J.D., 1989, Harvard Law School. Former Attorney, Federal Reserve Bank of New York, 1989-93. I would like to express my gratitude to the following individuals for the comments they provided on earlier drafts of this article: Professors David Frisch, Widener University School of Law, Patricia B. Fry, University of North Dakota School of Law, and Fred Miller, University of Oklahoma Law Center. I am also indebted to my colleague at William and Mary, Professor Glenn Coven, for his suggestions on an earlier draft. Finally, I am grateful for the research assistance of Ms. Nina Hval, third year student, William and Mary.

4. Helping the Sophisticated 43
5. The Tension with the Parol Evidence Rule 44
6. The Tension with the Battle of the Forms Rule 48
D. Linking the Resolutions of the Scope and Enforceability Problems 50
V. Extending the Pragmatic Strategy to Other Sales Law 51
A. Current Proposals to Revise U.C.C. Article 2 51
B. The United Nations Convention on Contracts for the International Sale of Goods 52
C. Private Contract Law 53
VI. Conclusion 55

Contract law has progressed and evolved sounder principles since the days of ritualistic and formalistic sealed instrument requirements.[1]

I. SALES LAW AND THE CURRENCY BAZAAR

The cheerful evaluation of contract law by the Court of Appeals in *V'Soske v. Barwick* is stunningly naive. At least one important branch of contract law, sales law, may be ill-equipped to serve the needs of the modern international financial marketplace. In the global currency bazaar, banks, corporations, investors, and governments buy and sell foreign currency from one another. Their transactions are conducted on tenuous legal grounds.[2] Fundamental issues of the scope of sales law and the enforceability of contracts are unresolved.

Perhaps if the currency bazaar were economically insignificant it would not matter whether sales law served the needs of foreign exchange traders. This bazaar, however, is the world's largest financial market. At the end of an average day, roughly $1 trillion worth of currencies has changed hands.[3] This figure represents a 35 percent increase in turnover in just three years (1989-92).[4] Hence, explosive growth, as well as enormous size, is a salient feature of the bazaar. Moreover, the currency bazaar never closes.[5] At any time of the day or night millions of U.S. dollars are

1. V'Soske v. Barwick, 404 F.2d 495, 499 (2d Cir. 1968), *cert. denied*, 394 U.S. 921 (1969).

2. For an overview of foreign exchange transactions, see Raj Bhala, *Risk Trade-offs in the Foreign Exchange Spot, Forward and Derivative Markets*, 1 THE FINANCIER 34 (1994); ROGER M. KUBARYCH, FOREIGN EXCHANGE MARKETS IN THE UNITED STATES (rev'd. ed. 1983).

3. MONETARY & ECONOMIC DEP'T, BANK FOR INT'L SETTLEMENTS (BASLE, SWITZERLAND), CENTRAL BANK SURVEY OF FOREIGN EXCHANGE MARKET ACTIVITY IN APRIL 1992 1, 5-6 (1993) [hereinafter, CENTRAL BANK SURVEY]. In April 1992, the total reported gross turnover in spot, forward, and derivative foreign exchange contracts was $1.354 trillion. (After correcting for double counting and estimated reporting gaps, the figure was $880 billion.) *Id.* at 6 tbl. I; *see also* James Blitz, *All Change in Foreign Exchanges: The Nature of the International Currency Dealing Has Altered*, FIN. TIMES (London), Apr. 2, 1993, at 17.

4. CENTRAL BANK SURVEY, *supra* note 3, at 6 tbl. I.

5. *See* Scott E. Pardee, *Internationalization of Financial Markets*, 72 ECON. REV. (Federal Reserve Bank of Kansas City, Mo.), Feb. 1987, at 3, 3 (noting that "[i]n the 1970s, foreign exchange became a 24-hour market, with the major banks dealing with each other each day through their offices in the markets in the Far East (Tokyo, Hong Kong, Singapore) the Middle East (once Beirut, now Bahrain), then Europe (London, Frankfurt, Zurich), and finally, the United States (New York for interbank trading and Chicago for futures trad-

traded for Japanese yen, millions of German marks are exchanged for French francs, millions of English pounds are sold against Swiss francs, and so forth.[6] The buy and sell decisions of banks, businesses, and institutional and individual investors establish exchange rates. In turn, the exchange rates critically influence the international flow of goods, services, and money. The currency bazaar is too big, growing too rapidly, and too important to the international economy to ignore.

This important bazaar, however, has been ignored by legal scholars.[7] Because foreign exchange trading raises fundamental contract issues, this neglect must end. A legal assessment of the needs of foreign exchange market participants is required, not only by the participants themselves, but also by judges and regulators who are trying to cope with a small but increasing number of breakdowns in foreign exchange trading.

Of course the participants expect to find currency trading profitable; otherwise they would channel their resources into a different financial market. Yet, participants can no longer dismiss the breakdowns as isolated incidents in their race for profits. Formal adjudication results from mishaps that involve high monetary stakes; but when a bank, business, or investor goes to court or to a regulatory agency over a foreign exchange deal gone sour, they get little comfort. Judges and regulators are groping. They do not know whether to apply sales law to the transactions, and if so, which sales law should be used. Judges and regulators do not know whether the transactions are enforceable. Worst of all, they do not have a conceptual framework for dealing with these issues.

This article attempts to fill the large and dangerous void in legal scholarship on foreign exchange transactions. It further seeks to provide a conceptual framework for judges and regulators that will help them resolve problems in the currency bazaar. It addresses two fundamental contractual problems that arise in the currency bazaar under Article 2 of the

---

ing.)"). Technically, the trading week begins in New Zealand on Monday morning at approximately 8 a.m. and ends in New York on Friday at approximately 5 p.m. While trading volumes are thin at other times, it remains possible to find counterparties with which to enter into foreign exchange transactions.

6. The five most widely traded currencies are the dollar, mark, yen, pound, and Swiss franc. The percentage of turnover accounted for by these currencies are 82, 40, 23, 14, and 9, respectively. CENTRAL BANK SURVEY, *supra* note 3, at 9 tbl. II(a).

7. With the exception of a fifteen year-old student Note, law reviews are largely devoid of articles on commercial law problems associated with foreign exchange transactions. *See* Michael L. Manire, Note, *Foreign Exchange Sales and the Law of Contracts: A Case For Analogy to the Uniform Commercial Code*, 35 VAND. L. REV. 1173 (1982). This Note asserts that Article 2 should apply to foreign exchange transactions, yet provides no theoretical basis for the assertion. *See id.* at 1174, 1188, 1192, 1200, 1206, and 1209. The Note fails to account for the highly significant fact that conversations between parties negotiating and concluding foreign exchange transactions typically are tape-recorded. It incorrectly suggests that all foreign exchange transactions are confirmed in writing. *See id.* at 1186-87. As page proofs of this article were being prepared, a brief article on foreign exchange transactions and Article 2 appeared. *See* Stephen C. Veltri, *Should Foreign Exchange Be "Foreign" to Article Two of the Uniform Commercial Code?*, 27 CORNELL INT'L L.J. 343 (1994). This welcome addition to the literature provides an excellent review of relevant pre-U.C.C. cases. It does not advocate a particular approach, like the pragmatic strategy argued for herein, to the issue of the scope of Article 2, nor does it dicuss the statute of frauds.

Uniform Commercial Code ("U.C.C."). First, should the scope of sales law cover foreign exchange? Second, should oral foreign exchange contracts be enforceable despite the statute of frauds?

This article proposes that these questions be answered using a *pragmatic strategy* which emphasizes the *needs of foreign exchange market participants* in relation to Article 2. The pragmatic strategy identifies an integral relationship between the problems of scope and enforceability in order to ensure that sales law does not impinge on the vitality and dynamism of the currency bazaar. Thus, according to this proposed strategy, the resolution of the first problem hinges critically on the outcome of the second problem.

Courts and regulators have failed to cope adequately with the problem of scope. Courts have adopted a "carelessly inclusive" approach. Regulators, most notably the Federal Reserve,[8] have used an "aggressively exclusive" approach. Instead, an examination of the key rules of Article 2 in the context of the currency bazaar is needed before the problem can be resolved. The statute of frauds is such a rule. How it functions in the currency bazaar is a critical indicator of whether Article 2 should govern transactions in this bazaar.

Currently, courts and regulators misguidedly adhere to what is best termed the "tangibility paradigm"—the fixation on a tangible document to satisfy the statute of frauds. The paradigm should be abandoned because it does not serve the needs of the currency bazaar. It is a strict approach to the problem of enforceability that clashes with the telephonic technology of the bazaar. It is also incongruous with the repeat-player, high-trust culture of the bazaar. Legislative amendments to, or a judicial re-interpretation of, the statute of frauds is needed. Unless these changes are made, the pragmatic strategy suggests foreign exchange transactions may be appropriately excluded from the scope of Article 2. In that event, other commercial laws, considered below, may be applicable.[9]

By no means is the thesis advanced herein limited to the foreign exchange market. The underlying and general theoretical question is at what point is it appropriate to codify a market or industry? The foreign exchange market is a case study of this question. Recently, Professor Raymond T. Nimmer set forth three criteria to determine when a market is ripe for codification: the area of the contract (i) achieves national scope (the nationality criterion); (ii) affects substantial commercial volume (the volume criterion), and (iii) would benefit from codification instead of common law governance (the relevance criterion).[10] Because the pragmatic strategy suggests that market needs must be examined, the strategy appears consistent with the relevance principle. Moreover, the foreign exchange market clearly satisfies the nationality and volume criteria. The

8. Hereinafter, unless otherwise noted, the "Federal Reserve" refers to the Federal Reserve Bank of New York and the Board of Governors of the Federal Reserve System.

9. *See infra* notes 218-235 and accompanying text.

10. Raymond T. Nimmer, *Intangibles Contracts: Thoughts of Hubs, Spokes, and Reinvigorating Article 2*, 35 WM. & MARY L. REV. 1337, 1367-73 (1994).

problem then becomes one of the relative weight to be given to the criteria. If two of Professor Nimmer's criteria are satisfied, is the conclusion in favor of codification inexorable? Not necessarily. The thesis herein suggests that the relevance principle may be relatively more important than other criteria. We cannot rush to include or exclude foreign exchange from Article 2 until we have determined whether the statute "works" for the market.

There are five remaining parts to this article. Part II sets up a hypothetical but highly realistic foreign exchange transaction. The hypothetical transaction raises the issues of scope and enforceability, and it is used in parts III-V to advance the thesis of this article.

Part III considers the scope issue, namely, whether foreign exchange should be a "good" within the ambit of Article 2. In part III, both the carelessly inclusive and aggressively exclusive approaches are rejected; but the pragmatic strategy, which centers on the relationship between contract enforceability rules and the needs of the currency bazaar, is advanced.

Part IV applies the pragmatic strategy to the statute of frauds issue. This part highlights the clash between a formal legal approach to the statute of frauds on the one hand, and the way transactions are negotiated and concluded in the currency bazaar on the other hand. The argument rejects the tangibility paradigm and advocates changing the statute of frauds (either legislatively or judicially) to meet the needs of the market.

Even though parts III and IV focus on the application of Article 2 to the hypothetical foreign exchange transaction, no shortage exists of other sales laws for judges to apply directly or by analogy. Accordingly, Part V extends the pragmatic strategy to three other potentially applicable sales law regimes:[11] first, the proposed revisions to Article 2 ("revised Article 2"), which are presently under consideration by the National Conference of Commissioners on Uniform State Laws and the American Law Institute;[12] second, the United Nations Convention on Contracts for the International Sale of Goods ("CISG"), to which approximately 34 countries are Contracting States, including the U.S., China, France, Germany, Switzerland, Italy, Canada, and Australia;[13] and third, private contract law, specifi-

11. The common law of contract is, of course, potentially applicable to disputes arising from foreign exchange transactions. However, there is no question of the scope of this legal regime; it purports to cover all contracts unless displaced by other law. There is not a common law statute of frauds for the sale of goods.

12. U.C.C. REVISED ARTICLE 2, pts. 1-3, 7 (Tentative Draft Dec. 21. 1993) and pts. 1-6 (Tentative Draft Sept. 10, 1993) [hereinafter, December 1993 Draft and September 1993 Draft, respectively] (on file with author) and Uniform Commercial Code Revised Article 2. Sales (Draft July 29-Aug. 5, 1994) [hereinafter, August 1994 Draft] (on file with author).

13. The United Nations Conference on Contracts for the International Sale of Goods, Apr. 10, 1980, 19 I.L.M. 668 (Final act, entered into force Jan. 1, 1988) [hereinafter CISG]. Documents and summary records of the Conference appear in U.N. CONFERENCE ON CONTRACTS FOR THE INTERNATIONAL SALE OF GOODS, U.N. Doc. A/CONF.97/19, U.N. Sales No. E.82.V.5 (1981). The text of the Conference and reservations taken by Contracting States are reprinted in Selected International Conventions, 1994 MARTINDALE-HUBBLE INTERNATIONAL LAW DIGEST, pt. VII, at IC-29 [hereinafter MARTINDALE-HUBBLE]. The legislative history of the Senate advice and consent action is found at *Proposed United Nations Convention on Contracts for the International Sale of Goods: Hearings on Treaty Doc. 98-9 Before the Comm. on Foreign Relations,*

cally, the International Foreign Exchange Master Agreement ("IFEMA") which is a standard-form contract devised recently by banks that actively trade foreign exchange.[14] Part V argues that the IFEMA is unhelpful in resolving the statute of frauds problem because that contract lacks the essential quantity term. Part V further argues that it may be appropriate to extend the scope of the CISG and revised Article 2 to cover foreign exchange transactions because these regimes omit a statute of frauds provision.

Part VI synthesizes the arguments of Parts III-V and draws conclusions from the application of the pragmatic strategy to the scope and enforceability problems in the currency bazaar.

II. THE HYPOTHETICAL SPOT TRANSACTION

> *The foreign-exchange market is one of the world's slickest. It is screen-based, genuinely international and open for business 24 hours a day. There are many buyers and sellers; prices adjust rapidly and for the most part smoothly. And it is huge.*[15]

United States Senate, 98th Cong., 2d Sess. (1984). A list of Contracting States is maintained by the Treaty Section, Office of Legal Affairs, United Nations.

Curiously, even though the U.S. is a Contracting State, no court or commentator has considered the applicability of the CISG to foreign exchange transactions directly or by analogy. *See, e.g.*, E. Allen Farnsworth, *The Vienna Convention: History and Scope*, 18 INT'L LAW. 17 (1984); Alejandro M. Garro, *Reconciliation of Legal Traditions in the U.N. Convention on Contracts for the International Sale of Goods*, 23 INT'L LAW. 443 (1989); James E. Joseph, *Contract Formation Under the United Nations Convention on Contracts for the International Sale of Goods and the Uniform Commercial Code*, 3 DICK. J. INT'L L. 107 (1984); Courtney P. Smart, Comment, *Formation of Contracts in Louisiana Under the United Nations Convention for the International Sale of Goods*, 53 LA. L. REV. 1339 (1993).

14. *See* FOREIGN EXCHANGE COMMITTEE, INTERNATIONAL FOREIGN EXCHANGE MASTER AGREEMENT § A (1993) [hereinafter IFEMA]. Even though the Foreign Exchange Committee first published the IFEMA in November 1993, it is already being used by foreign exchange market participants in New York, London, and Tokyo. It is not the only example of a privately-negotiated, standard-form written agreement in the global currency bazaar. In 1992, the International Swaps and Derivatives Association (ISDA) published a revised master agreement which can be tailored for use in spot foreign exchange transactions by altering the schedules to the agreement. FXNET, a London-based system for netting foreign exchange trades on a bilateral basis, has issued a Worldwide Agreement.

The IFEMA, however, appears to have the brightest prospects for widespread adoption by spot market participants. In part, the IFEMA bears the informal imprimatur of the Federal Reserve. The Foreign Exchange Committee (specifically, the Financial Market Lawyers Group) drafted the IFEMA. This Committee advises and operates under the auspices of the Federal Reserve. *See* FOREIGN EXCHANGE COMMITTEE, 1992 ANNUAL REPORT 5 (1993) [hereinafter FEC 1992]. In addition, the IFEMA reflects what seems to be commonly regarded as the best foreign exchange market practice. *See* FOREIGN EXCHANGE COMMITTEE, GUIDE TO THE 1993 INTERNATIONAL FOREIGN EXCHANGE MASTER AGREEMENT § C (1993) [hereinafter IFEMA GUIDE]; Ruth W. Ainslie, Foreign Exchange: The New Master Foreign Exchange Trading Agreements 1, (Apr. 20, 1994) (unpublished manuscript, presented at the International Monetary Fund Seminar on Current Legal Issues Affecting Central Banks, Washington, D.C., May 18, 1994, on file with the International Monetary Fund).

15. *The Last of the Good Times?*, ECONOMIST, Aug. 15, 1992, at 61.

A. *The Basic Terms*

The most basic and common foreign exchange transaction is a spot.[16] A spot contract involves a commitment by one party to deliver a specified quantity of one currency against another party's delivery of a specified quantity of a second currency. In effect, each party is buying one currency and paying for it with another currency. The date the commitment is made is the trade date. The value date, which is the date on which the reciprocal deliveries occur, is within two business days of the trade date.[17]

For example, suppose on November 1 Citibank and the Development Bank of Singapore ("DBS") enter into a dollar-yen spot foreign-exchange transaction.[18] Citibank agrees to buy 120 million yen from DBS in exchange for U.S. dollars at a price, or exchange rate, of 104 yen per dollar.[19] Hence, Citibank pays DBS $1,153,846.15 for the yen. Payment of the dollars, and the reciprocal yen payment, must occur on the value date, November 3.

---

16. *See* CENTRAL BANK SURVEY, *supra* note 3, at 16 ("The spot market is still the single most important segment of the foreign exchange market.").

17. *See id.* at 16. Thus, in a spot foreign exchange transaction, the settlement of payment obligations is said to occur on "T+2" (where "T" stands for the trade date). Settlement occurs, however, on T+1 in the spot markets for Mexican pesos and Canadian dollars. The analysis and arguments in this article are equally applicable to a forward foreign exchange transaction, i.e., one in which the value date is more than two days after the trade date.

18. Citibank is a commercial bank headquartered in New York. DBS is a commercial and investment bank headquartered in Singapore. The parties are deliberately put in different parts of the world not only to simulate real-world conditions but also to illustrate that trading in many different currencies occurs far away from the home countries of those currencies.

Non-bank parties—such as corporations, institutional investors, individual investors, and governments (mainly central banks)—actively participate in the foreign exchange market. Inter-bank dealing, however, accounts for 70 percent of total market activity. *See* CENTRAL BANK SURVEY, *supra* note 3, at 1, 11-12; Manire, *supra* note 7, at 1183. Of the transactions between interbank dealers, 41 percent are with dealers located abroad and 29 percent are with dealers in the same jurisdiction. CENTRAL BANK SURVEY, *supra* note 3, at 11-12.

19. In practice, exchange rates are quoted in more precise terms because finer movements are observed that result in large profits and losses. The dollar-yen rate is quoted in terms of hundredths of yen per dollar. For example, the rate on Friday, April 15, 1994 was 103.45 yen per dollar. *Currency Trading—Exchange Rates*," WALL ST. J., Apr. 18, 1994, at C15.

DBS need not actually possess 120 million yen at the time that it enters into a spot contract to sell yen. If it did not, it would be short selling—that is, selling foreign currency that is not held. For an example of selling 400 million pounds short against the dollar, see ANDREW J. KRIEGER, THE MONEY BAZAAR: INSIDE THE TRILLION-DOLLAR WORLD OF CURRENCY TRADING 65-79 (1992). If DBS sells yen short, it will have to obtain 120 million yen to cover its short position before it is contractually obligated to deliver the yen to Citibank's designated account on the appropriate value date.

DBS could cover its short position in a number of ways. First, it may take a long position in another spot transaction, for instance, buy yen in the spot market. DBS risks losing money if the spot rate for yen rises above 104 yen per dollar. Suppose the yen appreciates relative to the dollar to 103 yen per dollar. Then, DBS will have to buy 120 million yen at the market rate of 103 yen per dollar and sell them to Citibank at the previously agreed rate of 104 yen per dollar. The purchase will cost DBS $1,165,048.54, but the sale will fetch only $1,153,846.15, resulting in a loss of $11,202.39. Of course, DBS would not short sell yen if it thought yen would appreciate. It expects to profit by buying 120 million yen at, say, 107 yen per dollar and selling the yen at 105 yen per dollar. A second way for DBS to cover a short sale is to enter into a foreign currency option transaction. DBS could buy a call option on yen. This option would entitle DBS to obtain yen at a pre-set price.

B. *Negotiating the Deal*

The dollar-yen spot agreement is reached between foreign exchange traders at Citibank and DBS through direct telephone negotiations.[20] The telephone conversations are tape recorded by each of the parties.[21] Either Citibank or DBS may initiate the transaction. Assume a Citibank trader in New York calls a trader at DBS in Singapore. The Citibank trader asks for "a quote on dollar-yen," but does not necessarily indicate whether she intends to buy or sell yen.[22] The custom in the foreign exchange market is to quote "two-way rates," i.e., both the bid and offer prices.[23] The DBS trader does not learn the Citibank trader's intention until the Citibank trader proposes a specific trade.[24]

Assume the DBS trader tells her counterpart at Citibank that the bid-offer rates are "105-104."[25] The Citibank trader indicates a desire to buy 120 million yen. The DBS trader responds "120 million yen, yours, at 104." At that juncture, according to the custom in the foreign exchange market, the traders believe a spot contract is established for the sale of 120 million yen by DBS to Citibank at a price of 104 yen per dollar. Citibank is

---

20. *See* RUDI WEISWEILLER, INTRODUCTION TO FOREIGN EXCHANGE 14 (2d ed. 1984) (stating foreign exchange traders may transact by telephone); Thomas M. Campfield & John G. O'Brien, *Foreign Exchange Trading Practices: The Interbank Market, in* INTERNATIONAL FINANCE HANDBOOK § 2.4, at 3 (Abraham M. George & Ian H. Giddy, eds., 1983) (noting "[m]odern sophisticated communications" like telephones link market participants).

An alternative way for the Citibank and DBS traders to communicate directly is through an electronic messaging system. Such computer-to-computer communication requires the banks to be members of the same automated dealing system. The systems are sponsored by third-party vendors such as Reuters. About one-third of all foreign exchange deals are concluded through these systems. *See* CENTRAL BANK SURVEY, *supra* note 3, at 21, 24. The effect of automated dealing on the resolution of the scope and enforceability problems is beyond the scope of this article.

As an alternative to direct dealing, the Citibank and DBS traders might deal with each other indirectly through one or more foreign-exchange brokers who act as agents for their respective principals. *See* Bhala, *supra* note 2, at 100. Around one-third of all foreign exchange transactions between inter-bank dealers are arranged through brokers. *See* CENTRAL BANK SURVEY, *supra* note 3, at 21 tbl. VI, 23-24. All telephone conversations between the traders and their respective brokers, and between the brokers, would be taped. Consequently, the analysis presented herein is unaffected by the use of brokers.

21. Tape recording the telephone conversations between foreign exchange traders is customary practice in the currency bazaar. Telephone interview with Philip Hemnell, Kim Eng Securities, New York, N.Y. (June 14, 1994).

22. *See* KRIEGER, *supra* note 19, at 31.

23. Because Citibank is willing to act as either a buyer or seller in any foreign exchange transaction, it contributes importantly to market liquidity. It will take either side of the spot deal.

24. *See* KRIEGER, *supra* note 19, at 31.

25. These rates are expressed in terms of number of yen per dollar. The bid price of 105 yen per dollar reflects the exchange rate at which DBS is ready, willing, and able to buy yen. The offer (or asking) price of 104 yen per dollar reflects the rate at which DBS is ready, willing, and able to sell yen.

The difference of one yen per dollar between the bid and offer prices is the "spread." The bid price necessarily must be lower than the offer price. If it were not, then DBS would perpetually lose money. DBS would buy yen at a higher price and be forced to sell at a lower price. At the 105-104 rates, a 120 million yen transaction implies that DBS will pay $1,142,857.14 to buy the yen. The rates also imply that DBS will receive $1,153,846.15 from selling the yen. The spread assures that DBS can buy and re-sell the yen for a profit of $10,989.01.

obligated to deliver $1,153,846.15 to DBS in two days. Conversely, DBS is obligated to deliver 120 million yen to Citibank in two days. The word "yours" is the magic trade term indicating entry into a spot contract.[26]

After the traders conclude their conversation, each trader fills out a deal ticket.[27] The ticket states the name or initials of the trader, name of the counterparty, trade and value dates, type and amount of currency purchased, exchange rate, and payment (i.e., currency delivery) instructions. Citibank and DBS do not exchange deal tickets. Rather, the tickets are used by each bank for internal accounting and control purposes, for example, to monitor the total amounts of yen, dollars, marks, pounds, and other currencies bought and sold by the bank.

In addition, Citibank and DBS might use the deal tickets as a basis for preparing written confirmations of the transaction. In many, but not necessarily all, foreign exchange transactions, the parties exchange two confirmations.[28] First, the traders exchange confirmations of the terms of the transaction by telex. Second, the operations departments of the respective parties exchange confirmations, usually by mail.[29] This article analyzes

---

26. After the traders conclude the transaction, employees of the operations department of each bank may endeavor to confirm the terms of the transaction. Citibank's operations department may send a written confirmation of the transaction to the operations department at DBS, and vice versa. Each operations department may check the confirmation received with its own records of the transaction. These records will include the tape recording of the transaction, coupled with any written record. Naturally, the written confirmations should match the records of the deal.

How do Citibank and DBS ascertain the designated account of the other to which the currencies must be delivered on the value date? The currencies are delivered to designated accounts of each bank. DBS will inform Citibank of its bank account name and number at which the dollars are to be delivered, and DBS will inform Citibank of the account name and number to which the yen must be transferred. The exchange of account information, or settlement instructions, may occur as part of the written confirmation process. Alternatively, assuming Citibank and DBS routinely trade with one another, they may have previously exchanged settlement instructions. Finally, Citibank and DBS may have signed a master agreement such as the IFEMA and set forth settlement instructions therein.

27. The telephone negotiation lasts only several seconds or a few minutes. In order to profit, foreign exchange traders must conclude transactions rapidly for two reasons. First, foreign exchange rates are volatile even in the short-term. The price quoted on a currency can move dramatically in seconds. If there is a delay in negotiating a transaction, then a prospective buyer or seller of a currency may elect not to proceed because exchange rates have changed such that the originally quoted rate is now an off-market rate, or no longer a rate at which the transaction would be profitable.

Second, profits are generated by market participants like Citibank, DBS, and their brokers by entering into a large volume of transactions. With respect to highly liquid currencies like the U.S. or Canadian dollar, Swiss or French franc, Japanese yen, German mark, and English pound, the bid-offer spread is small. Consequently, a large volume of purchase and sale transactions are necessary. A smaller number of high-risk trades, such as betting correctly against perceived market trends, are also a source of profits. Naturally, the pressure to conclude transactions rapidly makes oral negotiation by telephone an ideal means in the decentralized, global currency bazaar.

28. Telephone interview with Philip Hemnell, Kim Eng Securities, New York, N.Y. (June 16, 1994).

29. This department is also known as the "settlements department" or "back office." In addition to exchanging confirmations, the operations departments exchange payment instructions, usually by electronic means, through the Society for Worldwide Interbank Financial Telecommunications (SWIFT).

the case where Citibank and DBS do not exchange these confirmations, as well as the case where they are exchanged.[30]

C. *The Dispute*

Shortly after the deal is concluded, the bid-ask spread moves from 105-104 to 107-106 yen per dollar. The Citibank trader observes that as a result of short-term volatility, the financial asset (yen) she is buying has depreciated relative to the currency (dollars) in which she will render payment. Accordingly, she backs out of the deal with DBS and buys 120 million yen from a third bank at the cheaper rate of 106 yen per dollar.[31] In the meantime, the DBS trader has rejected offers from other banks to buy the yen.

DBS learns of Citibank's unscrupulous action no later than the value date. At that point, DBS transfers 120 million yen to Citibank but does not receive the reciprocal payment of $1,153,846.15. The DBS trader insists on selling 120 million yen at the original exchange rate of 104 yen per dollar. After all, it relied on Citibank's oral representations and turned down other opportunities to sell the yen at this rate.[32] Citibank refuses to go through with the transaction.

Thus, DBS sues Citibank for breach of contract in the appropriate federal court. Citibank argues that U.C.C. Article 2 applies to the dispute and raises the affirmative defense that if there was a spot contract, then it is unenforceable under the statute of frauds set forth in section 2-201.[33] Because there is no signed writing to evidence the contract, Citibank argues, this statute is not satisfied.[34] DBS's response is that the scope of Article 2 of the U.C.C. does not encompass foreign exchange transactions, hence the statute of frauds is inapplicable.

The thesis of this article implies that Citibank's defense should be rejected and, while DBS's response may have some merit, it too should be rejected. DBS's response is an over-reaction. The pragmatic strategy suggests that whether foreign exchange transactions ought to fall within the scope of Article 2 depends on a comprehensive assessment of the provisions of Article 2 in relation to the needs of foreign exchange market par-

30. For the case where confirmations are not exchanged, see *infra* notes 107-156 and accompanying text. For the case where confirmations are exchanged, see *infra* notes 167-216 and accompanying text.

31. At the cheaper rate, 120 million yen cost $1,132,075.41, as compared with $1,153,846.15 at the original rate of 104 yen per dollar. If the yen appreciated relative to the dollar, then a dispute could arise because DBS sought to renege on the deal with Citibank and sell 120 million yen at the new, higher market rate to a third party.

32. Of course, it is not reliance that leads to the creation of a contract; there is a contract here. Citibank and DBS have certain expectations that contract law is designed to protect regardless of the reliance factor.

33. *See* U.C.C. § 2-201 (1990). Throughout this article, unless otherwise noted, the references to the U.C.C. are to the 1990 Official Text, approved and published by the American Law Institute and the National Conference of Commissioners on Uniform State Laws. Note that Citibank's defense raises the possibility of the judicial admissions exception to the statute of frauds. *See* U.C.C. § 2-201(3)(b). This issue is not addressed herein.

34. Citibank does not dispute that a contract was formed. *Cf.* U.C.C. § 2-204(1) (stating that a contract may be made in any manner, including conduct by the parties).

ticipants. Examining the statute of frauds in relation to these needs is an important part of this assessment. The results of the examination, however, cannot be dispositive because there is obviously more to Article 2 than section 2-201. Citibank's defense should be rejected, and the contract enforced, because the statute of frauds does not meet the needs of market participants. The statute should be modified or reinterpreted to conform with these needs.

III. SCOPE—SHOULD FOREIGN EXCHANGE BE A "GOOD"?

> *Without question the [Uniform Commercial] [C]ode was designed to bring the body of commercial law into the contemporary world of business. . . . Its principal purpose was to meet the contemporary needs of a fast moving commercial society.*[35]

The wisdom of the *General Electric Credit* decision is lost on courts and regulators struggling with problems in the currency bazaar. Courts and regulators understand that sales law is needed to establish precisely the rights and obligations of Citibank and DBS in the dollar-yen spot transaction. Yet, these decision-makers employ conventional, sequential reasoning to the dispute: first, decide the scope of U.C.C. Article 2; second, consider the enforceability of the contract under section 2-201. Thus, until the scope of sales law is clear, the risks of engaging in spot transactions cannot be allocated with precision.

In their haste to establish rights and duties, courts take a carelessly inclusive approach to the scope issue. Regulators, in contrast, take an aggressively exclusive approach. Both approaches are imprudent. The pragmatic strategy says the conventional, sequential reasoning is rigid. Courts and regulators should first determine if Article 2 works for the currency bazaar; in the words of *General Electric Credit*, does it meet the "contemporary needs of a fast moving" market?[36] Only then should courts and regulators consider whether the scope of Article 2 ought to include transactions in that bazaar.[37]

A. *The Carelessly Inclusive Approach*

Many courts fail to discuss adequately why U.C.C. Article 2 governs transactions in the currency bazaar. They assume it applies, but provide scant analysis of the language of Article 2 to support this assumption. Courts disregard the implications to the foreign exchange market of applying Article 2 to spot transactions.

---

35. General Electric Credit Corp. v. R.A. Heintz Construction Co., 302 F. Supp. 958, 967-8 (D. Or. 1969).

36. *Id.* at 968.

37. A short-cut solution to the problem of scope would be to rely on the first clause of U.C.C § 2-102 which states, "[u]nless the context otherwise requires." It could be argued that even if foreign exchange is not a "good," the context mandates the application of Article 2 to foreign exchange transactions.

1. The Five Careless Courts

No fewer than five decisions adopt this carelessly inclusive approach to the problem of applying U.C.C. Article 2 to foreign exchange transactions. Consider the opinion of a New York appellate court in *United Equities Co. v. First National City Bank.*[38] The court decided a claim arising from a dollar-yen forward foreign exchange transaction.[39] The parties did not dispute the applicability of Article 2. The court nonchalantly applied Article 2 without considering whether the transaction involved the sale of "goods."[40]

In *Saboundjian v. Bank Audi (USA),* a New York appellate court again neglected to explore carefully the reason Article 2 applies to foreign exchange transactions.[41] There, a bank failed to execute an oral foreign exchange order from its customer who unreasonably declined to mitigate his damages. The court asserted that "[t]he Uniform Commercial Code is applicable to foreign exchange transactions, since 'the Code excludes "money" only when it is a medium of payment, not when treated as a commodity.' "[42] The decision provided no theory for a distinction between money as the subject of a contract and money as the medium of payment.

In *Intershoe, Inc. v. Bankers Trust Co.,* the New York Court of Appeals considered whether a written confirmation slip of a foreign exchange transaction is the final expression of the parties' agreement for purposes of the parol evidence rule.[43] The court baldly stated: "[t]here seems to be no question that the UCC applies to foreign currency transactions."[44] In *Compania Sud-Americana de Vapores, S.A. v. IBJ Schroder Bank & Trust Co.,* the U.S. District Court for the Southern District of New York uncritically relied on the *Intershoe* decision and applied the Article 2 parol evidence rule to disputed foreign currency conversions.[45]

Finally, in *Koreag v. Refco F/X Associates, Inc.,* the U.S. Court of Appeals for the Second Circuit provided only a sketchy overview of the meaning of "goods" under section 2-105. This overview justified the court's application of Article 2 to foreign exchange transactions between a New York corporation, Refco, and a Swiss bank, Mebco. The court granted Refco, as the seller of U.S. dollars, a right to reclaim the dollars from Mebco, an

38. 383 N.Y.S.2d 6 (N.Y. App. Div. 1976), *aff'd,* 395 N.Y.S.2d 640 (N.Y. 1977).

39. United Equities entered into a contract on April 12, 1971 for the purchase of 360 million yen against $1,018,710. Under the six-month forward contract, the yen were to be delivered on Oct. 14, 1971. Between the trade and value dates the Japanese government declared that non-residents could not open a yen-denominated bank account if they did not already have one before Sept. 6, 1971. United Equities lacked such an account, and thus it was unable to receive delivery of the yen. 383 N.Y.S.2d at 7-8. *See also* discussion regarding forward transactions *supra* note 17.

40. *See United Equities,* 383 N.Y.S.2d at 9-13; *cf.* U.C.C. §§ 2-102 and 2-105(1) (discussing the scope of the Article and definition of "goods," respectively).

41. 556 N.Y.S.2d 258 (N.Y. App. Div. 1990).

42. *Id.* at 262 n.2 (citing New York Annotations to U.C.C. § 2-105 at 97).

43. 568 N.Y.S.2d 333 (N.Y. 1991); *see also* Steven Lipin, *Bankers Trust Wins Suit on Currency Deal,* AM. BANKER, Apr. 2, 1991, at 2. The facts of *Intershoe* are discussed *infra* note 193.

44. *Intershoe,* 568 N.Y.S.2d at 336 (citations omitted).

45. 785 F. Supp. 411, 431 (1992) [hereinafter *IBJ*]. The facts of *IBJ* are discussed *infra* note 202.

insolvent buyer.[46] The court failed to consider whether this remedy is inconsistent with other provisions of the U.C.C., namely, the receiver finality rule of Article 4A.[47]

Perhaps these courts carelessly applied Article 2 to foreign exchange transactions because the official commentary thereto sometimes invites application by analogy. Official comment 1 to section 2-105 states that while investment securities are expressly excluded from the scope of Article 2,

> [i]t is not intended by this exclusion . . . to prevent the application of a particular section of this Article by analogy to securities . . . when the reason of that section makes such application sensible and the situation is not covered by the Article of this Act dealing specifically with such securities (Article 8).[48]

Judges may reason that even if they are incorrect as a matter of law about the scope of Article 2, the statute encourages courts to apply it where appropriate by analogy.

Such reasoning is specious. There is no such invitation issued by the official commentary to judges adjudicating cases involving foreign currency. Moreover, direct application provides parties with the certainty that the entire statute governs. In contrast, application by analogy empowers the court to pick and choose among the provisions of Article 2. Selective application breeds uncertainty.

---

46. 961 F.2d 341, 356 (2nd Cir. 1992). In *Koreag*, Refco engaged in two types of spot foreign exchange transactions with Mebco. First, Refco bought foreign currency from Mebco for $7.4 million U.S. dollars. On April 28, 1989, Refco transferred $7.4 million by wire to Mebco. Second, Refco sold $4.1 million worth of foreign currencies to Mebco in exchange for U.S. dollars. Between April 28 and May 2, 1989, Refco transferred these currencies by wire to Mebco. Mebco, however, was declared insolvent and closed by the Swiss bank regulatory authority on April 27, 1989. The closure occurred before Mebco transferred the foreign currencies Refco had bought for $7.4 million, and before Mebco paid Refco for the $4.1 million in foreign currencies. Refco sought to reclaim the $7.4 million and the foreign currencies it had transferred to Mebco. *Id.* at 344-46.

With respect to the first type of transaction, the court found that Refco was a seller of U.S. dollars and, therefore, was entitled to reclaim the foreign currencies—the "goods"—under U.C.C. §§ 2-310(a), 2-507(2), and 2-702(2). *Id.* at 355-56. Sections 2-310(a) and 2-507(2) are relevant if the foreign exchange transaction is considered a cash sale. U.C.C. § 2-310(a) provides that "[u]nless otherwise agreed, payment is due at the time and place at which the buyer is to receive the goods. . . ." Section 2-507(2) indicates that "[w]here payment is due and demanded on the delivery to the buyer of goods . . . , [the buyer's] right as against the seller to retain or dispose of them is conditional upon his making the payment due." The *Koreag* court found the interaction of these two sections "to create a seller's right to reclaim goods from an insolvent buyer who takes possession of the goods, but fails to tender payment" [citations omitted]. 961 F.2d at 356. Section 2-702(2) is relevant if the transaction is considered a credit sale. The existence of a seller's right of reclamation in the case of a credit sale is clear from the statutory language; § 2-702(2) states, "[w]here the seller discovers that the buyer has received goods on credit while insolvent he may reclaim the goods upon demand made within ten days after [their] receipt." *See also infra* note 74.

With respect to the second type of transaction, the court stated that Refco was a buyer of U.S. dollars and, therefore, Article 2 did not provide a right equivalent to reclamation. *Koreaq*, 961 F.2d at 357.

47. *See infra* note 74.

48. U.C.C. § 2-105 cmt. 1; *see also* B.N.E., Swedbank, S.A. v. Banker, 1993 U.S. Dist. Lexis 2699 (S.D.N.Y. Mar. 10, 1993) (applying the parol evidence rule of § 2-202 in a case involving transactions in debt securities of less developed countries).

The failure of the five courts to assess carefully the scope of Article 2 cannot be tolerated because the stakes in the trillion-dollar-a-day currency bazaar are high. A plenary extension of Article 2 to the currency bazaar must be accompanied by a close reading of the relevant scope provisions of Article 2. Such a reading suggests the issue is more complex and subtle than the courts have acknowledged, and that non-conventional reasoning is needed.

2. Statutory Ambiguity

The first and most important point to acknowledge is that the language of U.C.C. Article 2 does not clearly indicate whether a foreign exchange transaction is included within its coverage. The starting point is section 2-102, which provides that Article 2 applies to "transactions in goods." Nowhere in Article 2 is "transaction" defined.[49] Surely the purchase of yen and sale of dollars by Citibank qualifies as a "transaction" as distinct from a "security transaction" governed by Article 9.[50] Moreover, the Citibank-DBS deal entails neither the provision of a service nor a lease of currency.

The lack of such complications has an important repercussion. The common law has developed the "essence" or "predominant factor" test for deciding whether hybrid contracts and sale-lease deals are subject to Article 2.[51] Problems in deciding whether to apply Article 2 to a hybrid sale-service contract, or categorizing a transaction as a sale governed by Article 2 or a lease governed by Article 2A, do not exist with respect to foreign exchange transactions.[52] Accordingly, the well-developed commercial law jurisprudence on these matters is inapposite. Similarly, frequently used treatises are unhelpful because they dwell on hybrid contracts and sale-lease agreements.[53] In the currency bazaar, an entirely new approach to the scope of Article 2, such as the pragmatic strategy advocated herein, is needed.

The crux of the scope problem, as the *Koreag* court acknowledged, is that the subject of a spot foreign exchange transaction (a currency issued

49. There are hints in Article 2 that the drafters meant that a "transaction" is a "sale." For example, U.C.C. § 2-101 provides that the title of Article 2 is "[s]ales." Section 2-106(1) indicates that a "contract" or "agreement" refer to a "sale of goods." Many sections of Article 2 refer to a buyer and seller. Nevertheless, it is curious that the drafters chose the word "transaction" in § 2-102 instead of "sale."

50. *See* U.C.C. § 9-102. U.C.C. § 2-102 and the official comment thereto make clear that a "transaction" does not include a security transaction.

51. This judicially-created test asks whether the sale of a good was the essence of the contract, or the predominant factor in the transaction, or whether it was merely incidental or collateral to the provision of a service, sale of real estate, or a lease arrangement. *See, e.g.*, Triangle Underwriters v. Honeywell, 604 F.2d 737 (2d Cir. 1979) (holding that the essence of a contract was the sale of goods and the provision of services was merely incidental to that sale); Bonebrake v. Cox, 499 F.2d 951 (8th Cir. 1974) (discussing the predominant factor test).

52. For a discussion of these problems, see Note, *Disengaging Sales Law from the Sale Construct: A Proposal to Extend the Scope of Article 2 of the UCC*, 96 HARV. L. REV. 470 (1982).

53. *See, e.g.*, JAMES J. WHITE & ROBERT S. SUMMERS, UNIFORM COMMERCIAL CODE § 1-1 at 25-26 (3rd ed. 1988) [hereinafter WHITE AND SUMMERS].

by the United States, foreign country, or internationally recognized economic zone[54]) is not necessarily a "good."[55] According to section 2-105(1), "goods" are "all things . . . which are *movable* at the time of identification to the contract for sale other than the *money in which the price is to be paid* . . . and *things in action.*"[56] Movability is central to the concept of a "good."[57] Undoubtedly, foreign exchange in the form of physical currency is movable.[58] On the value date, Citibank is entitled to receive delivery of 120 million yen in physical currency, and DBS is entitled to get $1,153,846.15 in physical currency. The yen and dollar notes could be shipped to the respective parties.

In practice, however, participants in the foreign exchange market—in contrast to a tourist buying foreign currency—do not receive physical delivery of cash. Foreign exchange is transferred from a seller to buyer by a funds (or wire) transfer.[59] A 120 million debit in yen is entered electronically to the bank account of DBS, and a corresponding credit is entered to Citibank's bank account. A debit of $1,153,846.15 is made to Citibank's bank account and a credit of that amount is made to the bank account of DBS. The transfer of yen and dollars around the world occurs in seconds. Thus, buying foreign exchange in the interbank market typically involves buying a bank balance, or bank deposit obligation, denominated in a foreign currency.[60]

---

54. There is an active spot market for the European Union's European Currency Unit (ECU); thus, it would be incomplete to think only in terms of currency issued by individual sovereign nations.

55. 961 F.2d 341, 355 ("[a]s a threshold matter, therefore, it must be determined whether the foreign currencies that Refco and Mebco agreed to exchange were 'goods' within the meaning of section 2-102.").

56. U.C.C. § 2-105(1) (emphasis added). Identification of goods occurs when the contract is made if, as in the hypothetical foreign exchange transaction, it is for the sale of goods already existing and identified. U.C.C. § 2-501(1)(a); *see also infra* note 60. "Money" is defined in U.C.C. § 1-201(24) as "a medium of exchange authorized or adopted by a domestic or foreign government as part of its currency." Thus, money is not narrowly viewed as legal tender, but rather that which has the sanction of government. U.C.C. § 1-201 cmt. 24.

57. U.C.C. § 2-105 official cmt. 1.

58. Manire, *supra* note 7, at 1193, 1196 (erroneously stating that foreign exchange transferred through banking channels—by which he presumably means wire transfer—is not moveable). *See also infra* note 59 and accompanying text.

59. Funds transfers are more commonly known as "wire transfers." *See generally* Ernest T. Patrikis, Thomas C. Baxter, Jr., and Raj Bhala, *Article 4A: The New Law of Funds Transfers and the Role of Counsel*, 23 UCC L.J. 219 (1990) (discussing U.C.C. Article 4A, which governs funds transfers, and other relevant laws, regulations, and private rules).

60. *See* GLENN G. MUNN, ENCYCLOPEDIA OF BANKING AND FINANCE 401 (F.L. Garcia ed., 8th ed. 1983). Foreign exchange transactions "are more accurately described as the transferring by individuals or corporations in one country of credits or debits through their banks by obtaining credits or debits on the books of the banks in other countries that are correspondents or branches of the banks through which the transmission is arranged." *Id.*

An interesting question is whether foreign exchange is identified to the contract. The definition of "goods" in U.C.C. § 2-105(1) indicates that the thing is movable "at the time of identification to the contract for sale." Trading foreign exchange entails entering debits and credits electronically to bank accounts maintained on computers. In a traditional sense—a tangibility paradigm—identification might mean that the buyer of foreign exchange withdraws the funds purchased, handles the physical currency, and re-deposits this currency. Arguably, if foreign exchange is not identified to the contract because of its intangible, electronic nature, then it cannot be a "good" for purposes of Article 2.

Are movable funds like yen and dollars the "money in which the price is to be paid" and, therefore, excluded from being "goods"? The official comment to section 2-105 indicates that foreign currency is not automatically excluded by this phrase:

> The exclusion of "money in which the price is to be paid" from the definition of goods does *not* mean that foreign currency which is included in the definition of money may not be the subject matter of a sales transaction. Goods *is* intended to cover the sale of money when money is being treated as a commodity but not to include it when money is the medium of payment.[61]

The official comment suggests that every foreign exchange transaction consists of two legs, a "commodity leg" and a "payment leg." The Practice Commentary to New York's codification of the definition of "goods" reinforces the bifurcation: " 'money', other than the money in which the price is to be paid, is included in the definition of 'goods': thus a contract for sale of coins or of *foreign currency is* a contract for sale of goods."[62]

Citibank buys yen and pays dollars. The transfer of yen is the "commodity leg," while the transfer of dollars is the "payment leg." From DBS's perspective as a buyer,[63] the situation is the reverse: it buys dollars with yen, thus the dollar transfer is the commodity leg and the yen transfer is the payments leg. If both parties are viewed as buyers, then there is an inconsistency in the legs. Accordingly, one party must be viewed as a seller.[64] DBS, for example, sells yen and receives dollars for the sale. Then, the commodity leg from both parties' perspective is the yen transfer, and the payment leg is dollars.[65]

Nevertheless, the movable yen in the commodity leg may be excluded from the U.C.C. section 2-105(1) definition because the yen are "things in action." The *Intershoe* and *Koreag* courts neglect to consider this possibility. Those courts read the definition of "goods" in U.C.C. section 2-105 as if it stopped with the phrase "other than the money in which the price is to be paid." A "thing in action" is not defined in Article 2. It "denotes a claim or right to personal property not in one's possession, as distinguished from property actually in one's possession."[66] The factual predicate for

---

61. U.C.C. § 2-105 cmt. 1 (emphasis added).

62. N.Y. U.C.C. § 2-105 Practice Commentary note 2 (McKinney 1964) (emphasis added).

63. A "buyer" is "a person who buys or contracts to buy goods." U.C.C. § 2-103(1)(a) (1989).

64. A "seller" is "a person who sells or contracts to sell goods." *Id.* § 2-103(1)(d).

65. It is not strictly necessary for the transaction to have a purchase price in money in order to be governed by Article 2. U.C.C. § 2-304(1) addresses barter transactions, i.e., situations where a good is paid for with something other than money. The purchase price "can be made payable in money or otherwise." U.C.C. § 2-304(1). Because the word "otherwise" includes any form of consideration sufficient to support a contract, it covers foreign currency. *See, e.g.*, Mortimer B. Burnside & Co. v. Havener Sec. Corp., 269 N.Y.S. 724, 726 (N.Y. App. Div. 1966). However, viewing the dollar-yen spot transaction as barter instead of a purchase of yen with money (namely, dollars), or a purchase of dollars with money (namely, yen) would be inappropriate because both dollars and yen are "money" as defined in U.C.C. § 1-201(24). *See supra* note 56.

66. GLENN G. MUNN ET AL., ENCYCLOPEDIA OF BANKING AND FINANCE 181 (9th ed. 1991). The above-quoted definition concerns a "chose" in action which is the same as a "thing" in

classifying yen as a thing in action is intimated above. Citibank is buying a bank credit (an intangible, electronic bank balance, not physical currency) denominated in yen.[67] As one court put it, "[e]ven when the obligation is performed and the credit established, the customer is only the owner of an obligation or chose in action and not of any actual foreign money."[68]

The rationale in two pre-U.C.C. cases seems to contemplate the distinction between physical currency and things in action. In the 1922 case of *Melzer v. Zimmerman*, a New York Supreme Court considered whether the sale of Austrian currency, kronen, was governed by New York's Personal Property Law ("PPL") as a sale of goods.[69] The contract called for delivery of physical currency to the buyer. The court stated the contract in dispute "was an agreement for the purchase by the plaintiff and a sale by the defendant over the counter of said Austrian kronen in the form of currency, and not a deposit account payable by such a draft."[70] Accordingly, the contract fell within the ambit of the PPL. Similarly, in a 1925 case, *Zimmerman v. Roessler & Hasslacher Chemical Co.*,[71] a New York appellate court considered whether a contract for the sale of marks involved "goods" under New York's former Sales of Goods Act. The court observed that the contract contemplated the delivery of marks, not the credit of marks to a bank account, because the parties did not specify a time or

---

action. A "chose in action" is defined in a manner similar to the meaning of a "thing in action," namely, as "a right of bringing an action or right to recover a debt or money." BLACK'S LAW DICTIONARY 241, 1479 (6th ed. 1990). Moreover, "chose" is the French word for "thing." The drafters of the U.C.C. undoubtedly sought to simplify and modernize terminology. *See* U.C.C. § 1-102(2)(a). "Things in action" are discussed in Alphonse M. Squillante, *Commercial Code Review*, 76 COM. L.J. 42 (1971).

67. *See* Manire, *supra* note 7, at 1192 (concluding that foreign exchange transactions involve the sale of choses in action). Indeed, in 1955 the New York Law Revision Commission stated in its study of the Uniform Commercial Code that:

> The more limited exclusion of "the money in which the price is to be paid" is designed to permit the Sales Article to govern a transaction when, in the language of Comment 1, "money is being treated as a commodity": a sale of an ancient Roman coin or a modern coin collection. This leads one on the question whether the Sales Article reaches a transaction for the exchange of dollars into pounds or pesos. Comment 1 (fourth paragraph) contains language which supports such coverage. But the provisions of the Sales Article hardly seemed designed to cope with foreign exchange transactions; perhaps they would be excluded from the article on the ground that they do not involve "things . . . movable" or in any event are "things in action". [sic]

1 STATE OF NEW YORK, REPORT OF THE LAW REVISION COMMISSION, STUDY OF THE UNIFORM COMMERCIAL CODE 362 (1955).

68. Samuels v. E.F. Drew & Co., 296 F. 882, 886 (2d Cir. 1924); *see also* Manire, *supra* note 7, at 1192 n.138 (citing other relevant cases). Of course, even if a court decides foreign exchange is a chose in action, it may well apply Article 2 by analogy. *See, e.g.*, Zamore v. Whitten, 395 A.2d 435 (Me. 1978) (applying the U.C.C. by analogy to a sale of a chose in action) *overruled by* Bahre v. Pearl, 595 A.2d 1027 (Me. 1991). For scholarly treatments of the application of the U.C.C. by analogy, see Jane P. Mallor, *Utility "Services" under the Uniform Commercial Code: Are Public Utilities in For a Shock?*, 56 NOTRE DAME LAW. 89 (1980); Daniel E. Murray, *Under the Spreading Analogy of Article 2 of the Uniform Commercial Code*, 39 FORDHAM L. REV. 447 (1971).

69. 194 N.Y.S. 222, 223-24 (N.Y. Sup. Ct. 1922), *aff'd*, 198 N.Y.S. 932 (N.Y. App. Div. 1923). The *Melzer* court relied on an 1847 decision, Peabody v. Speyers, 56 N.Y. 230 (1874).

70. *Melzer*, 194 N.Y.S. at 223.

71. 207 N.Y.S. 370 (N.Y. App. Div. 1st Dept.), *aff'd*, 148 N.E. 659 (1925).

bank account for such a credit. Accordingly, the contract was held to involve "goods."[72]

The concept of "things in action" and pre-U.C.C. cases suggest that the distinction between (1) a bank credit denominated in foreign currency; and (2) physical currency; should determine whether yen (or any other foreign currency) is a "good" under U.C.C. section 2-105(1). Foreign currency-denominated bank credits are "things in action." Such credits are rights to the payment of money, and these rights, which may be bought and sold in the market, are excluded from Article 2. The supporting logic is a strict interpretation of "goods": Article 2 applies only to paradigmatic goods, which in effect are tangible items, not to transfers of intangible interests such as interests in deposit accounts.

This distinction should not be dismissed lightly. An ill-considered expansion of Article 2 to include purchases and sales of bank credits would open the door to the inclusion of transfers of claims against insurers, corporate debt (to the extent not covered by U.C.C. Article 8), and accounts receivable (to the extent not governed by U.C.C. Article 9). The reach of Article 2 would extend to a far broader array of commercial and financial transactions than its drafters ever envisioned, and the provisions of Article 2 might be ill-suited to the needs of the many and varied transactors. Nevertheless, for the reasons discussed below, the distinction and supporting logic are problematic.[73]

B. *The Aggressively Exclusive Approach*

The leading advocate of the aggressively exclusive approach is the Federal Reserve. For over two years it has acted through the Subcommittee on Payments (Subcommittee) of the American Bar Association's Committee on the Uniform Commercial Code and urged the Subcommittee to adopt a report calling for the exclusion of foreign exchange transactions from the scope of U.C.C. Article 2.[74] The Federal Reserve's position suffers from four serious flaws.

---

72. 207 N.Y.S. at 371-72.

73. *See infra* notes 75-82 and accompanying text.

74. *See, e.g.*, Letter from Thomas C. Baxter, Jr., Deputy General Counsel and Senior Vice President, Federal Reserve Bank of New York, to Members of the Subcommittee on Payments (Mar. 30, 1994) and the accompanying *Outline of Report of Subcommittee on Payments Concerning the Application of the Uniform Commercial Code to a Foreign Exchange Trade* 2-7 (discussion draft, on file with author); Letter from Thomas C. Baxter, Jr., Counsel, Federal Reserve Bank of New York, to Members of the Subcommittee on Payments (Sept. 29, 1993) and the accompanying *Outline of Report of Subcommittee on Payments Concerning the Application of the Uniform Commercial Code to a Foreign Exchange Trade* 2-4, 9-10 (second discussion draft, on file with author). *See generally* Thomas C. Baxter, Jr. & James H. Freis, Jr., *Resolving Funds Transfer Disputes Related to Currency Exchange Transactions: What Law Governs?*, COM. L. ANN. (forthcoming).

The key reason for the Federal Reserve's position is a fear that if such transactions are subject to Article 2, then the receiver finality rule of Article 4A of the U.C.C. will be undermined. Section 4A-405(c) provides that a funds transfer is final and irrevocable upon acceptance of a payment order by the beneficiary's bank on behalf of the beneficiary. U.C.C. § 4A-405(c), 2B U.L.A. 532 (1991). Thus, for example, when DBS's bank accepts dollars on behalf of DBS, the credit of dollars to the DBS account is final. The fear arises from the conclusion that under §§ 2-310(a), 2-507(2), and 2-702(2), a seller of foreign currency has a right to

1. The Stalemate of Pre-U.C.C. Cases

The position relies on outdated, pre-U.C.C. cases—*Melzer* and *Zimmerman*—for guidance in resolving modern-day international commercial and financial problems. The *Melzer* and *Zimmerman* courts were not faced with a large and rapidly growing global currency bazaar. These courts focused on the difference between accepting physical delivery of currency and a deposit account denominated in foreign currency payable by a draft drawn on a foreign correspondent bank.[75] This distinction is arcane insofar as foreign exchange typically moves from seller to buyer by funds transfer.

Furthermore, there are pre-U.C.C. cases that plainly support the proposition that a foreign exchange transaction involves the sale of a "good" or commodity. Three such cases, *Reisfeld v. Jacobs*,[76] *Liepman v.*

---

reclaim currency from a buyer under certain circumstances. *See* Koreag v. Refco F/X Associates, Inc., 961 F.2d at 356.

The Federal Reserve's fear, however, is unfounded. To say that the payment of the dollar or yen leg of the transaction between Citibank and DBS is final for purposes of funds transfer law is one matter. To say that the seller of the currency at issue has a right to reclaim under sales law is a separate matter. There is no inconsistency between the two statements; rather, they simply reflect different legal effects of certain actions. *See* Letter from Patricia B. Fry, Associate Dean for Academic Affairs and Professor of Law, University of North Dakota School of Law, to Raj Bhala, Assistant Professor of Law, Marshall-Wythe School of Law (Nov. 2, 1993) with attached draft letter from Patricia B. Fry to Thomas C. Baxter, Jr., Deputy General Counsel, Federal Reserve Bank of New York 2-4 (on file with author).

The Federal Reserve's concern that *Koreag* conflicts with Donmar Enter., Inc. v. Southern Nat'l Bank of N. Carolina, 828 F. Supp. 1230 (W.D.N.C. 1993) is also unfounded. *See* Letter from Thomas C. Baxter, Jr., Chairman and Deputy General Counsel, Federal Reserve Bank of New York, to Members of the Subcommittee on Payments (Oct. 18, 1993) (on file with author) (stating that *Donmar* "adds fuel to the fire because it seems to conflict with *Koreag*"). In *Donmar*, the plaintiff bought 280,000 pounds from Stephen's Trading Corporation ("STC") for $540,680. The plaintiff transferred by wire the U.S. dollars in two installments, the second one in the amount of roughly $524,000, to the Southern National Bank (SNB). STC received this sum and, in turn, transferred it to a third party. When STC instructed SNB to pay 200,000 pounds to the plaintiff, SNB informed STC that STC's account lacked 200,000 pounds to make the payment. SNB informed the plaintiff of the insufficiency, and the plaintiff sought to recover the $524,000 it had paid to STC. 828 F. Supp. at 1233-34.

The Federal Reserve's fear of an inconsistency between *Donmar* and *Koreag* is unfounded because the plaintiff in *Donmar*, in contrast to Refco in *Koreag*, never argued it was a seller of dollars under Article 2 and thereby entitled to a right of reclamation. Indeed, no Article 2 provision was raised in *Donmar*. The plaintiff argued that it had a right to reclaim the $524,000 under § 4A-207 (which concerns payment orders that do not identify a beneficiary), and the court properly rejected the argument. *Donmar*, 828 F. Supp. at 1239.

Two additional concerns might lie behind the Federal Reserve's position. First, reclamation could allow a creditor of a failed bank to circumvent a foreign insolvency proceeding. Second, reclamation could undermine regulatory efforts to develop systems for netting (or off setting) foreign exchange delivery obligations among the players in the global currency bazaar. These concerns, like the fear discussed above, are dubious. *See* Raj Bhala, *Self-Regulation in Global Electronic Markets Through Reinvigorated Trade Usages*, 31 IDAHO L. REV. (forthcoming 1995) (manuscript at 33-35, on file with author).

75. *See Melzer*, 194 N.Y.S. at 223; *Zimmerman*, 207 N.Y.S. at 371.

76. 176 N.Y.S. 223 (N.Y. App. Div. 1919). In *Reisfeld*, the buyer purchased Russian rubles. At issue was whether the contract of purchase was enforceable under the statute of frauds in New York's Personal Property Law whose coverage excluded "money." The notes were issued by the Tsarist government that had been overthrown in the 1917 Bolshevik revolution and, therefore, could not be used as a medium of payment. The court decided that because the buyer purchased the notes for resale, they were not excluded from the

Rothschild,[77] and *Richard v. American Union Bank*,[78] established an intended use test. They focused on the intent of the buyer of foreign currency to determine whether foreign currency is a commodity or a means of payment under the relevant statute of frauds. Under the reasoning of these cases, if Citibank intends to resell the 120 million yen, then the yen are commodities and fall within Article 2. Alternatively, if Citibank plans to use the yen in Japan to purchase goods (other than foreign currency) and services, then the yen are a medium of payment and excluded from the statute. All three cases held that foreign currency was a commodity subject to the relevant statute.[79] Thus, it is not difficult to line up pre-U.C.C. cases against, as well as for, the Federal Reserve's position.

Of course, the pre-U.C.C. cases against that position are as outdated and unworthy of reliance as the cases in the Federal Reserve's favor. For example, classifying foreign exchange as a commodity versus a medium of payment based on the buyer's intent raises doctrinal and evidentiary problems. Should an objective or subjective test be used to determine intention? How should intention be proved? Moreover, the distinction between resale and use as a medium of payment makes little sense. Suppose Citibank intends to resell the 120 million yen to another bank. Why should this case be treated differently from a case where a U.S. importer of Japanese goods buys 120 million yen in order to pay a Japanese exporter for the goods? Finally, the intended use test suggested by the *Reisfeld*, *Liepman*, and *Richard* courts is at variance with the plain meaning of "goods" in section 2-105. The U.C.C.'s definition does not contemplate a focus on anything other than the immediate transaction.[80] In sum, the net result from reviewing pre-U.C.C. cases is an unhelpful and essentially irrelevant stalemate.

2. A Formalistic Distinction

The second problem with the aggressively exclusive approach is that it relies on an overly formalistic distinction—physical currency versus bank credit—to determine the scope of U.C.C. Article 2. Just because Citibank

---

statute. Its holding also rested on the characterization of the rubles as choses in action. *Id.* at 224.

77. 262 S.W. 685 (Mo. Ct. App. 1924). The contract in *Liepman* involved the purchase of marks and the issue was enforceability of the contract under a statute of frauds that covered goods but excluded "money." *Id.* at 685. Because the U.S. and Germany were at war, the court reasoned, the buyer could not have intended to use the marks as a medium of payment. Therefore, they were a commodity within the statute of frauds. *Id.* at 686.

78. 170 N.E. 532 (N.Y. 1930). This case involved a contract to buy two million Romanian lei. The contract called for a cable transfer of foreign currency into the buyer's bank account in Bucharest, but delivery was delayed and the buyer sued for damages. *Id.* at 533-34. The court held that the transaction was for a commodity because the buyer intended to resell the lei in the U.S., not spend the money in Romania. *Id.* at 535.

79. *See Resfeld*, 176 N.Y.S. at 224; *Liepman*, 2626 S.W. at 686; *Richard*, 170 N.E. at 535.

80. *Compare* U.C.C. § 2-105(1) (defining "Goods" as "all things . . . movable at the time of identification to the contract for sale.") *with Reisfeld*, 176 N.Y.S. at 224 (noting "that the rubles were bought for resale") *and Liepman*, 262 S.W. at 686 (because of the war with Germany, the marks were intended to be a commodity) *and Richard*, 170 N.E. at 535 (buyer intended to resell the lei in the U.S.).

obtains a credit of 120 million yen does not prohibit it from withdrawing that credit in physical currency. Presumably, unless a buyer and seller agree otherwise, a buyer of foreign exchange has the right to demand delivery of physical currency in place of a credit. The fact that Citibank elects to keep its yen in the form of a bank balance on deposit instead of physical currency should not affect Citibank's contractual rights and obligations. Citibank's decision is merely one of convenience, economy, and security. It is easier, cheaper, and safer to store 120 million yen electronically than in bills in a Brooklyn warehouse.

More fundamentally, the scope of Article 2 should not depend on whether the buyer converted bank credits to cash. If it did, then form would triumph over substance and form could be manipulated. Citibank could easily "shop" among legal regimes, opting into or out of Article 2, by choosing the form of delivery. Only an agreement with DBS that Citibank would maintain the yen in a specific form would limit Citibank's freedom to manipulate the legal rights and duties of the parties. A much more pragmatic evaluation—whether Article 2 serves the needs of Citibank and DBS—ought to determine the scope of the statute.

3. The Drafters' Intention

The exclusionary approach may be inconsistent with the intention of the drafters of the U.C.C. The official comment to section 2-105 clearly suggests that the drafters intended to include the sale of money as a commodity within the scope of the Article.[81] It is not a question of the drafters failing to foresee foreign exchange transactions—the comment evinces this foresight. Rather, the drafters did not anticipate that foreign exchange would occur electronically and that it would appear in computerized bank account records. If they had envisioned these developments, then they might have incorporated a definition of "things in action" to avoid confusion as to scope.

More generally, aggressively excluding foreign exchange transactions reflects an intransigence that is inconsistent with the drafters' goal of creating a workable and adaptable statute. They wanted the statute to be interpreted flexibly and liberally.[82] Only then could it could be adapted to new commercial contexts like the currency bazaar.

4. Underlying Principles

The exclusionary approach may be inimical to the underlying principles of the U.C.C.: simplification,[83] modernization,[84] uniformity,[85] cer-

81. *See, e.g.*, New York U.C.C. § 2-105, Practice Commentary, at 94-95 (McKinney 1964); § 2-105, New York Annotations, at 97 (McKinney 1964). *But see id.* § 2-201, New York Annotations, note (1)(a) at 119-20 (McKinney 1964) (stating that the original version of the Uniform Sales Act governed contracts for choses in action but later the Act was amended to exclude choses in action).

82. *See* U.C.C. § 1-102 cmt. 1; *infra* notes 98-105 and accompanying text.

83. U.C.C. § 1-102(2)(a).

84. *Id.*

85. *Id.* § 1-102(2)(c).

tainty,[86] and support for commercial transactions.[87] If foreign exchange transactions are wholly excluded from Article 2, then they become subject to non-uniform, obsolete legal regimes, namely, pre-U.C.C. common law or non-U.C.C. sales statutes.[88] Yet, pre-existing sales law was simplified and modernized by Article 2,[89] and Article 2 unified the sales law of various states.[90] Potentially applying the sales laws of fifty different states renders the law of sales as it pertains to foreign exchange transactions more complex, arcane, and disjointed.[91]

This possibility generates uncertainty for foreign exchange market participants. Providing a precise and predictable legal framework that enhances certainty for businesspersons is a founding principle of Article 2.[92] Taking foreign exchange transactions out of this framework obviously contradicts this principle. Uncertainty is further exacerbated if courts, seeking to keep a foreign exchange transaction within Article 2, manipulate facts to make foreign exchange more closely resemble a paradigmatic good.

The drafters of Article 2 sought a sales statute that would foster commercial development.[93] Tossing foreign exchange transactions to the vagaries of non-U.C.C. law may inhibit innovation in the foreign exchange market.[94] For example, market participants like Citibank and DBS may face difficulty devising a uniform practice of confirming transactions electronically because many different statutes of frauds potentially apply. The need to protect ongoing commercial relations is concomitant with the

---

86. While certainty is not expressly mentioned in U.C.C. § 1-102, the principles set forth in that section imply increased certainty as to legal rights and obligations. It has been argued, however, that simplification and clarification are fallacious principles, *see,e.g.*, Charles E. Clark, *The Restatement of Contracts*, 42 YALE L.J. 643, 653 (1933), and costly to achieve. *See, e.g.*, James Gordley, *European Codes and American Restatements: Some Difficulties*, 81 COLUM. L. REV. 140, 156-7 (1981).

87. U.C.C. § 1-102(2)(b).

88. An analogous problem exists for hybrid sales-service contracts and (to a lesser extent given the enactment of Article 2A) lease contracts. An example of a non-U.C.C. sales statute is New York's General Obligations Law. *See infra* note 108.

A tension may exist between providing certainty and predictability for commercial parties through codification, on the one hand, and stifling innovation in a market as a result of codification, on the other hand. The difficult, time-consuming nature of revising a code—as the current experience with Article 2 illustrates—reinforces concerns about inhibiting the expansion of new commercial practices. *See* Marion W. Benfield, Jr. & Peter A. Alces, *Reinventing the Wheel*, 35 WM. & MARY L. REV. 1405 (1994).

89. For a discussion of the need to simplify and update sales law, see Arthur L. Corbin, *The Uniform Commercial Code—Sales; Should It Be Enacted?*, 59 YALE L.J. 821, 834-5 (1950); Karl Llewellyn, *The General Scope of the Uniform Commercial Code*, 1950 N.J. ST. B.A.Y.B. 73, 75.

90. For a discussion of the importance of unifying sales law, see Llewellyn, *supra* note 89, at 73.

91. *See, e.g.*, Murray, *supra* note 68, at 456 (discussing wide variations in non-U.C.C. standards among states).

92. *E.g.*, In re Automated Bookbinding Serv. Inc., 471 F.2d 546, 552 (4th Cir. 1972).

93. *See, e.g.*, Llewellyn, *supra* note 89, at 73.

94. For a discussion of how the solar energy industry was harmed by the denial of the application of the implied warranties of Article 2 to the sale and installation of solar energy devices, see Harry R. Wright, Jr., Comment, *The Sales-Service Dichotomy: A Roadblock to Consumer Acceptance of Domestic Solar Energy Devices*, 30 MERCER L. REV. 547, 552-54 (1979).

need to support commercial development.[95] Currency bazaar participants like Citibank and DBS are repeat players—they are well known to each other and deal with one another on a daily basis. Application of Article 2 may place these relations, particularly with respect to matters of contract formation, on a firm legal foundation.[96]

C. *The Pragmatic Strategy*

Official comment 1 to section 1-102 aims to prevent U.C.C. Article 2 from aging while maintaining its strength as a foundation for commercial parties:

> This Act is drawn to provide flexibility so that, since it is intended to be a semi-permanent piece of legislation, it will provide its own machinery for expansion of commercial practices. It is intended to make it possible for the law embodied in this Act to be developed by the courts in the light of unforeseen and new circumstances and practices.[97]

The drafters knew subsequent generations of lawyers would face new transactional horizons. They realized Article 2 would be discarded if it failed to satisfy the needs of the parties.[98] Thus, they drafted Article 2 so it could adapt to a variety of contractual contexts that were not overtly contemplated by the drafters.[99] The statute's style is deliberately "loose" and "open-ended."[100] Many of its provisions articulate standards of contract law and principles of justice rather than specific, technical rules.[101]

Neither the carelessly inclusive approach of the *United Equities, Saboundjian, Intershoe, IBJ,* and *Koreag* courts, nor the aggressively exclusive approach of the Federal Reserve, heeds the advice of the drafters. The first approach lacks articulated standards and reasoning. The second approach is rigid. Both approaches neglect new ways of conducting business and the realistic needs of businesspersons. A new pragmatic strategy is

---

95. *See* Eugene F. Mooney, *Old Kontract Principles and Karl's New Kode: An Essay on the Jurisprudence of Our New Commercial Law,* 11 VILL. L. REV. 213, 230 (1966).

96. One practical mechanism that repeat players can use to avoid a statute of frauds problem is to waive the statute of frauds defense in a model or master agreement. Strangely, the IFEMA contains no such waiver. *See generally,* IFEMA, *supra* note 14. If it had the waiver, then the problem would be acute only for non-repeat players, *i.e.,* non-parties to the IFEMA.

97. U.C.C. § 1-102 cmt. 1.

98. Indeed, "[t]he origins of the Uniform Commercial Code lie in the law merchant, a specialized body of usages, or customs, that governed contracts dealing with commercial matters until the seventeenth century." 1 E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS § 1.9, at 34 (1990). The rules laid down in the U.C.C. can be viewed as a reflection of what the drafters thought was the best market practice of merchants. Not surprisingly, therefore, a number of provisions of Article 2 impose higher standards on merchants than on other sorts of parties. *Id.* § 1.10, at 42-43.

99. *See* William D. Hawkland, *Uniform Commercial 'Code' Methodology,* 1962 U. ILL. L.F. 291, 314 (1962); Mitchell Franklin, *On the Legal Method of the Uniform Commercial Code,* 16 LAW & CONTEMP. PROBS. 330, 333 (1951); *see also* Farnsworth, *supra* note 98, § 1.10, at 40-41 (discussing the adaptability of a general body of contract law to many different types of transactions).

100. Grant G. Gilmore, *In Memoriam: Karl Llewellyn,* 71 YALE L.J. 813, 814 (1962); *see also* Homer Kripke, *The Principles Underlying the Drafting of the Uniform Commercial Code,* 1962 U. ILL. L.F. 321, 328.

101. Note, *Disengaging Sales Law from the Sale Construct: A Proposal to Extend the Scope of Article 2 of the UCC,* 96 HARV. L. REV. 470, 484 (1982).

needed to decide whether foreign exchange fits the definition of "goods" and, accordingly, whether foreign exchange transactions are subject to Article 2. This model gains support from Article 2's invitation to courts and commentators to consider the needs of commercial parties by expanding the scope of the statute accordingly.[102]

Under the pragmatic strategy, the determining factor is whether Article 2 furnishes rules that meet the needs of participants in the currency bazaar. Insofar as Article 2 serves these needs, it retains its vitality in new transactional settings. In turn, it promotes certainty in foreign exchange dealings and the development of the foreign exchange market. Article 2 should be tailored to avoid the undesirable repercussions of applying the statute of frauds to foreign exchange transactions. Such a revision would follow the pragmatism countenanced by the drafters of the official comment quoted above.[103]

The *sine qua non* of the pragmatic strategy is the acknowledgment that foreign exchange trading is inherently fraught with uncertainties and risks. For example, Citibank buys yen and sells dollars because it expects yen to appreciate relative to the dollar. There is currency risk—yen could depreciate and Citibank could lose money on the transaction. Rules of sales law, such as the statute of frauds, potentially exacerbate the uncertainties and risks associated with foreign exchange trading. Uncertainties and risks should be minimized.

Put another way, problems that Citibank and DBS might incur in their dollar-yen deal should be resolved efficiently. Efficiency in the context of the foreign exchange market is a two-dimensional concept involving certainty and cost. Market participants need unambiguous rules setting forth whether foreign exchange contract obligations are enforceable. The rules should not impose unnecessary transaction costs on the participants. In sum, rules of contract enforceability that enhance certainty and reduce cost should be the legal cornerstone for wealth-generating foreign exchange transactions.[104] To the extent sales law does not serve as this cornerstone, it should be changed.

It is beyond the scope of this article to examine every provision of Article 2 from this pragmatic perspective in order to decide conclusively whether the statute ought to govern foreign exchange transactions. The

---

102. A feature of the pragmatic strategy is that it is designed for cases where it is arguable whether a transaction should be included within the scope of Article 2. Because real estate transactions, for example, clearly are excluded, the strategy is inapplicable. But how close is close enough? There is no attempt herein to set forth criteria as to how persuasive the arguments must be for and against inclusion in Article 2 before the pragmatic strategy should be used.

103. *See* U.C.C. § 1-102 cmt. 1.

104. Another important goal, appropriate risk-allocation, should be noted. The risks that the rules address should be allocated between two banks engaged in a foreign exchange transaction according to a "better position" criterion—the bank that can most cheaply insure against the loss should bear the risk in question. With respect to the statute of frauds, the risk that a contract is unenforceable is evenly distributed. Similarly, if the proposals discussed below regarding the statute of frauds are implemented, this risk would be evenly distributed.

pragmatic strategy, however, can yield a tentative resolution to the scope problem within the confines of one article. It can be applied to an Article 2 provision of fundamental importance, the statute of frauds. As Professor Farnsworth writes, "[i]t would be difficult to imagine a question more important to a person expecting to make agreements in an unfamiliar legal system than this: when is a writing required to make an agreement enforceable?"[105] Further, the statute of frauds set forth in section 2-201 has caused consternation among foreign exchange market participants. How this section affects the market is an important piece in the puzzle of the scope of Article 2.

IV. ENFORCEABILITY—DOES THE STATUTE OF FRAUDS SERVE MARKET NEEDS?

Had the statute of frauds been always carried into execution according to the letter, it would have done ten times more mischief than it has done good, by protecting, rather than by preventing, frauds.[106]

The statute of frauds, set forth in section 2-201(1), states:

> [e]xcept as otherwise provided in this section a contract for the sale of goods for the price of $500 or more is not enforceable by way of action or defense unless there is some *writing* sufficient to indicate that a contract for sale has been made between the parties and *signed* by the party against whom enforcement is sought or by his authorized agent or broker.[107]

---

105. 2 FARNSWORTH, *supra* note 98, § 6.1.

106. 1 WILLIAM BLACKSTONE, COMMENTARIES 601 (1844) (quoting Justice Wilmot's concurrence in Simon v. Metivier); Robert L. Misner, *Tape Recordings, Business Transactions Via Telephone, and the Statute of Frauds*, 61 IOWA L. REV. 941, 942 (1976).

107. U.C.C. § 2-201(1) (emphasis added). Because of the tense used in § 2-201 (namely, the words "has been made") it is evident that the statute of frauds does not require that a contract be in writing; it can be oral. Rather, the statute of frauds requires that a document exist to provide reliable evidence of the existence of the contract. Otherwise it is not enforceable. *See, e.g.*, Monetti, S.P.A. v. Anchor Hocking Corp., 931 F. 2d 1178, 1182, 1185 (7th Cir. 1991) (Posner, J.). This distinction between contract formation and enforceability is seen in the IFEMA. Section 8.3 of the IFEMA indicates a foreign exchange contract is oral and that the IFEMA is the written evidence thereof. *See* IFEMA, *supra* note 14.

In attempting to show that a writing sufficient to satisfy the statute of frauds exists, it is impermissible to rely on oral testimony. *See Monetti*, 931 F.2d at 1181; Southmark Corp. v. Life Investors, Inc., 851 F.2d 763, 767 n.5 (5th Cir. 1988). The writing, however, need not take the form of a single document. The integration of several documents prepared at different times may satisfy the statute of frauds. *See* Migerobe, Inc. v. Certina USA, Inc., 924 F.2d 1330, 1333 (5th Cir. 1991); Hunt Oil Co. v. FERC, 853 F.2d 1226, 1241 (5th Cir. 1988). The writing or writings need not be directed or delivered to the other party, nor do they have to be made for purposes of satisfying the statute of frauds. *See* 2 FARNSWORTH, *supra* note 98, § 6.7, at 406-07.

The statute of frauds applies prospectively as well as retrospectively. For example, it applies to future oral modifications of a contract that originally fell within the statute. Also, if a contract that did not originally fall within the ambit of the statute of frauds is modified, and the modified contract falls within the statute, then the requirements of the statute must be satisfied. U.C.C. § 2-209(3).

Some commentators suggest that the statute of frauds does not cover every transaction that is within the scope of Article 2. *See, e.g.*, 2 FARNSWORTH, *supra* note 98, § 6.6, at 402-03 & n.5. As discussed above, U.C.C. § 2-102 states that Article 2 governs "transactions" in goods. *See supra* note 50 and accompanying text. In contrast, § 2-201(1) quoted above refers to

If section 2-201(1) were applied strictly to the Citibank-DBS dollar-yen spot agreement, then it would be unenforceable.[108] Assuming written confirmations are not exchanged, Citibank and DBS have nothing to sign, and clearly the value of the transaction exceeds $500. The tape recording of the telephone conversations between the traders does not satisfy the writing requirement of the statute of frauds.[109]

More generally, all of the foreign exchange transactions in which written confirmations are not exchanged are unenforceable if the Article 2 statute of frauds is strictly construed.[110] Immediately, then, it is apparent

---

"contract[s]" for the sale of goods. This reasoning, however, does not appear to have been widely accepted; hence it seems unlikely that a court would rule that foreign exchange transactions are governed by Article 2 but excluded from the statute of frauds on the basis of the "transaction"/"contract" distinction.

108. Courts sometimes confuse enforceability with validity. *See, e.g.*, Tri-State Petroleum Corp. v. Saber Energy, Inc., 845 F.2d 575, 579 (5th Cir. 1988) (stating that the litigants are not "asserting that any oral contract is invalid because of the statute of frauds"). A contract is not void just because it fails to satisfy the statute of frauds. This failure means that the transaction cannot be "judicially enforced in favor of a party to the contract." U.C.C. § 2-201 cmt. 4; *see also* Glover School & Office Equip. Co. v. Dave Hall, Inc., 372 A.2d 221, 223 (Del. Super. Ct. 1977) (stating "the beginning premise is that an oral contract is valid and enforceable unless prohibited or restricted by some statutory provision [such as the Statute of Frauds]").

This result contrasts with that which would be obtained under § 5-701(a) of New York's General Obligations Law ("G.O.L."), a non-U.C.C. statute of frauds. The G.O.L. is potentially relevant to the global currency bazaar because the U.S. is one of the three largest foreign exchange trading centers in the world. The three countries with the largest average daily turnover of foreign exchange transactions are the United Kingdom ($300 billion daily), the U.S. ($192 billion daily), and Japan ($126 billion daily). Central Bank Survey, *supra* note 3, at 13-14. Undoubtedly, New York accounts for the bulk of the U.S. activity. Section 5-701(a) of the G.O.L. states that a contract is absolutely void "unless it or some note or memorandum thereof be in writing, and subscribed by the party to be charged therewith, or by his lawful agent, if such [contract]. . .[b]y its terms is not to be performed within one year from the making thereof." N.Y. GEN. OBLIG. LAW. § 5-701(a)(1) (McKinney 1989). This statute of frauds would not render a spot foreign exchange contract void because it is performed within two days. (Nor would § 5-701(a) of the G.O.L. render a forward foreign exchange contract void so long as the value date of the contract is within one year. Most forward contracts are completed in less than one year). Hence, there is a potential conflict between U.C.C. § 2-201, which renders a spot contract unenforceable, and G.O.L. § 5-701(a)(1), which does not render it void. Under U.C.C. § 1-103, which states that pre-U.C.C. law is applicable unless displaced by provisions of the U.C.C., § 2-201 prevails. *See* H & W Indus. v. Formosa Plastics Corp., 860 F.2d 172, 180 (5th Cir. 1988). For a discussion of the application of this statute of frauds, see Cathy L. Scarborough, *Foreign Exchange Contracts: What Statute of Frauds Applies in New York?*, 4 INT'L L. PRACTICUM 17, 20-21 (1991).

109. *See infra* notes 117-66 and accompanying text. For a recent discussion of electronic signatures and the requirement of a signed writing, see Sharon F. DiPaolo, Note, *The Application of the Uniform Commercial Code Section 2-201 Statute of Frauds to Electronic Commerce*, 13 J.L. & COM. 143 (1993).

110. Section 2-201 is not the only statute of frauds in the U.C.C. that would render such transactions unenforceable. For example, U.C.C. § 8-319 sets forth a writing requirement for the sale of investment securities. *See infra* note 130. Similarly, U.C.C. § 9-203 establishes a requirement for security agreements.

In the context of foreign exchange transactions, U.C.C. § 1-206 could be relevant. It is a residual, gap-filling provision that covers sales of personal property not otherwise covered by the aforementioned sections. U.C.C. § 1-206(2) & cmt. U.C.C. § 1-206(1) states that a contract for the sale of personal property is unenforceable beyond $5,000 "unless there is some writing which indicates that a contract for sale has been made between the parties at a defined or stated price, reasonably identifies the subject matter, and is signed by the party against whom enforcement is sought or by his authorized agent."

The official comment to U.C.C. § 1-206 states that two "principal gap[s]" are filled by that section. The first gap relates to the sale of general intangibles as defined in U.C.C. § 9-

that strictly applying section 2-201(1) could wreak destructive havoc in the currency bazaar.[111] This result indicates section 2-201(1) does not serve the needs of market participants: no law should threaten the foundations of an otherwise well-functioning market that is worthy of preservation. This conclusion is reinforced by the application of the pragmatic strategy—specifically, the analyses of the costs, benefits, and purposes of the statute—set forth below.[112] Thus, the model calls for the rejection of Citibank's affirmative defense. The model also suggests DBS's response has merit, namely, that foreign exchange should be excluded from section 2-105(1) and thus from the reach of section 2-201(1). This argument, however, cannot yet be accepted.

A. *The Clash of Cultures*

There is an uneasy tension between the technology and business practices of the foreign exchange market on the one hand, and the demands of contract enforceability rules in sales law on the other hand. The technology is telephonic. It expands the ways in which market participants negotiate and execute currency trades. Communications between Citibank, DBS and the like are not face-to-face meetings in which written draft contracts are exchanged and marked up by lawyers representing the

---

106. *See also* 2 FARNSWORTH, *supra* note 98, § 6.6, at 404. As discussed above, foreign exchange can be categorized as a thing in action. *See supra* notes 66-72 and accompanying text; U.C.C. § 9-106. Section 9-106 specifically includes things in action in the definition of "[g]eneral intangibles."

The second gap concerns transactions excluded from Article 9 by U.C.C. § 9-104. Section 9-104(l) of the U.C.C. indicates that Article 9 is inapplicable "to a transfer of an interest in any deposit account." A "deposit account" covers "a demand, time, savings, passbook or like account maintained with a bank." U.C.C. § 9-105(e). Certainly, accounts maintained by Citibank and DBS for foreign exchange trading purposes would be "deposit accounts" excluded from Article 9.

Consequently, contracts for foreign exchange could fall into either or both of the gaps and thereby come within the ambit of § 1-206. This result is obtained only if foreign exchange is a thing in action and not a good governed by Article 2 and its statute of frauds.

Such a result would not affect the analysis in this article. The same concerns about the costs, benefits, and purposes of the statute of frauds are relevant to U.C.C. § 2-201 or § 1-206. Indeed, because U.C.C. § 2-201 requires that a writing contain only a quantity term, while § 1-206 requires more, the arguments below apply to U.C.C. § 1-206 a fortiori. *Compare* U.C.C. § 2-201 cmt. 1 *with* U.C.C. § 1-206(1). *See generally* Note, *The Uniform Commercial Code, Section 1-206—A New Departure in the Statute of Frauds?*, 70 YALE L.J. 603 (1961) (comparing the requirements of § 1-206 and § 2-201).

111. This unhappy scenario has spurred legislative action in New York. Recently, the New York legislature passed an amendment to G.O.L. § 5-701 that excludes "qualified financial contracts" from the statute of frauds. Act of July 20, 1994, ch. 467, sec. 1, § 5-701(b), 1994 N.Y. LAWS 467 (codified as amended at N.Y. GEN. OBLIG. LAW § 5-701 (McKinney 1994). Such contracts include spot and forward foreign exchange contracts and currency swaps. The bill was transmitted to Governor Mario Cuomo for his signature on July 8, 1994. N.Y.A 11513, 215th G.A., 2d Sess. (1994). Similar amendments are proposed in the bill for §§ 1-206 and 2-201, which raises the specter of non-uniformity in New York's U.C.C. vis-a-vis the U.C.C. of other states. The bill is actively supported by the ISDA, a trade association representing participants in the over-the-counter derivatives markets. *See* Memorandum from Daniel Cunningham and Catherine Struve to ISDA Board of Directors re: Amendment of New York Statute of Frauds (June 2, 1994) (on file with author). The ISDA memo is accompanied by a form letter favoring the amendments that ISDA members are encouraged to send to New York legislators.

112. *See infra* notes 167-216 and accompanying text.

parties during endless rounds of coffee and take-out sandwiches. The trading floors of Citibank and DBS are entirely different from the conventional lawyers' conference room; traders often communicate by telephone. In sum, the deals made in the currency bazaar are oral and are concluded rapidly and informally.[113]

The statute of frauds must adapt to this telephonic technology. England accepted a similar proposition many years ago. In 1954, the English Law Reform Committee successfully advocated the repeal of the statute of frauds, in part because a writing requirement is " 'out of accord with the way in which business is normally done.' "[114] Unfortunately, this proposition has not gained respectability in the U.S. Except for the current noteworthy attempt at revising U.C.C. Article 2, "[t]here has been no serious movement to abolish the statutes of frauds in this country, though they have had many critics."[115] Foreign exchange market participants might not reduce their agreements to writing for good reason. Because bid-ask spreads are thin for trading in liquid currencies, profits are made through a high volume of trading. To maximize profits, market participants seek to conclude as many transactions as cheaply and quickly as possible. Outdated legal formalities like the statute of frauds requirements lead to higher transaction costs and delay the completion of transactions. Not surprisingly, many market participants prefer tape recordings of conversations among traders instead of written agreements.

The law also must account for the culture of the currency bazaar. Trust among participants in the foreign exchange market is high. Perhaps this aspect of business culture also distinguishes the trading floor from the conference room. The participants repeatedly deal with one another. To engage in fraudulent or deceptive practices is to invite ostracism: a trader's unctuous behavior quickly becomes widely known and other traders decide it is risky and imprudent to deal with the rogue trader. In 1966, an author of a textbook on foreign exchange observed that "[m]ost dealers with very long experience have never had a single lawsuit arising from misunderstandings in respect of foreign exchange transactions."[116] The observation remains true today.

Two variations of the hypothetical spot transaction presented in Part II will illustrate that the statute of frauds should be reformed or abolished in the context of the foreign exchange market. In Case One, the telephone conversation between the Citibank and DBS traders is taped. However, no written confirmations are exchanged between the parties. Case

---

113. Undoubtedly, these features, and the phenomenon of telephonic and computer-to-computer communications, are found in many other modern markets. Accordingly, the potential applicability of the pragmatic strategy is not limited to the foreign exchange market.

114. 2 FARNSWORTH, *supra* note 98, § 6.1, at 371 (quoting UNITED KINGDOM LAW REVISION COMMITTEE ON THE STATUTE OF FRAUDS AND THE DOCTRINE OF CONSIDERATION, SIXTH INTERIM REPORT, CMD. NO. 5449, 6-7 (1937)).

115. 2 FARNSWORTH, *supra* note 98, § 6.1, at 371. Farnsworth is referring specifically to non-U.C.C. statutes of frauds enacted in most states, but his remark is equally true with respect to U.C.C. § 2-201.

116. PAUL EINZIG, A TEXTBOOK ON FOREIGN EXCHANGE 41 (1966).

One raises the problem of compliance with section 2-201(1), the fundamental statute of frauds provision. The problem is solved if section 2-201(1) is judicially reinterpreted or legislatively modified so that tape recordings satisfy the "writing" and "signature" requirements of that section.

In Case Two, the operations departments of Citibank and DBS exchange such confirmations subsequent to the tape-recorded, telephonic conversation between their traders. This Case raises a the problem of compliance with section 2-201(2), the merchant's exception to section 2-201(1). It is argued that the microeconomic costs of such confirmations do not justify the requirement that they be used.

B. *Case One: Tape Recordings as "Writings"*

1. The Tangibility Paradigm

Twenty years ago Professor Misner argued that tape recordings should satisfy the statute of frauds.[117] His argument, however, is incongruous. He did not advocate that a tape recording be considered a "writing" for purposes of U.C.C. sections 1-201(46) and 2-201. Instead, Misner argued that the recording's voiceprint should satisfy the "signature requirement of U.C.C. section 1-201(39).[118] The real problem with Misner's argument, however, is that it is disappointingly conservative.

Misner's argument is mired in the tangibility paradigm. This paradigm does not consider the needs of a particular market like the currency bazaar. The market must fit the law, not the reverse. Most importantly, this paradigm requires some physically cognizable piece of paper to evidence a contract under the statute of frauds. A " 'voiceprint'—the graphic output of high-speed sound spectrograph"[119]—meets the requirement. Thus, Misner's argument does not strike at the heart of the statute of frauds: it will not allow a tape recording alone to satisfy the statute.

Ironically, in trying to cope with telephone business deals, Misner in effect advocated two writings to satisfy the statute of frauds—the voiceprint (which would be the "signature") combined with a written memorandum (which would be the "writing").[120] This solution is not only needlessly cumbersome and costly, but also inapposite to the special technologies and business practices of the currency bazaar. A more radical and efficient solution is required, whereby the tangibility paradigm is discarded and full legal effect is given to tape recordings.

2. Possible Legislative Amendments

One such solution—rejected by Misner with little reasoning[121]— is legislative action. Two straightforward legislative amendments might be appropriate. U.C.C. section 2-201(1) states the "writing" must be "signed"

117. Misner, *supra* note 106.
118. *Id.* at 942, 945-56, 964.
119. *Id.* at 946.
120. *Id.* at 950.
121. *Id.* at 943-46.

by the party (or her broker) against whom enforcement of the contract is sought. Under section 1-201(46), a " 'writing' includes a printing, typewriting or any other intentional reduction to tangible form."[122] A legislative amendment is needed to bring tape recordings squarely within the meaning of a "writing." The following phrase could be added at the end of section 1-201(46) to accomplish this goal: "and a tape recording or other recording on magnetic tape."[123]

Under U.C.C. section 1-201(39), " 'signed' includes any symbol executed or adopted by a party with present intention to authenticate a writing." To ensure consistency with the term "writing" as amended, the following phrase could be added at the end of section 1-201(39): "and any statement of a speaker identifying the speaker that is tape-recorded or recorded on magnetic tape."

Amending section 1-201(46) in the manner previously suggested, however, may be incongruous with the structure of U.C.C. Article 3 and other non-Article 3 provisions dealing with negotiable instruments. Rules regarding negotiable instruments are property rules that turn on the physical delivery of a tangible item. Such rules become nonsensical in an intangible world. Accordingly, adjusting section 1-201(46) could do more harm than good.[124] It is wise to consider two other legislative options in lieu of amending sections 1-201(46) and 1-201(39).

One alternative is to add a definition of "record" in section 1-201. Indeed, in April 1994 the American Bar Association's Working Group on Electronic Writings and Notices of the Subcommittee on Electronic Commercial Practices ("Working Group") approved a proposed definition of "record" that would embrace tape recordings: " '[r]ecord' means a durable representation of information which is in, or is capable of being retrieved or reproduced in, perceivable form. A record may be in writing or in any electronic or other media."[125] As the Working Group explains:

> The term "record" is new. Throughout the Code, numerous provisions of the various Articles have required parties to communicate in "writing" as defined in Section 1-201(46). Given the rapid development of electronic and other communication and storage technologies, the requirement that documents or com-

---

122. U.C.C. § 1-201(46). Obviously, a tape recording is not a "printing" or "typewriting." It can be transcribed, but then the transcription and not the tape itself is the writing. Moreover, a key feature common to printing and typewriting is that each is readable by the unaided human eye. This feature is not present with respect to tape recordings. *See* 1 RONALD A. ANDERSON, ANDERSON ON THE UNIFORM COMMERCIAL CODE § 1-201:134 (2d ed. 1970).

123. An alternative proposal may be derived from Professor Anderson's argument that a tape recording is not a "writing" because it "lack[s] the element of being 'readable' by the unaided human eye which is characteristic of printing and typewriting." 1 ANDERSON, *supra* note 122, § 1-201:134 This proposal calls for adding "or on any substance" after the words "tangible form" in § 1-201(46). *Id.* Anderson's argument has been criticized by Misner for citing no authority. *See* Misner, *supra* note 106, at 947-48.

124. *See* Patricia B. Fry, *X Marks the Spot: New Technologies Compel New Concepts for Commercial Law*, 26 LOY. L.A. L. REV. 607, 612-16, 622-24 (1993).

125. Letter from Patricia B. Fry, Professor of Law and Associate Dean for Academic Affairs, University of North Dakota School of Law, to Raj Bhala, Assistant Professor of Law, Marshall-Wythe School of Law, 5 (June 27, 1994).

> munications be "written" or "in writing" no longer reflects existing or developing commercial practices. Examples of current technologies commercially used to communicate or store information include, but are not limited to, magnetic media, optical discs, digital voice messaging systems, audio tapes and photographic media.[126]

With the addition of the term "record" to section 1-201, there should be no need to modify the term "writing." The Working Group's proposal is gaining acceptance among other American Bar Association committees charged with studying the revision of the U.C.C. For example, the term "record" may be used in revisions to U.C.C. Articles 2 and 5.[127]

Another legislative option would be to change the statute of frauds to allow parties to vary it by agreement. While freedom of contract is a foundation of Article 2, Citibank and DBS currently are not permitted to opt out of section 2-201(1).[128] Adding a phrase or sentence to section 2-201(1) that would allow them to vary the statute by agreement would provide the necessary freedom.[129]

3. Judicial Re-interpretation

A different way of accomplishing the same task would be for judges to reinterpret sections 1-201(39) and (46) to encompass tape recordings. While this may be a time-consuming and uneven process, there is a textual basis for such judicial action. U.C.C. sections 1-201(39) and (46) contain the word "includes." Because this word does not limit the definition, a court could reasonably extend the definition to include non-paradigmatic writings.

Unfortunately, some courts adhere to the tangibility paradigm. These courts take a formalistic view of tape recordings and, consequently, decide these non-paradigmatic writings do not satisfy the statute of frauds. For example, the issue in *Swink & Co. v. Carroll McEntee & McGinley, Inc.* was whether a tape-recorded oral contract for the sale of securities is enforcea-

---

126. *Id.*

127. *Id.* at 2.

128. *See* U.C.C. § 1-102 cmt. 2 ("[T]he statute of frauds found in Section 2-201 . . . does not explicitly preclude oral waiver of the requirement of a writing, but a fair reading denies enforcement to such a waiver as part of the 'contract' made unenforceable . . . ."); U.C.C. § 1-205 cmt. 4 (referring to the Article 2 statute of frauds as a "mandatory rule . . . whose very office is to control and restrict the actions of the parties, and which cannot be abrogated by agreement, or by a usage of trade").

129. The word "sign" (or "signature," "signatures," "signed," or "signer") appears in U.C.C. §§ 2-205 (the merchant's firm offer rule) and 2-209 (concerning modification, rescission, and waiver). The word "writing" (or "written") appears in U.C.C. §§ 2-202 (the parol evidence rule), 2-203 (providing that seals are inoperative), 2-205, 2-207 (concerning additional terms in an acceptance or confirmation), 2-209, 2-316 (regarding the exclusion or modification of warranties), 2-509 (concerning risk of loss in the absence of a breach of contract), 2-605 (regarding a waiver of a buyer's objections to a delivery of goods), 2-607 (relating to notice of a claim of litigation), and 2-616 (relating to notice of a material or indefinite delay). Modifying the statute of frauds as suggested above, or re-interpreting it as discussed below, is unlikely to upset these other provisions of Article 2. Foreign exchange contracts are between two parties and neither multiple writings nor signatures are involved. Thus, the rights and obligations of third parties would be unaffected by the proposals.

ble under the statute of frauds in Article 8 of the U.C.C.[130] The court assumed that a tape recording is an "intentional reduction to tangible form" and, therefore, a "writing" under section 1-201(46).[131] Nevertheless, the court held that the recording failed to satisfy the statute of frauds because it was not "signed" by the party against whom enforcement of the contract was sought.[132]

The court in *Roos v. Aloi* also refused to re-interpret the statutory language to reflect market practice.[133] Again, the issue was the enforceability of an oral agreement for the sale of stock.[134] The buyer argued the tape recorded conversation of the agreement satisfied the relevant statute of frauds.[135] The court rejected this argument in a highly ironic manner.[136] On one hand, the court noted that the parties (two equal shareholders in a closely-held corporation) were well known to each other. Consequently, strict adherence to the intricacies and formalities of corporate law was unnecessary.[137] On the other hand, the court slavishly followed a single precedent on the statute of frauds, even though the relationship of the parties and the surrounding facts and circumstances reliably indicated the existence of an enforceable contract.[138]

The facts of the *Swink* and *Roos* cases did not suggest that the tape-recorded voices were not genuine. Perhaps the courts' decisions were based on concerns that a voice on tape is more difficult to authenticate than a handwritten signature. For such courts, reliability is the core of the statute of frauds "signature" requirement. Obviously, contracts should not be enforced if it is uncertain whether the defending party assumed contractual obligations; but it is folly to think that authenticating recorded

---

130. 584 S.W.2d 393, 394-96 (Ark. 1979). U.C.C. Article 8, which governs transactions in investment securities, contains a statute of frauds in § 8-319:

> [a] contract for the sale of securities is not enforceable by way of action or defense unless:
>
> (a) there is some writing signed by the party against whom enforcement is sought or by his authorized agent or broker, sufficient to indicate that a contract has been made for sale of a stated quantity of described securities at a defined or stated price.

U.C.C. § 8-319. The fact that the statute of frauds at issue is U.C.C. § 8-319 does not render a case inapposite to the analysis of the Citibank-DBS dispute under U.C.C. § 2-201. The critical terms "writing" and "signed" are used in both sections and defined in exactly the same way in U.C.C. §§ 1-201(46) and 1-201(39), respectively. Moreover, U.C.C. § 8-319 cmt. 1 states that "[t]his Section is intended to conform the statute of frauds provisions with regard to securities to the policy of the like provisions in Article 2 (Section 2-201)." Indeed, examining cases arising under U.C.C. § 8-319 for clues about the interpretation of § 2-201 is commonly done. *See, e.g.*, Southmark Corp. v. Life Investors, Inc., 851 F.2d 763, 767 n.6 (5th Cir. 1988).

131. *Swink & Co.*, 584 S.W.2d at 398-99.

132. *Id.*

133. 487 N.Y.S.2d 637 (N.Y. App. Div. 1985).

134. *Id.*

135. *Id.* at 640. Those provisions were U.C.C. § 8-319 and New York's General Obligation Law. *Id.* at 642-43; *see also infra* note 130.

136. *Roos*, 487 N.Y.S.2d at 642. *See supra* note 108 for a discussion of the General Obligations Law.

137. *Roos*, 487 N.Y.S.2d at 640.

138. *Id.* at 642-43.

voices is any more difficult than authenticating a signature on paper.[139] Indeed, impersonating a voice is probably far harder than forging a signature; hence, a tape recording may well be more reliable than a signature. Further, the vast majority of foreign exchange transactions, like the agreement in the *Roos* case, involve repeat players whose voices are well-known to each other. Certainly the parties introduce themselves to each other before negotiating a spot foreign exchange deal. Courts that consistently apply the tangibility paradigm, such as the *Swink* and *Roos* courts, never entertain these considerations.

The court in the Citibank-DBS dispute should avoid this paradigm's inherently formalistic approach. Instead, it should follow the opinion rendered in *Ellis Canning Co. v. Bernstein* and hold that the tape-recorded conversations between Citibank and DBS satisfy section 2-201.[140] The facts of *Bernstein* closely resemble the disputed dollar-yen transaction. Bernstein orally agreed to sell the stock in United Packers, a company he owned and operated, to the Ellis Canning Company. Bernstein and the Ellis Canning representatives agreed to record the essential elements of the transaction.[141] After the taped conversation, letters and draft agreements were exchanged between Bernstein and Ellis Canning, though the parties' agreement was not memorialized in a signed writing.[142] Subsequently, a third party offered to buy the United Packers stock at a higher price than Ellis Canning had offered. Bernstein reneged on the deal with Ellis Canning and sold the stock to the third party.[143]

The court rightly rejected Bernstein's affirmative defense that the contract with Ellis Canning was unenforceable under the statute of frauds.[144] It held that the tape recording satisfied the statute of frauds writing requirement because the parties had previously agreed in writing to be bound by the tape recording.[145] The court reasoned that if the parties agreed to bind themselves to a recording of their agreement, then

---

139. *See, e.g.*, Misner, *supra* note 106, at 956-63 (discussing the scientific aspects of voiceprints).

140. 348 F. Supp. 1212 (D. Colo. 1972). Professor Misner wrongly rejected the *Bernstein* decision, discussed below, as too liberal and, therefore, unlikely to be followed in different jurisdictions in a uniform manner. Misner, *supra* note 106, at 949-50, 964. A court decision in one jurisdiction that is squarely consistent with the needs of the market and the technological and business culture of that market, and that makes rational economic sense, should be an attractive precedent for courts in other jurisdictions. Of course, possibly the best way to ensure that the *Bernstein* holding is uniformly adopted is to enact it legislatively. Yet, Misner rejected legislative solutions. *Id.* at 945-46.

141. *Bernstein*, 348 F. Supp. at 1215-17.

142. *Id.* at 1217-20. The letters and drafts are analogous to the written confirmations discussed in Case Two below. *See infra* notes 167-215 and accompanying text.

143. *Bernstein*, 348 F. Supp. at 1220.

144. Bernstein unsuccessfully argued that no contract had been formed with Ellis Canning because no meeting of the minds occurred. The court held that all the essential terms of the stock sale were agreed to notwithstanding certain objections to the structure of the transaction raised by Bernstein. *Id.* at 1221, 1225-28.

145. *Id.* at 1228. While the court characterized this decision as a holding, a more conservative interpretation of the case is possible. The court plainly held that the letters and draft contracts exchanged by the parties satisfied the writing requirement of U.C.C. § 8-319. The court stated that it went "a step farther" in deciding that the tape recording, agreed to by the parties, satisfied the statute of frauds. *Id.* This step could be regarded as dicta.

"the contract [was] 'reduced to tangible form' when it [was] placed on the tape."[146] Similarly, the signature requirement was met because its "clear purpose . . . is to require identification of the contracting party and . . . the identity of the oral contractors [was already] established" on tape.[147] Undoubtedly, the court's outrage at Bernstein's behavior compelled this holding. After the taped conversation was concluded, Ellis Canning had provided working capital and management expertise to United Packing. As a result, the financial performance of United Packing markedly improved.[148] Bernstein took advantage of this improvement by selling his stock in United Packing to a higher bidder.

In the hypothetical Citibank-DBS dispute, the court benefits from two additional recent precedents that break away from the tangibility paradigm. In *Londono v. City of Gainesville*, the court held that a tape recording of the city commissioner's action at a public meeting satisfied the signature requirement of the statute of frauds.[149] In *Color & Design Exchange Inc. v. Standish*, the court held that an oral statement made on the record in open court satisfies the statute of frauds.[150] The *Londono* and *Standish* decisions reflect a well-placed confidence in the reliability of tape recordings as evidence of a contract that is lacking in the *Roos* and *Swink* opinions.[151]

4. Certainty

By discarding the tangibility paradigm and enforcing the dollar-yen spot transaction, a court adjudicating the Citibank-DBS dispute ensures that the statute of frauds serves the needs of market participants. One such need is certainty as to the enforceability of spot deals. Knowing that existing obligations will be enforced, sellers of foreign exchange, like DBS, can comfortably reject new offers to buy currency. This certainty is particularly important in the volatile foreign exchange market where new offers are the inevitable result of dramatic, swift changes in rates.

---

146. *Id.*

147. *Id. But see* Swink & Co. v. Carroll McEntee & McGinley, Inc., 584 S.W.2d 393 (1979) (holding that a tape recording was a "writing," but the "signature" requirement was not met; hence the contract at issue was unenforceable).

148. *Bernstein*, 348 F. Supp. at 1218.

149. 768 F.2d 1223, 1227-28 n.4 (11th Cir. 1985).

150. 593 A.2d 169, 170 (Conn. Super. Ct. 1991). The statement was not made by the party to be charged, Standish, but by the attorney for the plaintiff seeking enforcement of the contract. The contract involved a personal guarantee made by Standish, the president of a corporation, for payment of a judgment rendered against his corporation.

151. The Court adjudicating the Citibank-DBS dispute also has the benefit of several precedents establishing a liberal approach to the signature requirement. *See, e.g.*, Barber & Ross Co. v. Lifetime Doors, Inc., 810 F.2d 1276 (4th Cir. 1987), *cert. denied*, 484 U.S. 823 (1987) (the trademark of a seller satisfied the signature requirement); Procyon Corp. v. Components Direct, Inc., 249 Cal. Rptr. 813 (Cal. Ct. App. 1988) (a buyer signed a letter of credit essentially by adopting the signature of the bank that issued the credit); Paloukos v. Intermountain Chevrolet Co., 588 P.2d 939 (Idaho 1978) (the business name of the seller printed on a worksheet satisfied the signature requirement); Automotive Spares Corp. v. Archer Bearings Co., 382 F. Supp. 513 (N.D. Ill. 1974) (letterhead satisfied the signature requirement); A&G Constr. Co. v. Reid Bros. Logging Co., 547 P.2d 1207 (Alaska 1976) (typed name was sufficient to satisfy the requirement).

5. Fraud Prevention or Fraud Promotion?

A decision in the Citibank-DBS dispute that is consistent with the *Bernstein, Londono,* and *Standish* cases serves a second important market need—promoting market integrity and preventing fraud. Like Bernstein, Citibank is asking a court to ratify its unscrupulous behavior through the statute of frauds. Notions of fairness and justice dictate that the tape-recorded telephone transaction between Citibank and DBS should be an enforceable contract.[152] Applying the statute of frauds in this case is at best a hinderance to achieving a just result. It gives a party like Citibank a legal basis for welshing.[153]

On the other hand, the statute of frauds was intended to prevent fraudulent claims by thwarting perjured testimony in contract cases and by generally avoiding "the maladies of fraud and deceit."[154] As Professor Llewellyn stated:

> The effort of the Code (2-201) has been to deal with the essential purposes for which the Statute was designed, while getting and keeping away from the abuses: to wit, to make utterly essential some evidence in writing and over signature, or else some pretty good other evidence that rests on something more tangible than words of mouth. . . . [T]he Code adds both the desire and a reasonable machinery for a businessman to be able to rely on what both parties sign and on the fact that he has procured a memo signed by the other party.[155]

Thus, one could argue that circumscribing or abolishing the statute of frauds as an affirmative defense would constrain the ability of a court to render just opinions in egregious cases. Carefully applying the fundamental rules of contract formation, however, is a protective substitute. Suppose that after the DBS trader provides the Citibank trader with a dollar-yen quote, the Citibank trader says she is "strongly interested" in buying 120 million yen and provides delivery instructions to the DBS trader. On the value date, DBS delivers 120 million yen to Citibank, but no reciprocal payment of $1,153,846.15 is made by Citibank. Assume that after the trade date, the yen depreciated relative to the dollar. Consequently, DBS is anxious to sell its yen at the higher exchange rate that prevailed on the trade

---

152. This position assumes that the tape recording meets the two basic requirements set forth in U.C.C. § 2-201 cmt. 1. First, there is "a basis for believing that . . . a real transaction" occurred. Second, the quantity term is set forth.

153. A contract should be enforced, even if the court must resort to the doctrine of promissory estoppel. *See* Monetti, S.P.A. v. Anchor Hocking Corp., 931 F.2d 1178, 1185-86 (7th Cir. 1991) (discussing whether promissory estoppel can be used to avoid limitations on the enforcement of oral promises placed by the statute of frauds); Southmark Corp. v. Life Investors, Inc., 851 F.2d 763, 771 (5th Cir. 1988) (stating that promissory estoppel should be invoked only when it would be inequitable for the court to apply the statute of frauds).

154. Misner, *supra* note 106, at 943; *see also* DF Activities Corp. v. Brown, 851 F.2d 920, 922, 924 (7th Cir. 1988) (Posner, J., plurality and Flaum, J., dissenting respectively); Holley Equip. Co. v. Credit Alliance Corp., 821 F.2d 1531, 1534 (11th Cir. 1987) (citing Campbell v. Campbell, 371 So.2d 55, 60 (Ala. Civ. App. 1979)); Levin v. Knight, 780 F.2d 786, 787 (9th Cir. 1986).

155. 2 WILLIAM D. HAWKLAND, UNIFORM COMMERCIAL CODE SERIES § 2-201:01, at 9-10 (1992).

date than at the currently-prevailing rate.[156] Therefore, DBS argues that a dollar-yen contract exists. The statute of frauds allows a court to avoid enforcing the alleged contract.

Yet, a court need not apply the statute of frauds to achieve this result. A court can reason that the offer-acceptance process was not completed. The statement of "interest," even when coupled with the provision of delivery instructions, is not an acceptance of DBS's offer to sell yen. Ostensibly, such reasoning is a doctrinal sleight-of-hand, that is, using contract formation rules to deal with enforceability problems. Any theoretical distinction between formation and enforceability, however, is of no practical moment to the Citibank trader. Whether a court decides that no valid contract ever was formed or that a contract exists but is unenforceable, the result is the same—Citibank has no obligation to buy 120 million yen.[157]

As the Citibank-DBS example illustrates, the statute of frauds can defeat its own purpose.[158] In effect, the statute of frauds is lop-sided. It is designed in part to prevent the fraudulent assertion of contractual claims, yet it allows a party to renege on a deal when an agreed price subsequently becomes unprofitable because of subsequent market developments. English law reformers have acknowledged this lop-sidedness. As Justice Wilmot's opinion (quoted at the outset of this Part) suggests, English jurists had complained for many years that the statute promotes fraud. The English Law Reform Committee finally agreed:

> "The Act," [the Statute of Frauds] in the words of Lord Campbell . . . "promotes more fraud than it prevents." True, it shuts out perjury; but it also and more frequently shuts out the truth. It strikes unpartially at the perjurer and at the honest man who has omitted a precaution, sealing the lips of both. Mr. Justice Fitzjames Stephen . . . went so far as to assert that "in the vast majority of cases its operation is simply to enable a man to break a promise with impunity, because he did not write it down with sufficient formality."[159]

In 1954, almost three centuries after it first enacted the statute of frauds, Parliament repealed the statute.[160]

---

156. DBS might cite U.C.C. § 2-204 in its favor, claiming sufficient appropriate conduct thereunder by Citibank to establish an agreement.

157. The statute of frauds is unnecessary to prevent injustices in other types of cases. Suppose a wrongdoer claiming she is a Citibank trader telephones a DBS trader and asks for a dollar-yen quote. The DBS trader provides the quote, and the wrongdoer says "mine, 120 million yen," thereby indicating a purchase of yen for dollars from DBS. The statute of frauds is a means for a court to ensure that the alleged contract is not enforced against Citibank. Citibank can argue, however, that under applicable agency law principles the wrongdoer lacked authority to bind Citibank to an enforceable agreement.

158. *See supra* note 31 and accompanying text.

159. UNITED KINGDOM LAW REVISION COMMITTEE, SIXTH INTERIM REPORT, CMD. NO. 5449, 9 (1937), *quoted in* Misner, *supra* note 106, at 942-43; *see also* E. ALLAN FARNSWORTH, CONTRACTS § 6.1, at 371-72 (1982).

160. FARNSWORTH, *supra* note 159, at 370-71; *see also* C. Grunfeld *Law Reform (Enforcement of Contracts) Act, 1954*, 17 MOD. L. REV. 451 (1954). For discussions of the history of the English statute of frauds pertaining to the sale of goods, see Thomson Printing Machinery Co. v. B.F. Goodrich Co., 714 F.2d 744 (7th Cir. 1983); Hugh E. Willis, *The Statute of Frauds—A Legal Anachronism*, 3 Ind. L.J. 427, 429-32 (1928); George P. Costigan, *The Date and Author-*

6. Results-Oriented Jurisprudence

Admittedly, rejecting the tangibility paradigm and applying the pragmatic strategy may amount to results-oriented jurisprudence inconsistent with the aims of the drafters of U.C.C. Article 2. After all, it would be incorrect to say the drafters intended the statute of frauds to be satisfied by a tape recording. The official commentary to section 2-201, as well as related definitional provisions in section 1-201(39) and 1-201(46), indicate that the drafters took a paradigmatic pencil-and-paper approach.[161]

This objection, however, must be answered using two fundamental aims of the drafters. First, the U.C.C. should be interpreted flexibly in order to promote commercial development. Second, the freedom of parties to contract with one another in a manner they find efficient should not be abridged.[162] The proposed legislative modifications and judicial re-interpretations call for Article 2 and the relevant definitions in Article 1 to be changed in the light of technology and culture in the currency bazaar. If legislatures adopt these changes, then the pragmatic justification for the inclusion of foreign exchange transactions in Article 2 is strengthened. Surely the drafters would prefer to see their sales law "work" for this market through some modest, constructive tinkering rather than wholly exempting the market from the law.

Judge Posner—hardly an exponent of judicial activism—provided an illustration of useful tinkering consistent with the drafters' fundamental aims. In *Monetti, S.P.A. v. Anchor Hocking Corporation*,[163] he confronted the issue of whether a memo that precedes the actual formation of a contract constitutes a writing which satisfies section 2-201. Posner overcame the perfect tense contained in the statutory language, which says the writing must be sufficient to show that a contract "has been" made. The plain meaning is obvious: contract first, writing second. Nevertheless, Posner held that a pre-contractual writing that indicates acceptance of all the essential terms of an offer satisfies section 2-201(1).[164] Posner reasoned that a rule of strict temporal priority is unnecessary where one party unilaterally performs its obligations under the alleged contract. The plaintiff who sought to enforce the contract had conveyed all of its inventory, records, and other assets to the defendant who invoked the statute of frauds. This unilateral performance is unthinkable unless a contract exists.

---

ship of the Statute of Frauds, 26 HARV. L. REV. 329 (1913); Crawford Henning, *The Orignal Drafts of the Statute of Frauds and Their Authors*, 61 U. PA. L. REV. 283 (1913); Justice Stephen & Frederick Pollack, *Section Seventeen of the Statute of Frauds*, 1 LAW. Q. REV. 1 (1885).

161. U.C.C. §§ 2-201, 1-201(39), 1-201(46). U.C.C. § 2-201 cmt. 1 states:

> The required writing need not contain all the material terms of the contract and such material terms as are stated need not be precisely stated. All that is required is that the writing afford a basis for believing that the offered oral evidence rests on a real transaction. It may be written in lead pencil on a scratch pad.

This comment also indicates the drafters' aim to minimize the number of terms a "writing" must contain. *See also* Bazak Int'l Corp. v. Mast Industries, Inc., 538 N.Y.S.2d 503, 508 (1989); 1 STATE OF NEW YORK, REPORT OF THE LAW REVISION COMM'N FOR 1954, at 117-18 (1954) (memorandum by K.N. Llewellyn).

162. *See* U.C.C. § 1-102(3) & cmt. 2; 1 HAWKLAND, *supra* note 155, § 1-102:12 (1984).

163. 931 F.2d 1178, 1182 (7th Cir. 1991).

164. *Id.* at 1182, 1185.

Posner's flexible approach to the language of section 2-201 is applicable to a tape-recorded spot foreign-exchange trade that lacks any evidentiary writing. The partial performance exception to the statute of frauds, set forth in U.C.C. section 2-201(3)(c), reinforces such an approach.[165] That exception states that a contract that fails to satisfy section 2-201(1) "but which is valid in other respects is enforceable . . . with respect to goods for which payment has been made and accepted or which have been received and accepted."[166] Suppose DBS delivers 120 million yen to Citibank on the value date whereas Citibank fails to deliver $1,153,846.15. Without a contract, DBS would not deliver the yen, just as the plaintiff in *Monetti* would not have turned over its entire business to the defendant without a pre-existing (albeit oral) contract. Accordingly, a formalistic interpretation of the writing requirement in the statute of frauds would be unwarranted. The only reasonable inference from the facts is that a contract exists and should be enforced. In sum, Posner-like tinkering is justifiable in contexts where the facts, and the basic aims of the drafters, demand enforcement of an oral contract.

C. *Case Two: The Costs and Benefits of Confirmations*

1. The Tangibility Paradigm Again

Case Two is a paradigmatic situation envisioned by section 2-201(2): a deal made orally, evidenced by a subsequent confirmation slip.[167] Because the confirmations are unsigned or transmitted electronically, they do not comply with the requirements of U.C.C. section 2-201(1). However, the merchant's exception of section 2-201(2) provides that:

> [b]etween merchants if within a reasonable time a writing in confirmation of the contract and sufficient against the sender is received and the party receiving it has reason to know its contents, it satisfies the requirements of subsection (1) against such party unless written notice of objection to its contents is given within 10 days after it is received.[168]

Strictly speaking, this provision is not an "exception" to the statute of frauds but rather "an alternate method of satisfying the writing requirement" of section 2-201(1) that is available for merchants.[169]

---

165. Curiously, Posner does not discuss this exception in detail in the *Monetti* opinion.

166. U.C.C. § 2-201(3)(c). A three-pronged definition of "acceptance" is set forth in U.C.C. § 2-606.

167. *Monetti*, 931 F.2d 1178; *see also* Mid-South Packers, Inc. v. Shoney's Inc., 761 F.2d 1117 (5th Cir. 1985) (involving an oral offer to sell followed by a written invoice). The nonparadigmatic situation is the reverse: a writing is prepared before the actual formation of the contract. *See supra* notes 163-66 and accompanying text.

168. U.C.C. § 2-201(2).

169. Migerobe, Inc. v. Certina USA, Inc., 924 F.2d 1330, 1334 (5th Cir. 1991). Accordingly, some commentators provide a misleading explanation of U.C.C. § 2-201(2). *See, e.g.*, Scarborough, *supra* note 108, at 20.

> A "merchant" is defined in U.C.C. § 2-104(1) as:
> a person who deals in goods of the kind or otherwise by his occupation holds himself out as having knowledge or skill peculiar to the practices or goods involved in the transaction or to whom such knowledge or skill may be attributed by his employ-

Accordingly, one way of satisfying the statute of frauds is to require a written confirmation between parties like Citibank and DBS.[170] Case Two, then, involves the exchange of written confirmations between dealing banks after the banks' traders conclude their recorded telephonic communication. The confirmations identify the currencies involved, the buyer and seller of the currencies, the exchange rate between the currencies, the amount of currencies to be delivered, and the value date. A cost-benefit analysis of the exchange of confirmations pursuant to the statute of frauds indicates that requiring this exchange does not serve the needs of the market.

Llewellyn envisioned parties like Citibank and DBS making contracts by telephone:

> These days we are making contracts over the long-distance telephone as an increasingly standard practice. Decent businessmen having made a contract over the long-distance telephone confirm before five o'clock or close of business that day. As the statute now stands, any crook who wishes to play it both ways against the middle has only to fail to communicate [i.e., to answer the counterparty] and the other guy is stuck. He can hold him or get out according to the market.
>
> This happy opportunity for fraud is unfortunately being indulged in to a considerable extent.
>
> We think that the machinery provided in the section [section 2-201(2)], not by any means wholly satisfactory, at least is a safeguard against this particular type of abuse and fits the practice of constantly closing deals at a distance, and orally.[171]

---

ment of an agent or broker or other intermediary who by his occupation holds himself out as having such knowledge or skill.

Surely, Citibank and DBS are merchants because they deal in currency, satisfy the "knowledge or skill" test, or meet the attribution test.

There is no bright-line test for what constitutes a "reasonable time." It "depends on the nature, purpose and circumstances" of the action that is required. U.C.C. § 1-204(2). Because of the short-term volatility of exchange rates, a "reasonable time" may be a shorter period in the context of the currency bazaar than in other markets. *See, e.g.*, Lish v. Compton, 547 P.2d 223, 227 (Utah 1976) (holding that twelve days was not a "reasonable time" with respect to the wheat market in which prices fluctuated rapidly).

The correct approach to U.C.C. § 2-201(2) appears to be that it is an exception to the signature requirement, not the writing requirement. U.C.C. § 2-201(2) states that a writing must be "sufficient against the sender." This phrase implies that while a written confirmation is required, the signature of the recipient on the confirmation is not needed. To enforce a contract, the sender of a confirmation must produce that confirmation and must have signed it; but the sender need not also show that the recipient signed the confirmation. In other words, to deprive the confirmation recipient of the statute of frauds defense, the confirmation need only be signed by the sender and indicate the existence of a contract. If the recipient receives the confirmation and does not make a timely objection to it, the recipient loses the statute of frauds defense. Therefore, U.C.C. § 2-201(2) excuses the need for the recipient of a confirmation to sign the confirmation.

170. *See, e.g.*, Lambert Corp. v. Evans, 575 F.2d 132 (7th Cir. 1978) (holding that written confirmation of an oral telephone contract satisfied § 2-201). *See generally* FARNSWORTH, *supra* note 159, § 6.7, at 405 (1982) (stating that "the usual way to satisfy the statute [of frauds] is still by a signed writing, commonly called a 'memorandum' ").

171. 1 STATE OF NEW YORK, REPORT OF THE LAW REVISION COMM'N FOR 1954 179.

But, like Misner, Llewellyn could not overcome the tangibility paradigm. In fact, Llewellyn successfully advocated the adoption of section 2-201(2). Yet, contrary to Llewellyn's view, there is nothing indecent about concluding a deal by telephone without exchanging written confirmations.

Llewellyn and Misner are not alone in defending the tangibility paradigm. The leading advocate of the use of written confirmations, as well as the exclusion of foreign exchange transactions from U.C.C. Article 2, is the Federal Reserve. Acting through the Foreign Exchange Committee ("FEC")—an informal advisory group of roughly thirty U.S. and foreign commercial and investment banks and foreign exchange brokers—the Federal Reserve repeatedly encourages market participants to exchange written confirmations.[172] The FEC "believes that the practice of confirming trades by personnel other than traders is the best protection against misdirected trades, payments problems, and other potentially costly mistakes as well as a deterrent to unauthorized dealing."[173]

In spite of the doubt cast below on written confirmations, the FEC continues to advocate the exchange of confirmations.[174] A possible explanation for this intransigence is the regulatory influence of the Federal Reserve on the FEC. The Federal Reserve is responsible for supervising many of the commercial banks (and their holding companies) that participate

---

172. *See, e.g.*, FEC 1992, *supra* note 14, at 9; FOREIGN EXCHANGE COMMITTEE, 1990 ANNUAL REPORT 5 (1991) [hereinafter FEC 1990]; FOREIGN EXCHANGE COMMITTEE, 1989 ANNUAL REPORT 9 (1990) [hereinafter FEC 1989].

173. FEC 1990, *supra* note 172, at 5. Accordingly, the IFEMA, which was drafted by the Financial Market Lawyers Group of the FEC, states that foreign exchange transactions governed by the IFEMA "shall be promptly confirmed by the Parties by Confirmations exchanged by mail, telex, facsimile or other electronic means." IFEMA, *supra* note 14, § 2.3. Section 8.15 of the IFEMA, allows parties to agree on a specific timing for the exchange, checking, and challenge of confirmations. Absent manifest error, confirmations are deemed correct three business days after receipt by a party. An example of manifest error would be where there is a conflict between the confirmation and a tape recording of the conversation between traders. *See* IFEMA GUIDE, *supra* note 14, § III.C, at 7. Under § 8.3 of the IFEMA, a tape recording is the preferred evidence of the terms of a transaction. The definition of "Confirmation" in § 1 of the IFEMA lists the elements that should be included in the document. But, no sample confirmation form is appended to the IFEMA because no single format is accepted in the foreign exchange market as a standard. IFEMA GUIDE, *supra* note 14, § III.C, at 8. Breach of the obligation to send a confirmation, however, carries no penalty; failure to exchange confirmations "shall not prejudice or invalidate" any foreign exchange transaction. *Id.*

174. *See infra* notes 176-216 and accompanying text. As the FEC recently reaffirmed:

> Nevertheless, the [Foreign Exchange] Committee felt strongly that *written* confirmations were still necessary and that tapes did not provide a sufficiently secure and continuous alternative record. . . . [I]n [a] . . . letter to foreign exchange market participants responding to the CIB proposal . . ., the Committee emphasized that it is as necessary as ever to have timely, written confirmations for all spot deals with banks and other dealers.

FEC 1990 *supra* note 172, at 5-6.

The only modification to this position concerns the means of transmitting confirmations. The Federal Reserve now acknowledges that transmitting confirmations in a timely manner, namely, electronically, by telex, or fax, is preferable to sending them in the mail. FEC 1990, *supra* note 172, at 6, 29; *see also* FEC 1992, *supra* note 14, at 9. Electronic transmission occurs through one of two linkages among trading banks: the Society for Worldwide Interbank Financial Telecommunications (SWIFT) system or a direct-dealing system. While such transmissions would entail a transaction cost, at least they would be available before the value date of a spot transaction.

in the foreign exchange market.[175] The Federal Reserve may suspect that the foreign exchange market is plagued by questionable and possibly illegal trading practices that threaten the safety and soundness of the participants it supervises. Written confirmations provide the Federal Reserve with an "audit trail," that is, potential evidence of improper practices. Determined wrongdoers, however, will not hesitate to falsify records. Hence, the practical value of written confirmations for Federal regulators and law enforcement agencies is limited. Tape recordings of traders' and brokers' conversations are themselves audit trails.

2. Delays and Costs

Exchanging confirmations is by no means a universal practice in the foreign exchange market. Many market participants find it time-consuming and costly. Traders seek to conclude their transactions quickly. It is infeasible to require traders to spend much time confirming their trades. After concluding one deal over the phone, their attention turns immediately to the next deal. Not surprisingly, the task of confirming—if it is performed—is left to the trading bank's operations department. The main point, however, is that the merchants exception provides little help to traders. In effect, the statute of frauds "interfere[s] with expeditious contracting by delaying mutual obligation from legally attaching until some later time."[176]

In addition to these delays, exchanging confirmations entails preparatory, transmission, and storage costs. Preparatory costs are those connected with the preparation of the confirmation. The officials in the operations department of Citibank and DBS must ascertain the terms of the trade (e.g., the currencies involved, exchange rates, value date, and delivery instructions) by listening to the tape recorded conversations of the traders, talking with the traders, and checking any written records like trade tickets. The officials must be paid for their time and effort. Sending the prepared confirmation via mail, telex, or fax entails a transmission cost. While this cost may be small for a single confirmation, the fact that Citibank and DBS enter into hundreds of deals every day means that the cumulative transmission cost could be significant. Finally, cautious market participants may seek to store confirmations for the statute of limitations period. Under U.C.C. section 2-725(1), the statute of limitations for an action involving a contract for sale is four years from the date the claim accrues. Either the writings must be stored in a warehouse, leading to inventory and property costs, or converted to microfiche, resulting in storage costs. The sum of preparatory, transmission, and storage costs is a sizeable transaction cost connected with every foreign exchange trade.[177]

175. Under the Bank Holding Company Act, the Federal Reserve has supervisory authority over all bank holding companies. 12 U.S.C. § 1841 (1988). It is also responsible for supervising commercial banks that are state-chartered and members of the Federal Reserve System. 12 U.S.C. § 1813(q) (1988) (defining "appropriate Federal banking agency").

176. Misner, *supra* note 106, at 945.

177. As the FEC admits:

Curiously, in *Monetti*, Judge Posner observed that one purpose of the statute of frauds is "to make the contractual process cheaper and more certain by encouraging the parties to contracts to memorialize their agreement."[178] Posner, however, failed to elaborate on this purpose. It could be that compliance with the statute of frauds leads to less litigation and, thus, greater certainty and lower legal costs.[179] Here, the statute of frauds acts in tandem with the parol evidence rule set forth in U.C.C. section 2-202. By forcing parties to reduce their deal to a writing, the terms thereof will be recorded. Parties will not have to rely on their memories or notes to check the terms. Because such sources are potentially inconsistent, reliance thereon could generate uncertainty. Thus, prudent parties might not only reduce their agreement to writing, but also include a merger clause to ensure their agreement is integrated.[180]

If this logic is what Posner had in mind, then it seems to be undermined by the statute of frauds itself. As discussed below, exactly what a document must contain to be a "writing" and to satisfy section 2-201(1) is unclear.[181] Accordingly, if disputes about the terms of the transaction arise, then surely the parties will rely on the tape recorded conversations of their transaction for guidance.[182] Yet, if the tape is the key evidentiary means for resolving disputed trades, then why bother with a writing in the first place?

> Some banks operate on the assumption that confirmation for a spot trade by a recorded telephone conversation is adequate as long as the contracts settle; they retain written confirmations only for use in the case of a disputed or failed trade. These banks have adopted this procedure in order to *reduce office costs.* They are willing to accept the risk that their more informal confirmation procedures may expose them to a larger number of misdirected spot trades.

FEC 1989, *supra* note 172, at 9 (emphasis added). The FEC's view that reliance on tape recordings exposes banks to greater risk is questioned below.

178. Monetti, S.P.A. v. Anchor Hocking Corp., 931 F.2d at 1181.

179. *See, e.g.*, DF Activities Corp. v. Brown, 851 F.2d at 925 (Flaum, J., dissenting). Indeed, in *Monetti* Posner cites Professor Farnsworth's treatise in support of the proposition that the statute of frauds "is largely based on distrust of the ability of juries to determine the truth of testimony that there was or was not a contract." *Monetti*, 931 F.2d at 1181 (citing FARNSWORTH, *supra* note 159, § 6.1, at 85 (1990)). The inference that Posner draws from this proposition is that it is more costly, and the outcome less certain, to leave this determination to a jury than to memorialize an agreement in writing. An alternative—and not necessarily inconsistent—inference is that a jury is less competent than the parties (or a judge) to make the determination.

180. This clause (also called an integration clause) is designed to prevent inconsistent sources from undermining the integrity of the agreement because of the parol evidence rule. *See generally* 2 FARNSWORTH, *supra* note 98, § 7.3.

181. *See infra* notes 232-35 and accompanying text.

182. Disputes about representations and warranties are unlikely because the participants are likely to be well-known to each other and to have dealt with each other on several previous occasions. A dispute could arise about the designated account to which the yen are to be delivered. Reference to the tape-recorded conversation of the traders might be fruitless in resolving this dispute. Officials of the operations departments of Citibank and DBS, and not the banks' traders, would be responsible for exchanging delivery instructions. Referral to their oral or written communications would be necessary.

3. Practical Irrelevance

There are very few benefits to requiring the exchange of written confirmations that offset the aforementioned delays. The FEC argues that "[c]onfirmations are an important defense against error and fraud."[183] The comments above regarding fraud promotion rebut this argument.[184] Moreover, the alleged benefit assumes that confirmations are, in fact, exchanged promptly. Suppose Citibank sends a written confirmation to DBS on November 1. DBS, which has reason to know of the contents of the confirmation, fails to respond to the confirmation by November 10. The confirmation would satisfy the writing requirement of section 2-201(1).[185] By failing to answer Citibank's written confirmation, DBS loses the defense of the statute of frauds, but Citibank must still prove that an oral contract was made prior to the written confirmation.[186]

As a practical matter, however, this result would be irrelevant. The value date of the spot transaction is November 3, and therefore, the dollar and yen legs should settle on that date. Thus, while confirmations exchanged by mail among foreign exchange market participants satisfy section 2-201(2), they rarely, if ever, arrive in time to identify problems before the value date of a spot transaction.[187] Not surprisingly, in 1989 the U.S. Council on International Banking ("CIB") recommended that banks discontinue the exchange of confirmations by mail because the practice serves no practical purpose.[188]

The only way to detect an error before November 3 would be to check the tape recording of the traders' oral agreement or to exchange confirmations electronically, by telex, or by fax on November 1 or 2. Again, written confirmations transmitted by these swift means are unnecessary where tape recordings of that transaction exist. The CIB's recommendation correctly pointed out that taped telephone conversations are a more efficient method of detecting problems with a trade.[189] They are immediately available for use by the operations departments.

4. Helping the Sophisticated

Given the trusting nature of currency bazaar participants, it is not surprising that many foreign exchange traders find written confirmations unnecessary. After all, repeat players are unlikely to attempt to defraud one another for fear of being ostracized from the marketplace. Citibank's attempt to renege on its agreement with DBS jeopardizes Citibank's own standing in the market. The foreign exchange market has been characterized (in gender-biased terms) as a "gentleman's market" where "a trader's

183. *IFEMA* GUIDE, *supra* note 14, § III.C, at 7.
184. *See supra* notes 152-60 and accompanying text.
185. The same writing standard set forth in U.C.C. § 2-201 cmt. 1 applies to sub-sections (1) and (2) of § 2-201. *See, e.g.*, Bazak Int'l Corp. v. Mast Indus. Inc., 538 N.Y.S.2d 503, 508 (1989).
186. U.C.C. § 2-201 cmt. 3.
187. FEC 1990, *supra* note 172, at 5.
188. *Id.*
189. *Id.*

word is his bond." When the DBS trader responds "120 million yen, yours, at 104," a deal is struck. To renege is to lose credibility.

Overall, market participants police themselves by refusing to deal with those who cannot be trusted to fulfill obligations to which they have committed verbally. Because the market has been generally free from the maladies of fraud and deceit, this self-regulating mechanism has worked. Rigid adherence to section 2-201(2) is, therefore, unnecessary.

This point leads to a reconsideration of one of the purposes of the statute of frauds, namely, precluding the enforceability of contractual claims where a party does not knowingly assume a contractual obligation.[190] The hidden presumption is that parties are unsophisticated and need protection from wandering unwittingly into a legally enforceable obligation. Yet, foreign exchange market participants hardly need such protection. They are large commercial and investment banks, corporations, and investment funds. Not surprisingly, they are acutely aware not only of the risks of trading in the currency bazaar, but also of how trades are negotiated, executed, and consummated.

5. The Tension with the Parol Evidence Rule

Written confirmations are not only irrelevant in practice and unnecessary for the sophisticated market participants of the currency bazaar, but they are also potentially dangerous. Insofar as the contractual terms in them are inconsistent with those stated in the tape recorded telephone conversations, they may generate problems of parol evidence. Suppose the dollar-yen transaction between Citibank and DBS is a thirty-day forward purchase of yen entered into on November 1, i.e., the value date is November 30.[191] The operations department official of Citibank records the terms of the transaction; however, instead of indicating that 120 million yen are purchased, she records that yen are sold. This confirmation is sent to DBS within a reasonable time, and DBS does not send an objection to the confirmation within ten days after it is received. Pursuant to U.C.C. section 2-201(2), the formal requirements of section 2-201(1) are met.

While enforceability is not an issue, what exactly should be enforced is in doubt. On the value date, Citibank demands $1,153,846.15 from DBS, which replies that Citibank is entitled to 120 million yen. Plainly, Citibank's written confirmation is incorrect. It is certain to be inconsistent with the tape recording of the conversation between the Citibank and DBS traders. It also may be inconsistent with Citibank's deal ticket. To ascertain whether Citibank bought, instead of sold, yen, the operations department officials of the two banks must speak with each other, examine their deal tickets, and check the confirmations against the tape recordings and deal tickets.

The point is that the greater the number of sources which evidence the agreement, the greater the probability of inconsistencies in the

190. *See, e.g.*, Scarborough, *supra* note 108, at 22.
191. For a discussion of forward transactions, see *supra* note 17.

sources. Requiring more confirmations of a trade is not necessarily a safeguard against error. To the contrary, it can foster errors, thus exacerbating risk and confusion in the currency bazaar.[192] Had the parties relied on the tape recordings, the mess may have been averted. This scenario is hardly far-fetched. Indeed, its facts resemble those in *Intershoe, Inc. v. Bankers Trust Co.*[193] Moreover, as one foreign exchange market observer stated, "[t]here cannot be any market dealer anywhere who has never done a deal 'the wrong way round', or for the wrong amount, or the wrong value date, or some other major error at some time."[194]

In this Citibank-DBS dispute, the parol evidence rule of section 2-202 must be applied to determine the terms of the transaction.[195] This application yields two principal difficulties. First, the methodology used when applying this rule is uncertain. Second, the rule can produce erroneous results.

With respect to methodology, the Court must decide whether Citibank's written confirmation is the final expression of the parties with respect to the terms of their agreement. Then, the court must decide whether the confirmation constitutes an integrated agreement. Assuming the writing itself does not indicate that it is or is not the complete, conclusive statement of the terms, the Court must make this determination. To be sure, the standard the Court must apply under Article 2 is clearer than

---

192. It can also foster disputes about whether the contents of a document are adequate to allow a court to conclude that the document constitutes a writing for purposes of the statute of frauds. *See, e.g.*, Levin v. Knight, 780 F.2d 786, 790 (9th Cir. 1986) (concerning whether a written memorandum contained enough of the essential terms, with sufficient specificity, to evidence a contract).

193. In *Intershoe*, 569 N.Y.S.2d 333 (N.Y. 1991), a shoe importer, Intershoe, entered into thirty-week dollar-Italian lira forward transactions with Bankers Trust. (The court erroneously referred to these transactions as "futures." Because they took place in the over-the-counter market and not on an organized exchange, they are forwards. *See* Bhala *supra* note 2, at 100). Bankers Trust sent Intershoe a confirmation indicating that it had bought 537,750,000 lira from, and sold $250,000 to, Intershoe. Intershoe signed the confirmation and returned it to Bankers Trust. Just before the delivery date, Intershoe said that it had bought, not sold, lira. Intershoe's attempt to introduce a supporting affidavit was rebuffed by virtue of Section 2-202. The court found the confirmation to be the final expression of the parties' agreement. *Intershoe*, 569 N.Y.S.2d at 335.

Curiously, while the disputed transactions were negotiated and concluded by telephone, the court did not refer to the tape recordings of the parties' conversations. Moreover, the court's dicta that "a confirmation slip or similar writing is usually the only reliable evidence of such transactions," *id.* at 336-37, is factually erroneous.

194. JOHN HEYWOOD, FOREIGN EXCHANGE AND THE CORPORATE TREASURER 109 (2d ed. 1979).

195. U.C.C. § 2-202 states:

> Terms with respect to which the confirmatory memoranda of the parties *agree* or which are otherwise set forth in a writing intended by the parties as a final expression of their agreement with respect to such terms as are included therein *may not be contradicted by evidence of any prior agreement* or of a contemporaneous oral agreement but may be explained or supplemented
>
> (a) by course of dealing or usage of trade . . . or by course of performance; and
>
> (b) by evidence of consistent additional terms unless the court finds the writing to have been intended also as a *complete and exclusive* statement of the terms of the agreement.

(emphasis added).

under the common law.[196] Official comment 3 to section 2-202 states that extrinsic evidence of contractual terms must be excluded "[i]f the additional terms are such that, if agreed upon, they would *certainly* have been included in the document."[197]

Unfortunately, the U.C.C. Article 2 standard may be non-sensical and lead to uncertain results in a dispute such as that between Citibank-DBS where one party has confused the deal. Surely if Citibank had agreed to sell yen, then it would have stated so in the confirmation; consequently, tape recordings must be excluded from consideration. Yet, in fact, selling yen is exactly what Citibank agreed to do, as the tape recordings can prove. Citibank did not state it sold yen because of a clerk's innocent error in making the confirmation, the dishonesty of a trader in writing a deal ticket, or some other reason. Thus, on the one hand, a court might hold the Citibank confirmation is not an integration and, therefore, extrinsic evidence such as the tape recording should be admitted.[198] But, on the other hand, following a strict construction of the test in official comment 3, a court could come to the opposite conclusion. In sum, because application of the parol evidence rule may yield inconsistent results in similar cases, the important goal of providing certainty and predictability to foreign exchange market participants is lost.[199]

---

196. One common law test, proposed by Professor Williston, focuses on whether reasonable parties, situated as were the parties to this contract, would have naturally and normally included the extrinsic matter in the writing. 4 SAMUEL WILLISTON, A TREATISE ON THE LAW OF CONTRACTS §§ 638-39 (3d ed. 1961). Williston's test leads to the result that merger clauses usually are conclusive evidence of the completeness of a writing.

Professor Corbin advocated a two-step inquiry, not an objective "reasonable person" test. The judge should consider extrinsic materials to determine whether there is "respectable" evidence that an antecedent agreement was made. If so, then the judge should determine whether the antecedent agreement was discharged by the subsequent writing. 3 ARTHUR L. CORBIN, CORBIN ON CONTRACTS § 582 (1960); RESTATEMENT (SECOND) OF CONTRACTS §§ 210(3), 214 (1981). Under Corbin's test, a merger clause is only one item of evidence weighed against other facts.

Yet another test, found in Sections 229 and 230 of the RESTATEMENT (FIRST) OF CONTRACTS (1932), asks whether reasonable persons in the parties' situation would have included the disputed provision in the contract.

Not surprisingly, Professor Murray, while discerning a movement toward the Corbin test, concluded that the case law "has been generally ineffective in articulating a workable rationale . . . ." JOHN E. MURRAY, MURRAY ON CONTRACTS § 107, at 235-36 (2d rev. ed. 1974) (footnote omitted).

197. U.C.C. § 2-202 cmt. 3 (emphasis added). This comment is commended as a "consistent starting point" that improves on the common law. *See, e.g.*, Manire, *supra* note 7, at 1204. U.C.C. § 2-202 manifests Llewellyn's approach to the problem of deciding whether a writing is integrated.

198. One ground for this conclusion could be that Citibank's confirmation is "merely to furnish an aid to the writer's recollection." 9 JOHN H. WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 2429, at 96 (Chadbourn rev. 1981). A different ground could be that the confirmation is designed solely to satisfy the statute of frauds under U.C.C. § 2-201(2). Some cases have adopted this approach. *See, e.g.*, Southern States Dev. Co. v. Robinson, 494 S.W.2d 777, 782 (Tenn. Ct. App. 1972) (holding that if a memorandum meets the Statute of Frauds, the entire contract may be explained and proved by parol evidence); Nathan v. Spector, 120 N.Y.S.2d 358, (N.Y. App. Div. 1953) (holding that parol evidence may even be used to determine whether a memorandum meets the Statute of Frauds).

199. An interesting question is whether reformation of the terms stated in the confirmation is possible based on the tape recordings.

Strangely, neither the *Intershoe* nor *IBJ* courts applied the "would certainly" standard. In addition, neither court explained why it rejected this standard.[200] Both courts were inevitably dragged into fact-specific inquiries, but each court emphasized different facts. The *Intershoe* court focused on the terms stated in the confirmation itself.[201] In effect, it applied the four-corners test, looking only at the writing to decide its completeness. The *IBJ* court examined all the evidence of completeness and exclusivity, including evidence beyond the writing. It considered the intention of the parties, the history of their negotiations and relationship, and the omission of a signature from the confirmations.[202]

Not only did each court highlight different facts, but they also rendered diametrically opposed judgments. In *Intershoe* the parol evidence rule barred out extrinsic evidence, while in *IBJ* such evidence was admitted. Consequently, these cases have created considerable uncertainty in the currency bazaar. Participants cannot predict exactly how a court might analyze whether a confirmation is an integrated agreement or the likely result. Suppose the Court determines that Citibank's written confirmation is an integration of the agreement. The second difficulty resulting from the application of section 2-202 is that it leads to an erroneous result. The tape recording cannot be introduced as evidence because it contradicts the confirmation. The result is that incorrect contractual terms are enforced, namely, that Citibank delivers rather than receives yen. DBS was in the best position to correct the confirmation but failed to do so. This factor, however, should not be dispositive.

---

200. One possibility arises from a close reading of U.C.C. § 2-202 cmt. 3, which refers to subsection (b). The comment indicates that the court should focus on whether the parties intended the writing to be a complete and conclusive statement of all the terms. If the parties so intended, then evidence of consistent additional terms must be kept from the trier of fact.

201. 569 N.Y.S.2d 333, 336. The facts of the case are discussed *supra* note 193.

202. *IBJ*, *supra* note 45, at 422-23. In *IBJ*, a Chilean company, Compania Sud-Americana de Vapores (CSAV) received foreign currencies in payment for its shipping services. The currencies were deposited in an account maintained by IBJ Schroder Bank & Trust ("Schroder"), which was responsible for converting the currencies into U.S. dollars and crediting CSAV's account with Schroder. Schroder confirmed each currency conversion transaction with CSAV. The gravamen of CSAV's complaint was that Schroder charged exchange rates that were in excess of spot market rates. *Id.* at 415-16.

These facts are distinguishable from those in *Intershoe* in certain key respects. *Intershoe* involved a written confirmation of a single transaction. The confirmation was an integrated document that reflected the terms to which the parties had agreed over the telephone. Consequently, the parol evidence rule barred the admission of extrinsic evidence to supplement or alter the terms of the transaction. *Intershoe*, 569 N.Y.S.2d at 337. *IBJ* involved a large number of currency conversions, and CSAV did not telephone Schroder to negotiate and conclude these transactions. Rather, the currency conversions were performed by Schroder pursuant to a prior overarching management agreement made with CSAV. The confirmations of each conversion were not intended to reflect this agreement. Hence, the parol evidence rule could not bar extrinsic evidence about its terms. *IBJ*, *supra* note 45, at 432.

The fact that *Intershoe* involved one confirmation whereas *IBJ* involved several is irrelevant for purposes of applying the parol evidence rule. A final expression of an agreement may be manifested in one or more documents. WHITE & SUMMERS, *supra* note 53, § 2-10, at 98. Accordingly, the decision in B.N.E., Swedbank, S.A. v. Banker, 794 F. Supp. 1291, (S.D.N.Y. 1992), *aff'd*, 996 F.2d 301 (2d Cir. 1993), is suspect. The *B.N.E.* court stated that *Intershoe* was irrelevant because it involved a single document whereas the case at hand involved several confirmation slips. *Id.* at 1292.

The effect of the parol evidence rule is "to give preference to the written version of [the] terms [of a contract]."[203] The justification for this preference is that "[w]ritings are more reliable than memories to show contract terms, and forgery is supposedly easier to detect than is lying on the witness stand."[204] Many critics of the rule emphasize that it is inconsistent with conventional processes of proof—juries should be allowed to hear all relevant evidence.[205] The criticism is even more poignant in the context of the currency bazaar where the conversation between the contracting parties is tape recorded. Obviously, a tape recording is the most reliable evidence of the transaction terms. The parol evidence rule compels a court to behave like an ostrich with its head in the sand. Application of the rule may specifically exclude the one form of evidence that can conclusively resolve whether Citibank bought or sold yen.

In sum, there is a tension between section 2-201(2) and the parol evidence rule. The former can be satisfied with a written confirmation to which there is no objection. The confirmation, however, generates a potential problem under section 2-202. It may transform the most reliable evidence of the transaction, the tape recording, into parol evidence. Therefore, the use of tapes to resolve disputes about terms becomes uncertain.

6. The Tension with the Battle of the Forms Rule

A similar tension exists between U.C.C. section 2-201(2) and Article 2's provision on the battle of the forms, section 2-207. Suppose both Citibank and DBS issue and exchange written confirmations on the trade date, November 1, after they reach an oral agreement evidenced by a tape recording.[206] These confirmations, printed on each bank's standard form, conflict. Citibank's confirmation says that it bought 120 million yen for value on November 30, whereas DBS's confirmation says that the value date is November 3. DBS does not notify Citibank of any objection to the terms of Citibank's confirmation, but on November 3 it sends 120 million yen to Citibank and asks Citibank for the reciprocal delivery of $1,153,846.15. Citibank objects, saying the deal involved the forward, not spot, sale of yen.[207]

---

203. WHITE & SUMMERS, *supra* note 53, § 2-9, at 95.

204. *Id.*

205. *Id.* at 95 & n.3.

206. The above hypothetical is a combination of Cases (1) and (6) in WHITE & SUMMERS, *supra* note 53, § 1-3, at 30-36, 43-46. *See also* Douglas G. Baird & George Weisberg, *Rules, Standards, and the Battle of the Forms; A Reassessment of Section 2-207*, 68 VA. L. REV. 1217-19 (1982) (discussing the scenarios to which the battle of the forms refers). For a recent survey of proposed revisions to U.C.C. § 2-207, see *Ending the "Battle of the Forms": A Symposium on the Revision of Section 2-207 of the Uniform Commercial Code*, 49 BUS. LAW. 1019 (1994). Except for the additional fact regarding the tape recording, the hypothetical is one of the two paradigmatic cases that U.C.C. § 2-207 is designed to deal with: "the written confirmation [situation], where an agreement has been reached either orally or by informal correspondence between the parties and is followed by one or both of the parties sending formal memoranda embodying the terms so far as agreed upon and adding terms not discussed." U.C.C. § 2-207 cmt. 1.

207. *See supra* note 17 regarding forward transactions.

Here the problem is the battle of the forms and application of section 2-207. Assuming the confirmations do not constitute an integrated agreement, if the terms of the Citibank and DBS confirmations conflict or if one confirmation omits a term that the other includes, then a court must determine which terms are part of the contract.[208] Applying section 2-207 again highlights the importance of the tape recorded conversation and the disruption caused by the very use of written confirmations.[209]

Because the confirmations state different delivery dates, the threshold question is whether section 2-207(2) covers different as well as additional contractual terms.[210] The answer is uncertain. Even though U.C.C. section 2-207(1) expressly refers to "different" terms, section 2-207(2) does not contain this language. "[T]he drafters could easily have inserted 'or different' if they had so intended."[211] Official comment 3 to U.C.C. section 2-207 and some case law, however, do indicate that different terms are covered.[212]

Assuming section 2-207(2) governs, its application is problematic. One interpretation is that the language "additional term" means "additional or different terms." Interpreted this way, U.C.C. sections 2-207(2)(b) and (c) indicate that an additional term automatically becomes part of the contract unless (i) that term "materially alter[s]" the contract or (ii) the recipient of the confirmation with the additional term objects to it within a "reasonable time" after receiving the confirmation. Because DBS failed to answer Citibank's confirmation "within a reasonable time after additional terms were proposed, it is both fair and commercially

---

208. Note that U.C.C. § 2-202 automatically construes the writings as integrated to the extent that they agree on certain terms. *See, e.g.*, Luria Bros. & Co. v. Pielet Bros. Scrap Iron & Metal,Inc., 600 F.2d 103 (7th Cir. 1979). Note also that U.C.C. § 2-207 has received considerable scholarly criticism. *See, e.g.*, Charles M. Thatcher, *Battle of the Forms: Solution by Revision of Section 2-207*, 6 UCC L.J. 237, 240 (1984) (arguing that § 2-207 has "discouraged the expansion of commercial practices through custom, usage, and agreement of the parties" and yielded "unsettled case law and consequent lack of uniformity . . . ."); W. David Slawson, *The New Meaning of Contract: The Transformation of Contracts Law by Standard Forms*, 46 U. PITT. L. REV. 21, 59 (1984) (arguing that offer and acceptance forms can be drafted to prevent the making of a contract, with the result that the common law mirror-image rule is reinstated); Baird & Weisberg, *supra* note 206 (arguing that § 2-207 is so vague that it leaves a number or questions unresolved).

209. According to the common law mirror image rule, one confirmation (presumably the first one sent) would be treated as an offer, while the other confirmation would be an effective acceptance only if it did not vary the terms of the offer. 1 RESTATEMENT (SECOND) OF CONTRACTS §§ 58-60 (1979); *see also* WHITE & SUMMERS, *supra* note 53, § 1-3, at 29-30. Because such a variation exists in the Citibank-DBS case, one of the confirmations would be treated as a counter-offer. To decide the terms of the contract, the Court would examine evidence of a prior oral agreement, including the tape recordings.

210. This section may be particularly relevant to the dispute because Citibank and DBS are "merchants." *See supra* note 169.

211. WHITE & SUMMERS, *supra* note 53, § 1-3, at 32 (footnote omitted).

212. U.C.C. § 2-207(1) operates automatically to convert a confirmation with different terms as a proposal for an addition to the contract. Official comment 3 indicates that "[w]hether or not additional or *different* terms will become part of the agreement depends on the provisions of subsection (2)." U.C.C. § 2-207 cmt. 3 (emphasis added); *see also* Westinghouse Electric Corp. v. Nielsons, Inc., 647 F. Supp. 896, 900 (D. Colo. 1986); Steiner v. Mobil Oil Corp., 569 P.2d 751 (Cal. 1977) (holding that § 2-207(2) applies to different and additional terms). *But see* WHITE & SUMMERS, *supra* note 53, § 1-3, at 32 nn.11-12.

sound to assume that their inclusion has been assented to."[213] The transaction, regardless of what the Citibank and DBS traders agreed to on tape, becomes a yen forward unless DBS persuades a court that the November 30 delivery date is a material alteration of the contract. Thus, the Citibank-DBS dispute becomes one of materiality of the delivery term.

This result is unsatisfactory. "Materiality" is a fact-specific determination that may involve protracted litigation.[214] While official comment 4 to U.C.C. section 2-207 provides limited guidance on what constitutes a "material alteration," the statute itself is silent on the matter. Thus, there is no guarantee of consistent results in other like cases.

An alternative and preferable approach to the application of section 2-207(2) is to focus on the advice provided in official comment 6:

> Where clauses on confirming forms sent by both parties conflict each party must be *assumed to object* to a clause of the other conflicting with one on the confirmation sent by himself. As a result the requirement that there be notice of objection which is found in subsection (2) is satisfied and the conflicting terms do *not* become part of the contract. The contract then consists of *the terms originally expressly agreed to,* terms on which the confirmations agree, and terms supplied by this Act, including subsection (2). The written confirmation is also subject to Section 2-201. Under that section a failure to respond permits enforcement of a prior oral agreement; under this section a failure to respond permits additional terms to become part of the agreement.[215]

Plainly, the conflicting terms stated in the confirmation are not part of the contract (though this may not be the case under the proposed revisions to Article 2). Rather, a court relies on the tape recordings of the agreement as evidence of the value date that was originally agreed to by the Citibank and DBS foreign exchange traders.

In sum, under the latter approach to its application, section 2-207(2) may be equipped to handle the problem of inconsistent confirmations. Yet, this resolution begs the question of the repercussions of using such confirmations. There is a tension between the statute of frauds (specifically, section 2-201(2)) and section 2-207(2) in the context of the currency bazaar: satisfying the former with written confirmations lays the foundation for battle of the forms problems. Such problems are best avoided by eliminating confirmations and relying on the tape recording for dispositive evidence of the terms of the dollar-yen transaction.

D. *Linking the Resolutions of the Scope and Enforceability Problems*

The pragmatic strategy links the resolution of the enforceability and scope problems. Whether the definition of "goods" in U.C.C. section 2-105(1) ought to encompass foreign exchange should depend in part on

---

213. U.C.C. § 2-207 cmt. 6.

214. *See, e.g.,* Luedtke Eng'g Co., Inc. v. Indiana Limestone Co., 740 F.2d 598, 600 (7th Cir. 1984); St. Charles Cable TV, Inc. v. Eagle Comtronics, Inc., 687 F. Supp. 820, 827 (S.D.N.Y. 1988).

215. U.C.C. § 2-207 cmt. 6 (emphasis added).

whether the application of the statute of frauds serves the needs of foreign exchange market participants. As presently constituted and interpreted, it does not. The statute of frauds would render unenforceable many foreign exchange transactions that should be enforced. The statute may lead to problems of parol evidence and the battle of the forms. Therefore, Llewellyn's argument that "after two centuries and a half the statute stands, in essence better adapted to our needs than when it was first passed"[216] is unpersuasive. At least in the context of the foreign exchange market, the statute of frauds in U.C.C. Article 2 must be reformed or abolished through legislative or judicial action.

Does this argument necessarily dictate that foreign exchange should not be considered a "good"? The argument strongly suggests an affirmative answer. Such an answer, however, would be an overreaction. Thus, the Court should reject DBS's response to Citibank's statute of frauds defense. A final resolution of the scope problem depends on more than the outcome of the enforceability problem. A complete assessment of other significant Article 2 provisions in relation to the needs of the currency bazaar is needed.[217]

V. EXTENDING THE PRAGMATIC STRATEGY TO OTHER SALES LAW

Parts III and IV considered the potential applicability of the U.C.C. to the Citibank-DBS dispute chronicled in Part II. This assumption is now relaxed. The application of three other sales laws to the scope and enforceability problems is considered below: revised U.C.C. Article 2, the CISG, and private sales law. Part V argues that the pragmatic strategy can be extended to deal with these problems under other sales law regimes.

A. *Current Proposals to Revise U.C.C. Article 2*

One proposed revision would decisively resolve the scope problem by excluding all foreign exchange transactions from U.C.C. Article 2. Section 2-102(a)(23) of the December 1993 and August 1994 Drafts of revised Article 2 defines "goods" as "all things . . . that are movable . . . [h]owever, the term does not include . . . *foreign exchange transactions.*"[218] Unfortunately, the Reporter's Notes do not indicate why foreign exchange transactions are expressly excluded. There is no suggestion that the exclusion reflects a calculated decision that the provisions of revised Article 2 would be inapposite to the currency bazaar, though this indeed may be the case.

Interestingly, the December 1993 and August 1994 Drafts also abolish the statute of frauds, thus resolving the enforceability problem. Section 2-201 states that "[a] contract . . . is enforceable whether or not there is a

---

216. Karl N. Llewellyn, *What Price Contract?—An Essay in Perspective*, 40 YALE L.J. 704, 747 (1931).

217. Such provisions include those relating to contract formation and remedies. *See* Bhala, *supra* note 74.

218. U.C.C. § 2-102(a)(23) (Tentative Draft Dec. 21, 1993) (emphasis added). Foreign exchange transactions are not excluded from the definition of "goods" in the September 1993 Draft.

writing signed or record authenticated by a party against whom enforcement is sought . . . ."[219] This change—a complete rejection of the tangibility paradigm—would, of course, meet the needs of foreign exchange market participants.

A more radical proposal regarding Article 2 is to revise it according to a "hub-and-spoke" model.[220] Core rules applicable to all contracts would form the hub from which spokes would emanate. The spokes would set forth principles designed for special transactions. Conceivably, one spoke could apply to foreign exchange transactions. The pragmatic strategy could support this radical approach if major provisions of Article 2, in addition to the statute of frauds, are not applied to the currency bazaar context.

B. *The United Nations Convention on Contracts for the International Sale of Goods ("CISG")*

A literal reading of Articles 1(1) and 2(d) of the CISG makes it impossible to argue that the scope of the CISG encompasses foreign exchange transactions. Article 1(1) states that the CISG applies to "contracts of sale of goods between parties whose places of business are in different States," but Article 2(d) excludes sales of money from the Convention.[221] Unlike U.C.C. section 2-105(1), Article 2(d) does not distinguish between money that is the subject of the contract (the commodity leg) and money that is used as payment (the payment leg). CISG Article 2(d) also fails to mention "things in action."

The crude scope clause in the CISG is unfortunate because the way in which the CISG resolves the enforceability problem serves the needs of the currency bazaar. There is no statute of frauds in the CISG—here again rejecting the tangibility paradigm. Article 11 of the CISG states that "[a] contract of sale need not be concluded in or evidenced by writing."[222] Thus, if the CISG applies to the Citibank-DBS dispute, then Citibank cannot argue that the dollar-yen contract is unenforceable. Moreover, questions about tape recordings as evidence of a contract, costs associated with written confirmations, and potential difficulties arising under Article 19 of

---

219. U.C.C. § 2-201 (Tentative Draft Dec. 21, 1993). The September 1993 Draft contains essentially the same revision of Section 2-201 as the December 1993 Draft.

220. *See* U.C.C. REV. ART. 2 (Discussion Draft Feb. 10, 1994); Nimmer, *supra* note 10.

221. CISG, *supra* note 13, at 672. The States must be Contracting States (i.e., they have ratified or acceded to the CISG), or the applicable choice of law rules must lead to the application of the law of a Contracting State. Pursuant to CISG Article 95, the United States has taken a reservation to the choice of law provision. *See id.* at 693; MARTINDALE-HUBBELL, *supra* note 13, at IC-35; Smart *supra* note 13, at 1344-46. Under Article 6, parties are free to exclude the application of the CISG. CISG, *supra* note 13, at 673.

222. CISG, *supra* note 13, at 674. However, under CISG Article 96, a state can take a reservation to Article 11 if the law of that state provides that contracts must be in writing. *Id.* at 693-94. (Article 96 refers to "legislation" which presumably means that the law must take the form of a statute or civil code.) If a party to a sales contract has its place of business in a reserving state, then CISG Article 11 is inapplicable. *Id.* at 674 (Art. 12). Argentina, Belarus, Chile, China, Hungary, the Russian Federation, and the Ukraine have taken reservations under CISG Article 96 to Article 11. *See* MARTINDALE-HUBBELL, *supra* note 13, at IC-34 to IC-35.

the CISG (the provision on the battle of the forms) are irrelevant. In sum, the CISG's resolution of the enforceability problem strongly suggests that the scope of the CISG should cover the Citibank-DBS dispute.

The literal language of CISG Article 2(d) remains problematic, however, thus requiring a new judicial interpretation of that Article. For example, a court could find that a distinction is implied between commodity leg and payment leg monies. Yen are the subject of the contract and "goods" for purposes of determining the scope of the CISG. This finding is not unprecedented. In *In re Midas Coin Co.*, the court distinguished between those coins sold as a commodity for numismatic purposes and those coins used as a means of payment.[223] The *Midas* court was confronted with the definition of "goods." Like the CISG definition, the court noted without further elaboration that the term " 'includes all things which are movable . . . but does not include money.' "[224]

Here again, a court that renders a *Midas*-type decision under the CISG may be criticized for engaging in judicial activism or subscribing to results-oriented jurisprudence.[225] The global nature of the currency bazaar, however, remains. Many foreign exchange transactions cross international borders.[226] The participants in different countries will benefit from the lack of a statute of frauds and from certainty that the CISG applies to their transactions, instead of wondering which country's contract law governs.

C. *Private Contract Law*

The outstanding example of private contract law in the currency bazaar is the International Foreign Exchange Master Agreement ("IFEMA").[227] Strictly speaking, there is no problem of scope under this law. Scope is a matter for the parties to decide. Parties can freely enter into the master agreement and designate the foreign exchange transactions that they want covered therein. Thus, for instance, Citibank and DBS can sign the IFEMA and indicate that it will govern dollar-yen spot transactions between their respective New York and Singapore offices.[228]

When a dispute arises concerning a transaction governed by the IFEMA, the scope problem resurfaces. The dispute must be resolved not by the IFEMA, but by interpreting the IFEMA under some other sales law regime, whether it be U.C.C. Article 2, revised Article 2, or the CISG.

---

223. 264 F. Supp. 193, 197-98 (E.D. Mo. 1967), *aff'd sub nom.* Zuke v. St. Johns Community Bank, 387 F.2d 118 (8th Cir. 1968). The application of Article 2 to such transactions is not controversial. *See, e.g.*, Morauer v. Deak & Co., 26 U.C.C. Rep. Serv. (Callaghan) 1142 (D.C. Super. Ct. 1979) (applying Article 2 to the sale of gold and silver foreign coins).

224. *Midas*, 264 F. Supp. at 195 (citing Mo. ANN. STAT. § 400.9-105(1)(f) (Vernon 1965).

225. *See supra* notes 160-66 and accompanying text.

226. *See supra* note 18.

227. *See* IFEMA, *supra* note 14.

228. Alternatively, they could specify that the IFEMA governs spot transactions in all currencies between these two offices, spot transactions between multiple offices of Citibank and DBS, or some other category of transactions and offices.

Therefore, the IFEMA does not necessarily afford the parties certainty as to whether their transactions are enforceable.

The utility of the pragmatic approach again becomes evident; whether a particular sales law should govern the disputed transaction should depend on whether that law meets the needs of the parties. The IFEMA expressly states that in the event of a dispute between the parties as to the terms of their transaction, the tape recording is "the preferred evidence" of the terms, "notwithstanding the existence of any writing [namely, the IFEMA] to the contrary."[229] This provision, which states that market participants signing the IFEMA agree in writing to be bound by a tape recording, parallels the criterion from *Ellis Canning Co. v. Bernstein.*[230] Accordingly, if the *Bernstein* holding is adopted by the Court in the Citibank-DBS dispute, then those parties may be afforded greater certainty as to the enforceability of their foreign exchange trades.

Regardless of the IFEMA applicability, the disputed dollar-yen transaction would be enforceable under either revised Article 2 or the CISG because a statute of frauds is absent under both regimes. This result suggests that including foreign exchange transactions governed by the IFEMA within the scope of revised Article 2 or the CISG would support market needs. A different result is reached with respect to U.C.C. Article 2. A court may feel compelled to accept Citibank's defense that the IFEMA does not satisfy the minimal writing requirements of section 2-201(1). It certainly qualifies as a "writing" under section 1-201(46) and is "signed" according to section 1-201(39).[231] Uncertainty remains, however, as to whether the IFEMA must state the quantity of currency purchased, and if so, whether the IFEMA meets that requirement.

Arguably, the statute of frauds does not require that a writing state a quantity term, because there is no reference to such a term in section 2-201. Moreover, official comment 1 to U.C.C. section 2-201 indicates that "[a]ll that is required is that the writing afford a basis for believing that the offered oral evidence rests on a real transaction." A liberal construction suggests that a quantity term is not essential for the establishment of an enforceable contract.[232] If true, then all of the transactions covered by the IFEMA are enforceable.

---

229. IFEMA, *supra* note 14, § 8.3.

230. 348 F. Supp. 1212, 1228 (D. Colo. 1972).

231. An interesting question arises as to whether the IFEMA is nothing more than a manifestation of intent to enter into a contract or is the contract itself.

232. *See* Caroline N. Bruckel, *The Weed and the Web: Section 2-201's Corruption of the U.C.C.'s Substantive Provision — The Quantity Problem*, 1983 U. ILL. L. REV. 811 (1983). Professors White and Summers adopt a similar approach: "a close reading of Section 2-201 indicates that all commentators may be wrong. An alternative interpretation is that only if the writing states a quantity term is that term determinative." WHITE & SUMMERS, *supra* note 53, at 76 n.12; *see also* American Original Corp. v. Legend, Inc., 652 F. Supp. 962, 966 (D.Del. 1986) (finding that the word "all" is sufficient to create an output contract under the statute of frauds and that parol evidence may be admitted to ascertain the exact amount); Riegel Fiber Corp. v. Anderson Gin Co., 512 F.2d 784, 789 n.11 (5th Cir. 1975) (stating that under the White and Summers reading, if a quantity term is present, then it controls, but if no quantity term is contained in the writing, then the party seeking enforcement of the agreement can establish the term by parol evidence).

This construction, however, conflicts with the plain meaning of a different passage from the same official comment: "The only term which must appear is the quantity term which need not be accurately stated but recovery is limited to the amount stated."[233] In the context of the currency bazaar, the "quantity term" refers to the amount of currencies bought or sold. Yet, parties sign the IFEMA before negotiating or concluding any particular spot transaction. The IFEMA cannot predetermine how many yen Citibank purchases from DBS in a particular transaction.[234] Consequently, no master agreement can be effective against Citibank's statute of frauds defense.[235]

VI. CONCLUSION

Private contract law does not resolve the problems of scope and enforceability. It simply leads judges, regulators, and market participants back to U.C.C. Article 2, revised Article 2, or the CISG. With respect to Article 2, absent legislative modification or judicial re-interpretation, the statute of frauds and associated tangibility paradigm are inimical to the technology and business practices of the currency bazaar. This fact is strong, but not conclusive, evidence that including foreign exchange transactions in the scope of Article 2 is at variance with the needs of market participants. Thus, the pragmatic strategy suggests that Citibank's defense under the statute of frauds should be rejected. It indicates that DBS's response to that defense, while perhaps an over-reaction, has merit.

With respect to revised Article 2 and the CISG, the pragmatic strategy suggests that it may be appropriate to include the foreign exchange transactions in the scope of these regimes. They properly reject the tangibility paradigm. Because these sales laws omit a statute of frauds, neither revised Article 2 nor the CISG renders important transactions in the global currency bazaar, like the Citibank-DBS dollar-yen deal, unenforceable.

---

233. U.C.C. § 2-201 cmt. 1. The same requirements for a writing apply to § 2-201(1) and (2). To be sure, the official comment is deceptively simple. In effect, more may be required in a writing in order to satisfy the statute of frauds. *See* FARNSWORTH, *supra* note 159, § 6.7, at 409 (stating that the identity of the parties and the nature, subject matter, and the essential terms of the contract must be expressed).

234. Indeed, there is no separate quantity clause in the IFEMA. The definition of "FX Transaction" refers to a transaction between the parties "of an agreed amount" of one currency in exchange for another currency. IFEMA, *supra* note 14, § 1, at 4. Section 3.1 discusses the obligation of each party to deliver an "amount" of currency. *Id.* at 7.

235. The uncertainty is compounded by the fact that, assuming a quantity term is required, it is not clear what types of phrases constitute a "quantity term." At one extreme is a general quantity clause such as "a quantity of yen to be determined." The polar opposite is a specific quantity term, such as "120 million yen." The issue in choosing these alternatives (or some intermediate language) is whether to adopt "a mechanical construction of the quantity language of section 2-201(1)." Bruckel, *supra* note 174, at 815. For example, a mechanical construction is used in New York. *See* Int'l Commercial Resources, Ltd. v. Jamaica Pub. Serv. Co., 612 F. Supp. 1153, 1155 (S.D.N.Y. 1985) (holding that even though the total dollar amount of a transaction was established, a writing that referred to "various goods that [defendant] intends to purchase" and "various material and equipment" lacked a quantity term and, therefore, did not satisfy the statute of frauds) (citations omitted).

HITTING 'EM WHERE IT HURTS: USING TITLE IX LITIGATION TO BRING GENDER EQUITY TO ATHLETICS

MELODY HARRIS*

TABLE OF CONTENTS

Introduction 58
I. History of Title IX As It Applies to Athletics 60
II. Recent Athletics Litigation Under Title IX 63
A. *Roberts v. Colorado State University* 63
B. *Cohen v. Brown University* 63
C. *Cook v. Colgate University* 64
D. *Favia v. Indiana University of Pennsylvania* 64
III. Parties to Title IX Athletics Litigation 65
A. Private Right of Action Established in *Cannon* 65
B. Title IX Plaintiffs 65
1. Female Athletes: Individual Claims and Class Action Suits 66
2. Coaches and Tutors 66
3. Women's Organizations and Player Associations 67
4. Men 68
C. Title IX Defendants 69
IV. Substantive Issues in Title IX Athletics Cases: Case Elements and Affirmative Defenses 70
A. Stating a Prima Facie Athletics Case Under Title IX: Three Types of Violations 70
1. Financial Assistance Compliance Area 71
2. Other Athletic Benefits Compliance Area 74
3. Interests and Abilities Compliance Area 80
a. Substantially Proportionate 83
b. History and Continuing Practice of Program Expansion 87
c. Full and Effective Accommodation of Interests and Abilities 90
4. Discriminatory Intent: Not a Required Element 91
B. Affirmative Defenses to a Title IX Athletics Case 92
1. Team-by-Team Differences: Allowed 93
2. Temporary Annual Fluctuations: Allowed 93

* J.D. Harvard Law School, 1990; B.A. University of Denver, 1986. The author served as trial co-counsel for plaintiffs in Roberts v. Colorado State University, 814 F. Supp. 1507 (D. Colo. 1993), discussed *infra*. The author wishes to thank Karen Robertson, Pamela Gagel, Julie Trent, Ellen Vargyas and John Kobayashi (also counsel for plaintiffs) for their invaluable assistance in preparation for this article.

3. Voluntary Affirmative Action: Allowed 93
4. Compliance with NCAA or Other Rules: Rejected 94
5. Financial Constraints: Rejected 94
6. Equal Reductions in Male/Female Athletic Programs: Rejected 94
V. Remedies for Title IX Violations in Athletic Programs..... 95
A. Injunctive Relief 96
1. Success on the Merits..... 100
2. Irreparable Harm..... 100
3. Balancing the Interests 100
4. Public Interest..... 100
B. Monetary Relief..... 101
C. Attorneys' Fees 109
Conclusion..... 110
Appendix 111

INTRODUCTION

The important thing in [athletic competition] is not to win but to take part; the important thing in life is not the triumph but the struggle.
Pierre de Coubertin, founder of modern Olympics

Over twenty years ago, Congress passed Title IX of the Education Amendments of 1972 mandating gender equity in federally funded educational institutions.[1] Since that time, however, women have made little real progress in achieving equality in intercollegiate or interscholastic athletics. Seven years after Congress passed Title IX, the Department of Health, Education and Welfare found women made up only 30% of the intercollegiate athletes, but made up 48% of national undergraduate enrollment.[2] In the 1990s, despite over twenty years of Title IX's clear mandate against sex discrimination, women still make up only one third of Division I intercollegiate athletes, even though women actually make up over 50% of all college students enrolled.[3] Women athletes still lag well behind men in terms of financial assistance, operating dollars, and athletic participation opportunities afforded to them by their schools. Through weak enforcement of Title IX by the Office of Civil Rights, and alleged "voluntary compliance" efforts by educational institutions, Title IX had proved nothing more than placebo legislation with no real remedial effect on gender inequality in athletics.

Effective litigation under Title IX has begun to change all that. In the last few years, women athletes have taken Title IX enforcement into their own hands and taken their educational institutions to court. No longer satisfied with administrative complaints, informal requests, or strategic

1. 20 U.S.C. §§ 1681-88 (1988).

2. *See* Title IX of the Education Amendments of 1972; a Policy Interpretation; Title IX and Intercollegiate Athletics, 44 Fed. Reg. 71,413, app. A at 71,419 (1979) [hereinafter Final Policy Interpretation] (stating 1977-78 national averages).

3. *See* Alexander Wolff, *The Slow Track*, SPORTS ILLUSTRATED, September 28, 1992, at 52, 54-55.

protests[4] as their weapons against gender inequality in athletics, women athletes have enlisted attorneys to champion their cause for equality through litigation.

Although Title IX is over two decades old, litigation aimed at federally funded educational institutions' athletic departments is a relatively recent phenomenon. Four new cases have addressed some of the issues presented by this type of litigation.[5] These cases, however, have only scratched the surface of the issues presented by Title IX litigation aimed at intercollegiate or interscholastic athletics. Many issues, both procedural and substantive, remain unresolved by the courts. Using the course of litigation as a logical outline, this article explores the substantive and procedural issues presented by Title IX litigation in athletics. As background, the article first discusses the history of Title IX as it applies to athletics and the recent victories and defeats by women athletes in litigation aimed at bringing gender equity to intercollegiate athletics.

The article further discusses the prima facie elements of Title IX discrimination claims brought against educational institutions' athletic departments and legitimate affirmative defenses to those claims. Although courts in recent federal cases have found Title IX violations by schools, those courts failed to provide consistent methods of analysis by which future courts and litigants can gauge an educational institution's conduct. Thus, using statutory, regulatory and practical guidelines, this article presents a method of analysis for evaluating whether Title IX has indeed been violated. In particular, the article explores at what level of female athletic participation an educational institution will be found to be providing equal athletic opportunity to women in compliance with Title IX.

Finally, the article explores the remedies which are or should be available to athletes who successfully prove Title IX violations. In particular, this article presents scholarly debate of the availability of monetary damages for Title IX violations, extrapolating from and analyzing analogous statutes and Supreme Court decisions. Unlike other commentaries, this article concludes monetary damages should be available under Title IX even in instances of unintentional, or disparate impact, discrimination.

---

4. In 1976, the Yale women's crew team sought shower facilities in their boathouse. Polite protest went unheard. To get shower facilities, the women resorted to what has become known as the "Title IX strip"; the women painted the letters of Title IX on their bare backs and chests and disrobed for, among others, a New York Times photographer. Before resorting to such drastic measures, they had been reminded by the school administrators that the athletes were there to improve their athletic skills, not to powder their noses. *See* Deborah L. Rhode, *The "No-Problem" Problem: Feminist Challenges and Cultural Change*, 100 YALE L.J. 1731, 1762 (1991); *see also Yale Women Strip to Protest a Lack of Crew's Showers*, N.Y. TIMES, March 4, 1976, at A33.

5. Roberts v. Colorado State Univ., 814 F. Supp. 1507 (D. Colo.), *aff'd in part, rev'd in part sub nom.* Roberts v. Colorado State Bd. of Agric., 998 F.2d 824 (10th Cir.), *cert. denied*, 114 S. Ct. 580 (1993); Cohen v. Brown Univ., 809 F. Supp. 978 (D.R.I. 1992), *aff'd*, 991 F.2d 888 (1st Cir. 1993); Cook v. Colgate Univ., 802 F. Supp. 737 (N.D.N.Y. 1992), *vacated as moot*, 992 F.2d 17 (2d Cir. 1993); Favia v. Indiana Univ. of Pa., 812 F. Supp. 578 (W.D. Pa.), *aff'd*, 7 F.3d 332 (3d Cir. 1993).

I. HISTORY OF TITLE IX AS IT APPLIES TO ATHLETICS

Congress passed Title IX as a result of perceived gender discrimination in federally funded educational institutions. In relevant part, Title IX reads:

> No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance . . . .[6]

Title IX serves two main federal objectives: (1) to restrict the use of federal funds from supporting institutions engaging in discrimination; and (2) to protect individuals from that discrimination.[7] To meet these objectives, Congress explicitly granted authority to effectuate the provisions of Title IX to federal agencies that extend financial assistance to any educational institution, program, or activity.[8] The Department of Education has been the primary agency enforcing the Title.[9]

In 1974, the Department of Education's predecessor, the Department of Health, Education and Welfare ("HEW"), submitted proposed regulations implementing Title IX.[10] The regulations ("Regulations"), significantly revised from those proposed, were finally signed into law by President Gerald Ford on May 27, 1975.[11] Unlike Title IX itself, the Regulations explicitly address Title IX's application to athletics offered by federally funded educational institutions.[12] In this regard, the Regulations require educational institutions to provide "equal athletic opportunity for members of both sexes."[13] The Regulations also explicitly address the provision of scholarship monies by federally funded educational institutions.[14] The Regulations gave educational institutions three years to comply with its equal athletic opportunity requirements.[15] That period expired on July 21, 1978.[16]

Because Title IX called for the withdrawal of federal financial assistance from educational institutions that violated the statute's provisions, many educational institutions cried out for a clear delineation of the statute's requirements.[17] In response, HEW published a proposed Policy In-

---

6. 20 U.S.C. § 1681(a) (1988).

7. Cannon v. University of Chicago, 441 U.S. 677, 704 (1979).

8. 20 U.S.C. § 1682 (1988).

9. *See* Note, Franklin v. Gwinnett County Public Schools*: The Supreme Court Implies a Damages Remedy for Title IX Sex Discrimination*, 45 VAND. L. REV. 1367, 1368 & n.11 (1992).

10. 39 Fed. Reg. 22,228 (1974) (codified at 45 C.F.R. pt. 86 (1993)) (proposed June 20, 1974).

11. *See* North Haven Bd. of Educ. v. Bell, 456 U.S. 512, 531-33 (1982); Final Policy Interpretation, *supra* note 2, at 71,413 (detailing history of regulation); Diane Heckman, *Women & Athletics: A Twenty Year Retrospective on Title IX*, 9 U. MIAMI ENT. & SPORTS L. REV. 1, 12-13 (1992).

12. 34 C.F.R. § 106.41 (1993).

13. 34 C.F.R. § 106.41(c) (1993).

14. 34 C.F.R. § 106.37 (1993).

15. 34 C.F.R. § 106.41(d) (1993).

16. *See* Final Policy Interpretation, *supra* note 2, at 71,413.

17. *See* Heckman, *supra* note 11, at 13; *see also* Final Policy Interpretation, *supra* note 2, at 71,413 (discussing filing of 100 complaints against over 50 educational institutions during

terpretation on December 11, 1978, to give further guidance to educational institutions in providing equal athletic opportunity to members of both sexes.[18] After comments and revisions, HEW issued a Final Policy Interpretation on December 11, 1979, specifically addressing intercollegiate athletic programs.[19] HEW intended the Final Policy Interpretation's general principles to apply to club, intramural, and interscholastic athletics as well.[20] Although not reviewed or approved by Congress, the Final Policy Interpretation, as the definitive, published interpretation of the agency charged with implementing Title IX and its accompanying Regulations, is entitled to substantial deference by the courts.[21]

Despite the explicit language of the Regulations and the Final Policy Interpretation, several courts and commentators struggled with the issue of whether Title IX could apply to a specific program within an educational institution if that program did not receive direct federal financial assistance.[22] The "program-specific" debate was seemingly resolved when the Supreme Court decided *Grove City College v. Bell*.[23] There, the Supreme Court definitively ruled that federal financial assistance received by students did not inure to the benefit of the college as a whole but only to the financial aid department. Thus, Title IX only prohibited sex discrimination in the financial aid department and not in the school as a whole.[24] Following the Supreme Court, lower courts refused to apply Title IX to athletic programs because few, if any, receive direct federal funding.[25]

---

three year grace period and university community's lack of guidance in complying with Regulations).

18. *See* Final Policy Interpretation, *supra* note 2, at 71,413 (describing process of proposed Policy Interpretation and comments).

19. *Id.*

20. *Id.*

21. *See, e.g.*, Roberts v. Colorado State Univ., 998 F.2d 824, 828 (10th Cir.) (deferring substantially to the Final Policy Interpretation), *cert. denied*, 114 S. Ct. 580 (1993); *see also* Martin v. Occupational Safety & Health Review Comm'n., 499 U.S. 144, 149-50 (1991) (deferring to Commission's interpretation). This deference to the Final Policy Interpretation is supported by the Supreme Court's analysis of the analogous EEOC Guidelines interpreting Title VII. *See, e.g.*, Meritor Sav. Bank v. Vinson, 477 U.S. 57, 65 (1986) (holding the EEOC Guidelines "constitute a body of experience and informed judgment to which courts and litigants may properly resort for guidance") (citations omitted); Griggs v. Duke Power Co., 401 U.S. 424, 433-34 (1991) ("The administrative interpretation of [Title VII] by the enforcing agency is entitled to great deference."). Congress's failure to take any action disapproving the Final Policy Interpretation, particularly since Congress has addressed Title IX twice since 1979, creates a strong inference that Congress endorses HEW's interpretation. *See* North Haven Bd. of Educ. v. Bell, 456 U.S. 512, 535 (1982).

22. *Compare, e.g.*, Grove City College v. Bell, 687 F.2d 684, 700 (3rd Cir. 1982) (supporting institution-wide approach), *aff'd*, 465 U.S. 555 (1984) *with* Seattle Univ. v. HEW, 621 F.2d 992 (9th Cir. 1980) (per curiam) (supporting program-specific approach), *vacated as moot*, 456 U.S. 986 (1982).

23. 465 U.S. 555 (1984).

24. *Id.* at 572.

25. *See, e.g.*, Bennett v. West Tex. State Univ., 799 F.2d 155 (5th Cir. 1986) (holding ministerial relationship between financial aid department, which received federal funds, and athletic program insufficient to bring the latter under Title IX). On the general lack of direct receipt of federal financial assistance by intercollegiate athletic programs, see Note, *Sex Discrimination in Intercollegiate Athletics*, 61 IOWA L. REV. 420, 469 (1975); B. Glenn George, *Miles to Go and Promises to Keep: A Case Study in Title IX*, 64 U. COLO. L. REV. 555, 558 & n.16

Congress, however, had the last word in the "program-specific" debate when it overruled the Supreme Court's decision in *Grove City College v. Bell* with the Civil Rights Restoration Act of 1987:

> The Congress finds that—(1) certain aspects of recent decisions and opinions of the Supreme Court . . . cast doubt upon the broad application of title IX of the Education Amendments of 1972 . . . and (2) legislative action is necessary to restore the prior consistent and long-standing executive branch interpretation and broad, institution-wide application of those laws as previously administered.[26]

Under the Civil Rights Restoration Act, Title IX's prohibitions against sex discrimination apply institution-wide to any educational institution receiving federal financial assistance in any form.[27]

In addition to the Regulations and Final Policy Interpretation, the Department of Education, through the Office for Civil Rights ("OCR"), further interpreted Title IX as it applied to athletics by issuing the Title IX Athletics Investigator's Manual ("Investigator's Manual").[28] The manual was designed to assist OCR investigators in their investigations of intercollegiate and interscholastic athletic programs.[29] For the same reasons the Final Policy Interpretation is given great deference,[30] so too should the Investigator's Manual receive deference from the courts. Unlike the Final Policy Interpretation, however, the Investigator's Manual is an internal agency document which was not subject to public notice and comment. Therefore, any inconsistencies between the Final Policy Interpretation and the Investigator's Manual must be resolved in favor of the Final Policy Interpretation as the implementing agency's authoritative interpretation of the Regulations.

The Investigator's Manual is divided into thirteen areas that address specific components of athletics programs. These areas are discussed below in relation to litigation initiated under Title IX.

(1993). At least one lower court, however, found Title IX applied to an institution's athletic department. Haffer v. Temple Univ., 678 F. Supp. 517, 537-38 (E.D. Pa. 1987) (holding athletic scholarships are subject to Title IX scrutiny because federal financial support is given to financial aid department and athletic scholarships are a part of that department).

26. Civil Rights Restoration Act of 1987, Pub. L. 100-259, § 2, 102 Stat. 28 (1988) (codified at 20 U.S.C. § 1687 (1988)); *see also* West Virginia Hosps., Inc. v. Casey, 499 U.S. 83, 113-14 (1991)).

27. *Id.* The retroactive application of the Civil Rights Restoration Act is now irrelevant. Only cases pending before the Act would need to address retroactivity. Whether the Act is retroactive or not, evidence of discrimination prior to the enactment of the Civil Rights Restoration Act is relevant and admissible to prove that an institution's present conduct perpetuates past discrimination. *See* Bazemore v. Friday, 478 U.S. 385, 396 n.6 (1986) (Brennan, J., concurring); EEOC v. University of N.M., 504 F.2d 1296, 1301 (10th Cir. 1974).

28. VALERIE M. BONNETTE & LAMAR DANIEL, OFFICE FOR CIVIL RIGHTS, U.S. DEP'T OF EDUC., TITLE IX ATHLETICS INVESTIGATOR'S MANUAL (1990) [hereinafter INVESTIGATOR'S MANUAL].

29. *See id.* at [i].

30. *See supra* note 21 and accompanying text.

II. Recent Athletics Litigation Under Title IX

Federal Courts have heard and decided a handful of key Title IX cases dealing with gender equity in athletics within the last few years. These cases have been and will continue to be central to the evolving legal issues in determining the application of Title IX's gender equity directive to athletics. This section briefly discusses these important cases and sets the scene for a discussion of issues that have been or will be pursued by women athletes under Title IX.

A. Roberts v. Colorado State University[31]

In *Roberts*, the Colorado State University women's softball team filed suit in their individual capacities to have their team reinstated after its abrupt termination by the University. The women softball players sued under the interests and abilities compliance area of Title IX,[32] alleging Colorado State failed to provide equal athletic opportunity to women in its federally funded athletic department.[33]

The federal district court found Colorado State in violation of the interests and abilities compliance area and granted the softball players a permanent injunction requiring Colorado State to reinstate the softball team with all incidental benefits accorded other varsity sports at Colorado State.[34] The Tenth Circuit affirmed in part and reversed in part. The Tenth Circuit ruled the district court had overstepped its authority by ordering Colorado State to organize a fall season of play for the softball team that had not been the practice at Colorado State prior to the team's termination.[35] Nevertheless, the Tenth Circuit agreed Colorado State was in violation of Title IX's interests and abilities compliance area and fully supported the district court's order to reinstate the softball team.[36]

B. Cohen v. Brown University[37]

In *Cohen*, members of the women's gymnastics and volleyball teams brought a class action suit seeking a preliminary injunction to reinstate the varsity status of their teams. The women athletes alleged a violation of the interests and abilities compliance area of Title IX and presented evidence of unequal treatment in receiving other athletic benefits.[38] The district court agreed with the women athletes and granted a preliminary injunction restoring women's gymnastics and volleyball to varsity status and mandating adequate funding and benefits to those reinstated teams.[39] The

31. 814 F. Supp. 1507 (D. Colo.), *aff'd in part, rev'd in part sub nom.* Roberts v. Colorado State Bd. of Agric., 998 F.2d 824 (10th Cir.), *cert. denied*, 114 S. Ct. 580 (1993).
32. *See infra* section IV.A.3.
33. *Roberts*, 814 F. Supp. at 1510-11.
34. *Id.* at 1518-19.
35. *Roberts*, 998 F.2d at 835.
36. *Id.* at 832-34.
37. 809 F. Supp. 978 (D.R.I. 1992), *aff'd*, 991 F.2d 888 (1st Cir. 1993).
38. *Id.* at 994-97.
39. *Id.* at 1001.

First Circuit affirmed the district court's preliminary injunction order in full.[40]

C. Cook v. Colgate University[41]

Cook involved claims by the women's club ice hockey team, in their individual capacities, that Colgate's refusal to upgrade them to varsity status violated Title IX. The women athletes alleged a violation of the other athletic benefits compliance area,[42] stating that the men's ice hockey team, and the men's athletics program as a whole, received far greater benefits in such areas as equipment, funding, travel, and coaching.[43] The magistrate judge at the district court level ordered Colgate to grant varsity status to the women's ice hockey team and to supply the team with "equivalent athletic opportunities."[44] The Second Circuit vacated the order and remanded the case.[45] Since the athletes brought their claims in their individual capacities, not as representatives of the hockey team, and because the last of the plaintiffs would graduate before the order to upgrade the team could take effect, the Second Circuit ruled their claims were moot and remanded the action with orders to dismiss the complaint.[46]

D. Favia v. Indiana University of Pennsylvania[47]

In *Favia*, athletes on the women's gymnastics and field hockey teams brought a class action suit against Indiana University of Pennsylvania ("IUP") seeking reinstatement of their teams. Like the plaintiffs in *Roberts*, the *Favia* plaintiffs alleged a violation of the interests and abilities compliance area of Title IX. The district court agreed that IUP failed to provide equal athletic opportunity to women and ruled IUP was in violation of Title IX.[48] The Third Circuit affirmed and denied IUP's motion to modify the injunction.[49] Nevertheless, the Third Circuit suggested an appropriate order for final injunctive relief should focus not so much on the interests of the representative plaintiffs, but instead, should allow IUP to institute women's sports other than gymnastics or field hockey which could meet the needs of more women athletes.[50]

---

40. Cohen v. Brown Univ., 991 F.2d 888, 907 (1st Cir. 1993).
41. 802 F. Supp 737 (N.D.N.Y. 1992), *vacated as moot*, 992 F.2d 17 (2d Cir. 1993).
42. *See infra* section IV.A.2.
43. *Cook*, 802 F. Supp. at 744-45.
44. *Id.* at 751.
45. Cook v. Colgate Univ., 992 F.2d 17, 20 (2d Cir. 1993).
46. *Id.*
47. 812 F. Supp. 578 (W.D. Pa.), *aff'd*, 7 F.3d 332 (3d Cir. 1993).
48. *Id.* at 584-85.
49. Favia v. Indiana Univ. of Pa., 7 F.3d 332, 344 (3d Cir. 1993).
50. *See id.* at 344.

III. Parties to Title IX Athletics Litigation

A. *Private Right of Action Established in* Cannon v. University of Chicago

Title IX began as a public remedy statute wherein federal administrative agencies were given authority to withdraw federal funds from educational institutions that discriminated on the basis of sex.[51] For several years after Title IX's passage, courts interpreted Title IX's relief provisions to foreclose any private right of action by individuals who were harmed as a result of an educational institution's discriminatory practice.[52] Thus, Title IX was relegated to "administrative legislation" whereby administrative agencies had discretion to pursue individual claims against an educational institution. In such claims, an individual complainant had no participation in the investigation or subsequent enforcement proceedings.[53]

In 1978, the Supreme Court gave the reins of Title IX to the individuals it was meant to protect.[54] In *Cannon v. University of Chicago*,[55] the Supreme Court found Title IX implied a private right of action for individuals harmed by the discriminatory practices of a federally funded educational institution.[56] The Court so held because (1) Title IX was enacted for the benefit of a specific class,[57] (2) Congress intended to create a private right of action,[58] (3) a private action would not frustrate the underlying purpose of Title IX,[59] and (4) implying a federal remedy would not infringe states' rights.[60] Thus began the enforcement of individuals' rights under Title IX through private litigation.

B. *Title IX Plaintiffs*

Under Title IX, an appropriate plaintiff is any person who is excluded from, denied the benefits of, or subjected to discrimination under any educational program which receives federal funds.[61] As expressed by the Court in *Cannon*, persons who are "private victims of discrimination" are

---

51. 20 U.S.C. §§ 1681-1682 (1988).

52. *See, e.g.*, Cannon v. University of Chicago, 559 F.2d 1063, 1073 (7th Cir. 1976) (holding female applicant twice denied admission to medical school had no right of action against the institution under Title IX), *rev'd*, 441 U.S. 677 (1979).

53. Cannon v. University of Chicago, 441 U.S. 677, 707 n.41 (1979) (describing administrative enforcement process).

54. Congress's intent to protect individual women from discrimination is evident from Title IX's legislative history. Congress passed Title IX with the express intent to provide solid legal protection to women as they seek education and training for later careers. *See* 118 Cong. Rec. 5806-07 (1972) (comments of Sen. Bayh); *see also* Final Policy Interpretation, *supra* note 2, at 71,423 (stating legislative history of Title IX demonstrates it was enacted because of discrimination against women in educational institutions).

55. 441 U.S. 677 (1979).

56. *Id.* at 709.

57. *Id.* at 694.

58. *Id.* at 703 (analogizing to Title VI remedies in which the Court understood Congress authorized private causes of action for victims of prohibited discrimination).

59. *Id.* at 703-04 (identifying purposes of Title IX as (1) avoiding use of federal funds to support discrimination, and (2) providing individual citizens effective protection against discrimination).

60. *Id.* at 708-09.

61. *See* 20 U.S.C. § 1681(a).

entitled to bring a claim under the statute.[62] This section explores the standing of individuals and organizations to bring Title IX claims against educational institutions' athletics departments allegedly violating Title IX.

1. Female Athletes: Individual Claims and Class Action Suits

Individual female athletes who allege they have been excluded from, denied the benefits of, or subjected to discrimination under their educational program obviously fit squarely within Title IX's standing provision.[63] Indeed, Title IX's focus, as it applies to athletics, is on the protection of individual students' rights, not the rights of any group or subgroup of students.[64]

This focus on the individual's rights, however, does not prevent individuals from joining together against an educational institution in a single lawsuit. Indeed, courts have allowed women to maintain class action suits on behalf of entire teams or on behalf of all present and future female students at an institution who participate in or are deterred from participating in intercollegiate athletics.[65] Certain consequences will result at the remedial stage of trial depending on whether plaintiffs bring their claims as individuals or in class actions. For example, an individual plaintiff may be able to recover monetary damages and specific injunctive relief, but may be faced with mootness problems for prospective injunctive relief. A class of athletes or students, on the other hand, may not face the mootness argument, but may lack standing to request specific injunctive relief, such as the reinstatement of a particular team.[66] Thus, special attention must be paid early on to the form of the suit and the remedy sought.

2. Coaches and Tutors

Coaches and tutors are an integral part of interscholastic and intercollegiate athletics and, as such, are afforded protection under Title IX from sex discrimination.[67] Where a coach is denied benefits by an educational institution on the basis of his or her sex, Title IX offers a remedy to that individual.[68] Nevertheless, these coaches or tutors must seek to vindicate their own rights, not those of the student athletes they coach or teach. The Regulations governing individual athletes' rights do not afford coaches and tutors any protection. Instead, coaches and tutors are af-

62. *Cannon*, 441 U.S. at 709.

63. *See id.*

64. *See* Final Policy Interpretation, *supra* note 2, at 71,422.

65. *See Favia*, 812 F. Supp. at 579; *Cohen*, 809 F. Supp. at 979.

66. *See Roberts*, 998 F.2d at 833 (dictum). For a full discussion of the remedial consequences of the plaintiffs' status, see *infra* notes 259 to 297 and accompanying text.

67. *See* North Haven Bd. of Educ. v. Bell, 456 U.S. 512, 530 (1982) (reading Title IX to give rights to employees of educational institutions as well as students).

68. *Id.*

forded protections as *employees* of the educational institutions and are governed by different provisions of the Regulations.[69]

Recently, female coaches have filed lawsuits under Title IX and other employment provisions alleging unequal pay and treatment as a result of their sex.[70] While these suits are a fascinating new development in the movement toward gender equity in intercollegiate athletics, this article does not discuss the elements of these employment disputes.

3. Women's Organizations and Player Associations

Organizations or associations whose members are likely to suffer harm as a result of unlawful discrimination by educational institutions are potential plaintiffs in a Title IX suit. For example, organizations that represent women's interests or, more specifically, women athletes' interests, may bring suit under Title IX. An association may have standing to sue as the representative of its individual members: "injury to the association's members will satisfy Article III and allow that organization to litigate in federal court on their behalf."[71] Thus, as long as the organization can satisfy the three-part test required for associational standing, the organization may file suit to enforce Title IX.

The three-part test requires that "(a) [the association's] members would otherwise have standing to sue in their own right; (b) the interests it seeks to protect are germane to the organization's purpose; and (c) neither the claim asserted nor the relief requested requires the participation of individual members in the lawsuit."[72] The National Organization for Women employed this associational standing doctrine to represent its members in a California suit brought under Title IX.[73]

In choosing to bring an action on behalf of their members, however, associations must be prepared to accept limited remedies that may give ultimate control over compliance methods to the educational institutions. Like the remedies available in class actions, the remedies available to a representative organization are limited to general prospective injunctive relief that will benefit the class as a whole.[74] Monetary relief will likely not be recoverable.[75] Moreover, the educational institution may be allowed to

---

69. 34 C.F.R. § 106.51 (1993) (prohibiting sex discrimination in employment in educational institutions).

70. *See, e.g.*, Stanley v. University of S. Cal., 13 F.3d 1313 (9th Cir. 1994).

71. International Union, UAW v. Brock, 477 U.S. 274, 281 (1985); *see also* Hunt v. Washington State Apple Advertising Comm'n, 432 U.S. 333, 343 (1977) (discussing in general an organization's standing to sue on behalf of its employees); Simon v. Eastern Ky. Welfare Rights Org., 426 U.S. 26, 40 (1976) (same); Sierra Club v. Morton, 405 U.S. 727, 729 (1972) (same).

72. *Hunt*, 432 U.S. at 343.

73. *See* Complaint for Injunctive and Declaratory Relief, California Nat'l Org. for Women v. Evans (No. 728548) (filed in the Superior Court of the State of California for the County of Santa Clara, February 3, 1993) (settled before trial).

74. *See infra* notes 259 to 297 and accompanying text.

75. *Id.*

direct its own compliance efforts rather than being judicially directed or "micromanaged."[76]

4. Men

Men have standing to bring suits under Title IX if they can establish they have been discriminated against or denied equal benefits on the basis of sex. This begs the question, however, whether a suit under Title IX can be successful when brought on behalf of men. Male athletes will have difficulty establishing a prima facie case[77] under Title IX or its Regulations since male athletes are the overrepresented gender in virtually every athletic department in the country.[78] The only court that has decided a case brought by male athletes under Title IX granted summary judgment against the plaintiffs because it found Title IX afforded no protection to the male athletes before it. In *Kelley v. Board of Trustees of the University of Illinois*,[79] male athletes brought a Title IX suit alleging sex discrimination because the University of Illinois had terminated their varsity swimming team while retaining the women's swimming team. In granting summary judgment for the university, the district court ruled that male athletes were not entitled to protection at the University of Illinois because male athletic participation percentages exceeded male undergraduate enrollment percentages.[80] While the plain meaning of Title IX would seem to outlaw a school's decision to cut one team over another on the basis of sex, the Regulations and the Final Policy Interpretation allow (and cases have allowed) affirmative actions taken on behalf of the underrepresented gender,[81] traditionally women. "Quite frankly, these interpretations [in the Regulations and Final Policy Interpretation] have converted Title IX from a statute which prohibits discrimination on the basis of sex (defined as the elimination of or exclusion from participation opportunities), into a statute which provides 'equal opportunity for members of both sexes.' "[82]

Despite this conclusion by the *Kelley* court that Title IX has somehow been transformed by the Regulations and the Final Policy Interpretation, the legislative history behind Title IX demonstrates Congress intended it to serve primarily as a remedy for current discrimination in educational

---

76. *See infra* notes 283 to 288 and accompanying text.

77. *See infra* part IV.A.

78. *See* Wolff, *supra* note 3, at 54-55 (stating that national average for Division I women athletic participation is approximately 30%, while national female college enrollment is approximately 50%); *see also* Steve Wieberg, *Title IX: 20 Years After, A New Call for Action*, USA TODAY, June 8, 1992, at C10 (reporting on 86 Division I-A schools and finding that only 12 provided women's athletics more than 30% of total athletic funding).

79. 832 F. Supp. 237 (C.D. Ill. 1993), *aff'd*, No. 93-3205, 1994 WL 473875 (7th Cir. Sept. 1, 1994).

80. *Id.* at 242.

81. *See id.*; *see also* 34 C.F.R. §§ 106.3(b), 106.41(c) (1993) (allowing affirmative action and requiring equal opportunity, respectively); Final Policy Interpretation, *supra* note 2, at 71,416 (voluntary affirmative action measures are legitimate nondiscriminatory reasons for unequal treatment of male and female teams).

82. *Kelley*, 832 F. Supp. at 241.

institutions *against women.*[83] Thus, whether Title IX is labeled as "anti-discrimination" or "affirmative action" legislation, the Regulations and the Final Policy Interpretation have expanded Title IX's legislative directive to equalize federally financed educational opportunities for women with those opportunities already provided to men. This expansion must be regarded as an implementation of the legislative purpose behind Title IX and not as a transformation of the statute as concluded by *Kelley.*

Consequently, although male athletes have technical standing to bring suit under Title IX, their suits will be subject to an immediate motion to dismiss, which will likely be granted, unless they can prove they are underrepresented in the athletic department.[84] While men may argue this "affirmative action" approach violates constitutional principles of equal protection, federal courts have long recognized that acts taken and legislation passed by governmental entities for the benefit of an historically deprived group will withstand constitutional challenge.[85] Moreover, the current exclusion of men from Title IX's remedy is not offensive as long as men continue to receive disproportionately higher athletic opportunities than women. Title IX was passed as a result of women's need for solid legal protection as they seek education and training for their later careers.[86] Once an educational institution establishes athletic participation equality, either gender will be able to bring suit under Title IX for actions taken by an educational institution motivated by gender considerations or which unfairly impact one gender. Until that time, however, men are protected by institutional and historical biases operating in their favor, obviating the current need for Title IX protection.

C. *Title IX Defendants*

Any educational institution receiving federal funds that allegedly fails to provide a discrimination-free athletic department will be subject to suit under Title IX.[87] An "educational institution" is defined as:

> any public or private preschool, elementary, or secondary school, or any institution of vocational, professional or higher education, except that in the case of an educational institution composed of more than one school, college, or department which are admin-

---

83. Final Policy Interpretation, *supra* note 2, at 71,423. "Millions of women pay taxes into the Federal treasury and we collectively resent that these funds should be used for the support of institutions to which we are denied equal access." 117 CONG. REC. 39,252 (1971) (comments of Rep. Mink). *See also* Cannon v. University of Chicago, 441 U.S. 677, 704 n.36 (1979) (discussing legislative intent of Title IX). Title IX was passed "in view of the scope and depth of discrimination" against women in educational institutions. H.R. REP. No. 554, 92d Cong., 1st Sess. 51 (1971), *reprinted in* 1972 U.S.C.C.A.N. 2462, 2512.

84. In *Cohen,* the First Circuit rejected Brown's claim that effective accommodation for women athletes violates the equal protection rights of men athletes. *Cohen,* 991 F.2d at 900-01. The court held that Congress had broad authority to remedy past discrimination and had intended to do so under Title IX. *Id.* at 901.

85. *See, e.g.,* Mississippi Univ. for Women v. Hogan, 458 U.S. 718, 728 (1982).

86. *See* 118 CONG. REC. 5806-07 (1972) (comments of Sen. Bayh).

87. 20 U.S.C. § 1681(a) (1988).

> istratively separate units, such term means each such school, college, or department.[88]

Certain types of educational institutions, such as military institutions, are excluded from Title IX's reach.[89] Individual universities, colleges, and schools are traditional defendants in Title IX litigation. Depending on the charter of the educational institution, however, the proper defendant in place of the institution may be a governing body, for example, the Colorado State Board of Agriculture (which governs Colorado State University).[90]

At the secondary and post-secondary levels, separate schools and colleges often form together to form athletic groups or associations. The National Collegiate Athletic Association ("NCAA") is such a group; it is "a voluntary association of public and private institutions."[91] In defining the term "educational institution," Title IX refers to the individual schools or colleges within these groups. Thus, Title IX directs that the proper defendant is not the NCAA itself, but the individual members of the NCAA. This is especially true since the rules of the NCAA, or other such group, are subject to change by vote of the individual members.[92]

The Eleventh Amendment[93] serves as no protection to liability. State run schools are deemed to have waived Eleventh Amendment immunity by accepting federal funds.[94]

IV. Substantive Issues In Title IX Athletics Cases: Case Elements and Affirmative Defenses

A. *Stating a Prima Facie Athletics Case Under Title IX: Three Types of Violations*

The Regulations, Final Policy Interpretation, and the Investigator's Manual all set out three areas in which an educational institution must comply in order to avoid a Title IX violation in its athletics department.[95] These areas, referred to throughout this article as compliance areas, are: (1) athletic financial assistance; (2) other athletic benefits and opportunities; and (3) accommodation of athletics interests and abilities.[96] The Final Policy Interpretation sets the three compliance areas apart as separate considerations, even though the Regulations combine the second and

---

88. 20 U.S.C. § 1681(c) (1988).

89. 20 U.S.C. § 1681(a)(1)-(9).

90. *See Roberts*, 998 F.2d at 827.

91. Arlosoroff v. National Collegiate Athletic Ass'n, 746 F.2d 1019, 1021 (4th Cir. 1984).

92. *See* Final Policy Interpretation, *supra* note 2, at 71,422 (suggesting suit must be against individual institutions to change NCAA rules, since compliance with NCAA rules does not serve as defense to member institutions).

93. "The Judicial power of the United States shall not be construed to extend to any suit in law or equity, commenced or prosecuted against one of the United States by Citizens of another State . . . ." U.S. Const. amend. XI.

94. *See* 42 U.S.C. § 2000d-7(a) (1988).

95. 34 C.F.R. §§ 106.37(c), 106.41(c)(1)-(10) (1993); Final Policy Interpretation, *supra* note 2, at 71,415-17; Investigator's Manual, *supra* note 28, at 7.

96. Final Policy Interpretation, *supra* note 2, at 71,415-17; Investigator's Manual, *supra* note 28, at 7.

third compliance areas.[97] Moreover, the Investigator's Manual states, "the [Final] Policy Interpretation does permit separate investigations and findings for three major areas" and goes on to delineate the three compliance areas.[98]

Educational institutions have argued that Title IX requires a program-wide analysis, pointing to the Regulations and the Investigator's Manual as support for this approach.[99] Specifically, the Investigator's Manual calls for a balancing of all compliance areas against each other to determine if there is overall compliance with Title IX.[100] Thus, schools have argued that a violation of a single compliance area cannot state a valid claim under Title IX.[101]

Nevertheless, courts have found that failure to provide equal treatment to female athletes in any one of the three compliance areas laid out in the Final Policy Interpretation gives rise to a violation under Title IX.[102] "Although § 106.41(c) [of the Regulations] goes on to list nine other factors that enter into a determination of equal opportunity in athletics, an institution may violate Title IX simply by failing to effectively accommodate the interest and abilities of student athletes of both sexes."[103] The same logic dictates that a violation of the financial assistance or other athletic benefits compliance areas states a separate, valid claim under Title IX, although no court has addressed the issue.

This article analyzes each compliance area. Particular emphasis is given to the third compliance area, accommodation of athletic interests and abilities, because it forms the cornerstone of Title IX as it applies to athletics.

1. Financial Assistance Compliance Area

The first compliance area relates to the financial assistance or scholarships an educational institution provides to its student athletes. Scholarships are specifically governed by section 106.37(c) of the Regulations.[104] "[Institutions] must provide reasonable opportunities for [athletic scholar-

97. *Compare* Final Policy Interpretation, *supra* note 2, at 71,415-17 *with* 34 C.F.R. § 106.41(c)(1)-(10).

98. INVESTIGATOR'S MANUAL, *supra* note 28, at 7.

99. *See, e.g., Cohen*, 809 F. Supp. at 988-89 (defendants arguing "that Title IX requires a program-wide analysis to determine compliance," not a single violation of 34 C.F.R. § 106.41(c)(1)).

100. INVESTIGATOR'S MANUAL, *supra* note 28, at 7 (suggesting, in general, an overall approach that reviews the entire athletic program).

101. *See, e.g., Roberts*, 814 F. Supp. at 1510-11 (defendants arguing that "the Investigator's Manual requires that plaintiffs demonstrate an overall violation of either 34 C.F.R. § 106.37(c) or 34 C.F.R. § 106.41(c)(1)-(10) in order to sustain a claim of discrimination under Title IX").

102. *See Cook*, 802 F. Supp. at 742 (rejecting Colgate's argument that overall compliance with Title IX in the athletic department as a whole insulates Colgate from liability for specific Title IX violations); *see also Roberts*, 998 F.2d at 828 (holding interests and abilities violation sufficient to find Title IX violation); *Cohen*, 991 F.2d at 897 (holding failure to accommodate interests and abilities sufficient for Title IX violation).

103. *Roberts*, 998 F.2d at 828 (citation omitted).

104. 34 C.F.R. § 106.37(c).

ship] awards for members of each sex in proportion to the number of students of each sex participating in . . . intercollegiate athletics."[105]

The Final Policy Interpretation clearly defines the standard for measuring compliance in this area. The OCR and courts are to divide the amount of financial assistance to members of each sex by the number of male or female participants[106] in the athletic program. If the comparison results in a substantially equal amount or the school can offer nondiscriminatory reasons for a disparity, the school has satisfied this compliance area.[107] The compliance area does not require a proportionate *number* of scholarships; rather, it requires an *amount* of financial assistance proportionate to men's and women's participation rates.[108] In analyzing these amounts, non-grant aid, such as work-study, also must be compared.[109]

In assessing an institution's purported nondiscriminatory reasons for disparities in the amount of awards to men and women, the courts should exercise caution. The Final Policy Interpretation sets out examples of alleged nondiscriminatory reasons that educational institutions put forth during the notice and comment period for the proposed Policy Interpretation. For example, the Final Policy Interpretation suggests a school may provide unequal amounts of aid where more students of one gender hail from out of state.[110] Yet, the OCR's experience in investigating schools revealed that this alleged nondiscriminatory reason for unequal financial assistance to women athletes could be the result of another discriminatory practice by the institution, namely greater out-of-state recruiting for the men's athletic program.[111] "A disparity in recruitment of student athletes may not be used to justify a disparity in athletic financial assistance."[112] Likewise, other discriminatory practices by the institution cannot be used to justify disproportionate amounts of financial assistance to members of each sex.

In measuring the proportionality of financial assistance, the OCR employs complicated statistical analyses known as the "Z"[113] and "T"[114] tests.[115] Where the percentage of athletes who are female is not propor-

105. *Id.*

106. The Final Policy Interpretation defines "participants" as those athletes who: "receive institutionally-sponsored support normally provided to athletes" (e.g. coaching, equipment); "are participating in organized practice sessions and other team meetings . . . on a regular basis"; "are listed on the eligibility or squad lists maintained for each sport"; or "[w]ho, because of injury, cannot meet [any of the above requirements] but continue to receive financial aid on the basis of athletic ability." Final Policy Interpretation, *supra* note 2, at 71,415; INVESTIGATOR'S MANUAL, *supra* note 28, at 14.

107. Final Policy Interpretation, *supra* note 2, at 71,415.

108. *Id.*

109. *Id.*

110. *Id.*

111. INVESTIGATOR'S MANUAL, *supra* note 28, at 20.

112. *Id.*

113. The "Z" test measures whether the difference between the percentage of total aid awarded to one sex and the percentage of participants of that sex in the athletic program is significant. INVESTIGATOR'S MANUAL, *supra* note 28, at 153.

114. The "T" test measures "whether the difference between the average [financial] award to male and female athletes is significant." *Id.*

115. *Id.* at 19, app. D at 153-63.

tionate to the percentage of financial assistance provided to female athletes, the statistical tests are employed by OCR to determine whether there is a violation of the financial assistance compliance area. The tests are not used to determine whether there is a difference between the awards of financial assistance; that determination is assumed by the application of the "Z" or "T" test. Instead, the tests are "used for determining whether differences in the awards of athletic financial assistance to male and female athletes are significant."[116]

Neither the Regulations nor the Final Policy Interpretation requires a plaintiff to use this complicated statistical analysis to demonstrate a violation of the financial assistance compliance area. Instead, a plaintiff need merely show that the ratio of aid to participants is not substantially proportionate.[117] Where, for example, women make up 34% of the athletic participants, but receive only 28%[118] of the athletic financial assistance dollars, a court may hold, based on the 6% disparity alone, that the aid to participants ratio is not substantially proportionate.[119] Although statistical analysis can serve as useful evidence in determining whether the aid to participants ratio is substantially proportionate, the "Z" and "T" tests are not bright-line standards a plaintiff must meet in order to state a violation.[120]

On the surface, schools appear to be meeting this compliance area. In or about 1992, women athletes made up approximately one-third of the athletes in NCAA Division I schools and received approximately one-third of athletic scholarship dollars.[121] Nevertheless, women made up approximately 50% of the undergraduate populations at those schools.[122] As will be discussed below, this nearly 20% differential between enrollment and female athletic participation violates the interests and abilities compliance area stated in section 106.41(c)(1) of the Regulations.[123] It therefore defies logic that section 106.37 of the Regulations would permit schools to provide scholarships in proportion to a number of women athletes that

116. *Id.* at 153.

117. 34 C.F.R. § 106.37(c).

118. INVESTIGATOR'S MANUAL, *supra* note 28, at 19 (sample numbers used in manual's analysis).

119. Indeed, courts have made decisions on the "substantially proportionate" analysis required under the third compliance area discussed *infra* at section IV.A.3.a. *See, e.g., Roberts*, 814 F. Supp. at 1511-13 (10.6% disparity between female athletic participation and female undergraduate enrollment unacceptable absent a showing of effective accommodation); *Favia*, 812 F. Supp. at 584-85 (school failed to provide opportunities to participate when 36.51% of women were involved in intercollegiate athletics while 55.61% of students enrolled were women). Courts have made these decisions without resorting to the complicated analyses proposed in the "Z" and "T" tests. Courts have, however, considered expert testimony on whether the disparities in the proportionality are statistically significant. None have found that statistical significance is a necessary element of a valid claim of violation of the third compliance area. Likewise, statistical significance is not an element under the financial assistance compliance area.

120. On discouraging bright-line standards for the substantially proportionate test of the third compliance area, see *infra* notes 201-14 and accompanying text.

121. Wolff, *supra* note 3, at 54-55.

122. *Id.* at 54.

123. *See infra* part IV.A.3.

violates section 106.41(c)(1). Thus, even where schools are fully funding scholarships for its women athletes in proportion to their participation rates, this will not constitute compliance with Title IX if those participation rates are not proportionate to the enrollment rates, in violation of section 106.41(c)(1).[124]

Commentators have noted this problem with the financial assistance compliance area in their calls for reform of the Title IX Regulations.[125] Short of legislative reform, however, courts should follow the Investigator's Manual's caution against accepting a school's nondiscriminatory reasons for disparities in financial assistance provided to female athletes.[126] Where the athletic participation numbers themselves are in violation of Title IX, those numbers cannot justify failure to provide more financial assistance to women athletes.[127]

2. Other Athletic Benefits Compliance Area

The second compliance area deals with the provision of other athletic benefits and opportunities. Here, the Regulations set out ten criteria to consider in determining whether the educational institution is providing "equal athletic opportunity" to members of both sexes.[128] These criteria are: (1) effective accommodation of interest and ability; (2) provision and maintenance of equipment and supplies; (3) scheduling of games and practice times; (4) travel and per diem expenses; (5) opportunity to receive coaching and academic tutoring; (6) assignment and compensation of coaches and tutors; (7) provision of locker rooms, practice, and competitive facilities; (8) provision of medical and training services and facilities; (9) provision of housing and dining services and facilities; and (10) publicity.[129] The Regulations also allow the consideration of other factors in measuring compliance.[130]

In addition to these ten criteria, the Final Policy Interpretation adds two other factors useful in determining compliance under this compliance area: recruitment of student athletes and provision of support services.[131] Furthermore, the Final Policy Interpretation specifically separates the first criterion, effective accommodation of interests and abilities, from the other nine criteria set out in the Regulations.[132] The nine remaining cri-

124. *See Haffer v. Temple Univ.*, 678 F. Supp. 517, 539 (E.D. Pa. 1987).

125. *See, e.g.*, Heckman, *supra* note 11, at 63.

126. INVESTIGATOR'S MANUAL, *supra* note 28, at 20 (disallowing disparities in recruitment to justify failure to provide more athletic financial assistance and scrutinizing the award of scholarship funds based on "reasonable professional decisions" when decision negatively affects athletes of one sex).

127. *See Haffer*, 678 F. Supp. at 539.

128. 34 C.F.R. § 106.41(c)(1)-(10).

129. *Id.*

130. *Id.*; *see also* Final Policy Interpretation, *supra* note 2, at 71,415.

131. Final Policy Interpretation, *supra* note 2, at 71,415.

132. *Id.* at 71,415 n.3. The Final Policy Interpretation makes effective accommodation of interests and abilities a separate compliance area. *Id.* at 71,417. For discussion of the interests and abilities compliance area, see *infra* part IV.A.3.

teria plus the two added by the Final Policy Interpretation make up the other athletic benefits compliance area.[133]

The Investigator's Manual treats each of the eleven criteria of this compliance area separately.[134] Nevertheless, the criteria should be evaluated together as a whole to determine whether an institution satisfies the other athletic benefits compliance area.[135] Indeed, this compliance area provides for an equitable weighing of each of the eleven criteria against the others to determine whether, in totality, the institution provides equal opportunity.[136] The main focus of the other athletic benefits compliance area is on the *treatment* of female athletes and whether that treatment is equal in effect to the treatment of male athletes.[137] The treatment need not be identical; instead, an institution must provide services and benefits which are "equal or equal in effect."[138]

Title IX permits nondiscriminatory differences in the provision of athletic benefits.[139] Indeed, these differences are necessary to sustain the individual athletic teams offered by the institution. A women's soccer team, for example, need not have the same amount of equipment as a men's hockey team. The nature of the sport may also require a greater or lesser degree of recruiting to fill existing positions.[140] For example, a large football team will often require greater recruiting than any women's teams. Publicity, event management, and competitive facilities also may differ according to individual teams. The Final Policy Interpretation takes into account these team-by-team differences and directs courts and the OCR to consider factors inherent in the basic operation of the sport including rules of play, replacement of equipment, rates of injury, nature of facilities, and others.[141] Although unequal expenditures on necessary services and benefits may be considered in evaluating a school's compliance, unequal aggregate expenditures alone will not constitute non-compliance with Title IX.[142] As long as the differences in the provision of services and benefits are not based on discriminatory reasons, the differences are allowed and should be encouraged.[143]

---

133. *See* Final Policy Interpretation, *supra* note 2, at 71,415-17.

134. INVESTIGATOR'S MANUAL, *supra* note 28, at 29-102.

135. *See id.* at 7 ("OCR will investigate all 13 program components. However, the Policy Interpretation does permit separate investigations and findings for three major areas, specifically: 1) athletic financial assistance - § 106.37(c) 2) accommodation of athletics interests and abilities - § 106.41(c)(1) 3) other athletic benefits and opportunities - § 106.41(c); (this encompasses § 106.41(c)(2)-(10), support services and recruitment).").

136. *Id.*; *see, e.g.*, *Cook*, 802 F. Supp. at 744-46.

137. *See* INVESTIGATOR'S MANUAL, *supra* note 28, at 14.

138. Final Policy Interpretation, *supra* note 2, at 71,415.

139. *Id.*

140. *See id.*

141. *Id.* at 71,415-16.

142. 34 C.F.R. § 106.41(c).

143. Some feminist scholars encourage and celebrate differences between the genders and find the drive for identical treatment as a threat to women's rights. "Women are being asked to settle for some pared-down version of models that men have established. Developing alternatives that are more participatory and less exploitive presents a continuing challenge. The problem, thus reformulated, is how to gain equality without relinquishing difference." Rhode, *supra* note 4, at 1763 (citation omitted). Another scholar questions

Educational institutions, however, often invoke differences between sports to justify discriminatory disparities in treatment between men and women athletes.[144] Indeed, female athletes still seem to find themselves being treated as second-class, even when compared to their male counterparts in the identical sport.[145] One author noted vast disparities between the men's and women's basketball programs at the University of Colorado at Boulder.[146] Specifically, the women's program received 62% of the budget the men's basketball program received.[147] The women's basketball coach's total compensation package was only 38% of the total compensation package of the men's basketball coach.[148] Women received less money for equipment, less than half the men's budget.[149] Moreover, men's training tables, university provided meals, had a budget twenty times greater than women's.[150] Finally, the men's recruiting budget was nearly twice the size of the women's recruiting budget, and growing larger each year.[151] While discrepancies between single sports cannot give rise to a claim under Title IX,[152] the discrepancies tend to bear out in the women's athletics programs as a whole.[153]

The disparities between men's and women's athletic programs are often rationalized by universities and colleges under a recurring theory: revenue generation. Since men's sports, particularly football and basketball, generate revenues, universities argue special treatment is justified for

---

whether the men's athletics program should be the standard of equality for women athletes. Wendy Olson, *Beyond Title IX: Toward an Agenda for Women and Sports in the 1990's*, 3 YALE J.L. & FEMINISM 105, 107 (1990); *see also* Lyn LeMaire, *Women and Athletics: Toward a Physicality Perspective*, 5 HARV. WOMEN'S L.J. 121, 123-27 (1982) (stating that the traditional three-step model of athletics deters women from participating more than it does men); CATHARINE A. MACKINNON, FEMINISM UNMODIFIED 122 (1987) (arguing that women should create a new model of institutional athletics).

144. When the women's softball team sought reinstatement through informal means, Colorado State responded that hard economic times and the public's desire to see men's football and basketball required the university to cut other sports. In commenting on this response to the softball players, Colorado State's Title IX Coordinator wrote these expressed purposes were often used to " 'justify and sustain discrimination.' " Plaintiffs' Trial Brief at 5 n.4., Roberts v. Colorado State Univ., 814 F. Supp. 1507 (D. Colo.) (No. 92-Z-1310), *aff'd in part, rev'd. in part sub nom.* Roberts v. Colorado State Bd. of Agric., 998 F.2d 824 (10th Cir.), *cert. denied*, 114 S. Ct. 580 (1993).

145. *See* George, *supra* note 25, at 562; *see also* Rhode, *supra* note 5, at 1762-63 (comparing the inadequate facilities appropriated for the women's crew, tennis, and field hockey teams at Yale); Wolff, *supra* note 3, at 54-55 (reporting dismal results of NCAA's gender equity study).

146. George, *supra* note 25, at 557-58.

147. *Id.* at 560.

148. *Id.* at 562.

149. *Id.* at 563.

150. *Id.* (noting the university spent $30,000 on training tables for men and only $1,500 on tables for women).

151. *Id.* at 564.

152. *See infra* notes 169-72 and accompanying text.

153. *See* Wolff, *supra* note 3 (discussing the differences in benefits received by men's and women's athletic programs); *see also* Wieberg, *supra* note 78, at C10 (noting that only 12 out of 52 NCAA Division I schools provided women's programs with as much as 30% of the men's budget).

these athletes, particularly in recruiting and expenditures.[154] The revenue generation argument must be rejected for three reasons.

First, the argument itself is a fiction. The true financial situation of most men's athletic teams, including football and basketball, is far from in the black. In 1989, only 13% of the 524 Division I football programs managed to cover their expenses.[155] Indeed, many men's programs lost far more money than they earned, in some instances close to one million dollars per year.[156] Thus, men's athletic programs, like women's programs which the institutions complain produce little or no revenue, are net expenses to most institutions.

Second, and more important for purposes of legal analysis, the Final Policy Interpretation has rejected the revenue generation argument as a valid defense under Title IX.[157] Revenue producing sports are neither exempt from nor provided special treatment under Title IX despite urging from numerous commentators to the contrary.[158] As the authoritative statement on Title IX from the implementing agency, the Final Policy Interpretation's rejection of the revenue generation argument should be followed by courts.

Finally, allowing revenue generation as a defense fundamentally changes the purpose of funding educational institutions with federal dollars. Tax dollars are provided to these institutions not to make a profit but to provide an education. Intercollegiate athletics departments, just like music or theater departments or science labs, are an integral part of the education experience and should not be converted to mini pro-sports franchises. Otherwise, Congress and the Internal Revenue Service may very well resolve to tax those universities that treat athletic programs as a means of generating profit rather than a means of providing and supplementing a federally funded education.[159]

No court has directly addressed the revenue generation argument under Title IX.[160] When presented with the argument, however, courts should reject it. Sports offered by federally funded educational institutions should complement the educational process, not supersede it. Like

---

154. *See* Stanley v. University of S. Cal., 13 F.3d 1313, 1321-23 (9th Cir. 1994) (accepting the revenue generation argument under the Fair Labor Standards Act); *see also* George, *supra* note 25, at 567-70 (mentioning the "market demands" analysis for why women's coaches are paid substantially less than men's coaches); Wolff, *supra* note 3, at 55 (discussing traditional argument that men's teams, which produce revenue, should be excluded from gender equity concerns).

155. Wolff, *supra* note 3, at 55.

156. *Id.* (stating that the University of Colorado's football program lost more than $800,000 in 1990, the year it shared the national championship with Georgia Tech); *see also* George, *supra* note 25, at 567 (noting that the University of Colorado men's basketball program lost over $400,000 in 1992).

157. Final Policy Interpretation, *supra* note 2, at 71,421.

158. *Id.*

159. *See* Wolff, *supra* note 3, at 64 (noting some congressmen, the IRS, and the Federal Trade Commission have considered taxing intercollegiate athletics); *see also* George, *supra* note 25, at 568-69 (rejecting revenue generation argument and stressing importance of the educational aspects of athletics as opposed to financial profitability).

160. *Cf. Stanley*, 13 F.3d at 1320-21 (accepting revenue generation defense under Fair Labor Standards Act).

the *Favia* court, future courts should focus on the educational benefits of interscholastic and intercollegiate competition: the development of self-confidence, physical and mental well-being, and the feeling of team cohesion and accomplishment.[161] Simply because an all-male sport may bring a school additional revenue, prestige, or alumni contribution cannot justify the denial of the educational benefits of athletics to women. Title IX mandates that no person shall be denied benefits of any educational program on the basis of sex.[162] Allowing schools to buy exemption from Title IX for individual male teams makes a mockery of the plain language of Title IX.

To analyze properly compliance with Title IX under the other athletics benefits compliance area, courts must review differences and/or disparities under the criteria set out in the Final Policy Interpretation and the Investigator's Manual without reference to revenue generation. In *Cook v. Colgate University*, the district court performed this analysis and found Colgate in violation of this compliance area.[163] The women's hockey team was provided far fewer benefits than the men's hockey team and, on that basis, the district court found a violation.[164]

Interestingly, the *Cook* court directly compared men's hockey to women's hockey and did not evaluate the athletic programs as a whole.[165] Where the teams are identical, arguably there are no sport differences which justify different treatment.[166] Thus, one could argue, when an educational institution sponsors separate programs for members of each sex in the same sport, courts may apply the team-to-team analysis to determine whether a school meets the requirements of this compliance area. If an institution maintains men's and women's basketball teams, for example, the women's basketball team need only prove a disparity between the treatment of its team and the treatment of the men's basketball team.

Although the team-to-team approach may seem compelling at first, the approach may end up doing more harm than good to women's sports by needlessly hampering an educational institution's choice to foster a particular team or teams. The team-to-team approach forces educational institutions to focus their equal treatment efforts on women's teams which have an identical male counterpart at the institution. Thus, women's basketball and soccer may get elevated in treatment, while women's field hockey or gymnastics get overlooked. Yet, women's teams with male coun-

161. *Favia*, 812 F. Supp. at 583.

162. *See supra* note 6 and accompanying text.

163. *Cook*, 802 F. Supp. at 744-45.

164. *Id.* at 745. In *Cook*, the court stated, "[t]he men's ice hockey players at Colgate are treated like princes. The women ice hockey players are treated like chimney sweeps." *Id.*

165. *See id.* at 744-45. Plaintiffs provided examples in six areas where they contended Colgate discriminated against them, including expenditures, travel, equipment, locker rooms, practice times, and coaching. *Id.*

166. Differences in crowds drawn by either team may support different treatment for event management and security. Nevertheless, no team differences justify disparate treatment with regard to equipment, travel, lockers, team practice, etc. *See* Final Policy Interpretation, *supra* note 2, at 71,416.

terparts may not be teams the institution or the bulk of its student body wishes to foster.

For example, an educational institution may choose to operate and foster a women's gymnastic team (which has no male counterpart) while, at the same time, offering women's and men's basketball. Given the historical success of the gymnastics team accompanied by the relatively high cost of maintaining the sport, the educational institution may choose to devote more funding to maintain the regional or national success of that team rather than divert funds to women's basketball. The school may wish to provide greater recruiting efforts and coaching to women's gymnastics to continue as a "power house" in that particular sport. Moreover, female students may wish to continue the focus on the gymnastics team as demonstrated by the number of interested gymnasts, attendance at meets, and the like.

On the other hand, if women's basketball is less successful and historically has gained less recognition from the institution and its students, the institution should be free to choose to foster the gymnastics team over the women's basketball team. This is true even where men's basketball may be the institution's "power house" counterpart to women's gymnastics. Simply because a school chooses to foster men's basketball over men's hockey or football should not require the school to equally prioritize the women's side of the athletics program.

Nothing in Title IX prevents schools from fostering certain athletic teams; Title IX merely prevents them from providing unequal treatment between the sexes. The team-to-team analysis gives an educational institution the limited option of downsizing the successful men's basketball team, or upgrading the less successful women's basketball team, by taking funds from other teams such as the women's gymnastics team. Under a program-to-program analysis, however, schools need not downgrade successful teams to upgrade less successful women's teams simply because there is a male counterpart team at the institution. As long as the institution is providing benefits to women that are equal, or equal in effect, to the men's athletic program as a whole, a disparity between a single women's athletic team and its male counterpart should not violate this compliance area. Equality should be measured by opportunities and benefits to the group as a whole, not the individual or team.[167]

To state a valid claim under Title IX, athletes must present prima facie evidence that the entire women's program receives unequal treatment in the criteria listed above.[168] In practice, women will be able to present prima facie evidence since most educational institutions are not providing women's programs equal benefits. The Final Policy Interpreta-

167. *See* Note, *Sex Discrimination and Intercollegiate Athletics: Putting Some Muscle on Title IX*, 88 YALE L.J. 1254, 1265 (1979) (discussing why a selection of separate teams for men and women is necessary).

168. *See generally* Wolff, *supra* note 3 (stating that women's sports receive less money for scholarships and programs overall). Moreover, individual teams can request specific treatment even though the violation goes to the program as a whole. *See infra* notes 275-89 and accompanying text.

tion and Investigator's Manual provide detailed guidance to courts in evaluating the criteria in which a plaintiff has alleged a violation.[169] Differences in a single criterion will not necessarily result in a violation of the compliance area. Only where the single difference is substantial, such as wholesale failure to recruit for women's sports, will a single difference result in a disparity or violation.[170] Otherwise, the Final Policy Interpretation and the Investigator's Manual essentially use a "totality of circumstances" approach to assess an institution's compliance.[171] This approach is appropriate because, unlike the other compliance areas, the other athletic benefits compliance area requires evaluation of eleven different criteria. Furthermore, this totality of circumstances evaluation allows courts to balance shortfalls in one criterion against adequate services and benefits in other areas.[172]

The totality of circumstances approach is a workable solution to conserving scarce educational resources. This approach allows an institution to balance the relatively minor provision of equipment to the women's soccer team with other provisions like higher recruiting or newer uniforms in women's soccer or more equipment in women's gymnastics. Moreover, when applied on a program-to-program basis, the totality approach allows an educational institution to retain flexibility in choosing which sports it chooses to foster. The goal behind this compliance area is to look for overall balance and an attempt to spread scarce resources equally, not identically, between the sexes. A totality of circumstances approach helps achieve that goal.

3. Interests and Abilities Compliance Area

The third and final compliance area is the effective accommodation of athletic interests and abilities (interests and abilities compliance area). The Regulations require educational institutions to ensure their "selection of sports and levels of competition effectively accommodate the interests and abilities of members of both sexes."[173] Although the Regulations combine this compliance area with nine other criteria, the Final Policy Interpretation, Investigator's Manual, and case law make clear this criterion is to be assessed as a wholly separate compliance area.[174] Failure to provide full and effective accommodation of women's athletic interests and abilities can create a separate and distinct violation of Title IX.[175]

---

169. Final Policy Interpretation, *supra* note 2, at 71,415-17; INVESTIGATOR'S MANUAL, *supra* note 28, at 29-96.

170. Final Policy Interpretation, *supra* note 2, at 71,415.

171. *Id.* at 71,417; *see also* INVESTIGATOR'S MANUAL, *supra* note 28, at 7-8.

172. *See* INVESTIGATOR'S MANUAL, *supra* note 28, at 8.

173. 34 C.F.R. § 106.41(c)(1).

174. Final Policy Interpretation, *supra* note 2, at 71,417; INVESTIGATOR'S MANUAL, *supra* note 28, at 7, 21; *see, e.g., Roberts*, 814 F. Supp. at 1510-11.

175. *See supra* notes 103-04 and accompanying text.

The interests and abilities compliance area is the cornerstone of Title IX as it applies to athletics.[176] "There is little question that this factor [of effective accommodation of interests and abilities] is the most important criteria listed in section 106.41(c)."[177] Without this compliance area, the other compliance areas would be of little meaning. A simple example demonstrates the point. Assume an educational institution has 48% men and 52% women in its undergraduate population. The educational institution's athletic program, however, is 70% male and 30% female. Assume further the female athletes each receive 100% scholarship (while their male counterparts receive a range of 0% to 100% scholarships). The female athletes also have brand new showers, brand new equipment, and first class travel. Under the analyses above, the educational institution is in full compliance with the financial assistance and other athletic benefits compliance areas. Nevertheless, the educational institution will probably not comply with the interests and abilities compliance area, as discussed below.[178] Without the interests and abilities compliance area, upon which the other two compliance areas depend, educational institutions can feign gender equity by treating a few individual female athletes on a level equal to or better than the male athletes. Indeed, educational institutions would find it far less troublesome and less expensive to treat a single individual, or small group of individuals, as "equals" than to bring true gender equity to an entire athletic department.[179]

The Final Policy Interpretation states, "it is the achievement of 'equal opportunity' for which recipients are responsible and to which the final Policy Interpretation is addressed."[180] The goal behind Title IX and the Final Policy Interpretation is to see that educational institutions provide equal athletic opportunities to women as an entire group, not just a few select individuals hand-picked by educational institutions.[181] It would defeat the purpose of Title IX to allow institutions to doctor the symptoms of

176. The interests and abilities compliance area delineates the "heartland" of Title IX. *Cohen*, 991 F.2d at 897; *see also* Heckman, *supra* note 11, at 45 (stating that the interests and abilities factor is the most significant of the 10 criteria).

177. *Cohen*, 809 F. Supp. at 989.

178. At these enrollment participation number ratios (a 22% point differential), the educational institution unquestionably fails the "substantially proportionate" prong of the interests and abilities compliance area. *See infra* part IV.A.3.a. For purposes of this example, it is assumed the educational institution cannot show a history and continuing practice of program expansion or full accommodation of interests and abilities as these requirements are discussed below. *See infra* parts IV.A.3.b, IV.A.3.c. Courts, however, should not find a violation of Title IX based *solely* on lack of substantial proportionality between female athletic participation rates and female undergraduate enrollment. Instead, courts must also look to the other two prongs of the interests and abilities compliance area to determine whether there is unlawful discrimination. *Cohen*, 991 F.2d at 895.

179. "Satisfaction of [the compliance area] factor is the only means by which greater athletic opportunities for women and girls may be established. If women are afforded greater athletic opportunities, by necessity more equipment, supplies, facilities, scheduling, coaches, and scholarships will be required to insure equal athletic opportunity." Heckman, *supra* note 11, at 45.

180. Final Policy Interpretation, *supra* note 2, at 71,414.

181. *Cohen*, 991 F.2d at 897. "[A]n institution that offers women a smaller number of athletic opportunities than the statute requires may not rectify that violation simply by lavishing more resources on those women or achieving equivalence in other respects." *Id.*

discrimination—less money and fewer benefits for women's sports—without combating the source of the discrimination: lack of equal opportunity to participate in athletics. Accordingly, this compliance area should be of primary concern to the courts. Only where the playing field is level through equal athletic opportunity can courts effectively evaluate compliance in the other two compliance areas.

The Final Policy Interpretation divides assessment of the interests and abilities compliance area into three categories: (1) "the determination of athletic interests and abilities of students;" (2) "the selection of sports offered;" and (3) "the levels of competition available including the opportunity for team competition."[182] Determination of interests and abilities[183] and selection of sports,[184] while important, are secondary to the levels of competition category; the Investigator's Manual includes the two former categories in its discussion of the levels of competition category.[185]

The levels of competition category of the interests and abilities compliance area is intended to assess whether the institution provides overall equal opportunities to men and women to participate and compete in athletics. Institutions must provide equal *opportunities* to participate (called opportunities to compete) and equal *levels* of team competition (called competitive team schedules).[186] Competitive team schedules refers primarily to the level of play at which the athletes compete.[187]

The opportunity to compete subcategory is the real focus of the interests and abilities compliance area and has been the focus of the recent Title IX litigation in athletics. This subcategory addresses the question whether women and men are afforded the same athletic opportunities.[188]

---

182. Final Policy Interpretation, *supra* note 2, at 71,417. A schematic diagram of Title IX is set out in the Appendix to this article.

183. The determination of interests and abilities discusses the manner in which an educational institution may assess the athletic interests and abilities of its underrepresented sex. *Id.* The institution may choose any method of assessing interests and abilities, including surveys or sections within the enrollment application, so long as the method, among other things, does not discriminate against the underrepresented sex. *Id.*

184. The selection of sports category discusses the institution's responsibilities when it offers a sport for one gender that it does not offer for the other and distinctions are made for contact and non-contact sports. *Id.* at 71,417-18. For an analysis of this category, see Heckman, *supra* note 11, at 45-62.

185. Determining interests and abilities is discussed under the third prong of the interests and abilities compliance area's competitive opportunities test set out below. *See* INVESTIGATOR'S MANUAL, *supra* note 28, at 25. Selection of sports is discussed in the "Cautions" section of the interests and abilities compliance area. *Id.* at 26-27. The Investigator's Manual indicates that a survey or assessment of interests and abilities is not required by the Title IX regulations or the Final Policy Interpretation. *Id.* at 27.

186. Final Policy Interpretation, *supra* note 2, at 71,418. A violation of either the opportunity to compete or competitive team schedules subcategories gives rise to separate violations of Title IX. *Cohen*, 809 F. Supp. at 990-91. A violation occurs when requirements of one or the other, or both subcategories are not met. *Id.* at 991.

187. *See* Final Policy Interpretation, *supra* note 2, at 71,418; *see also* INVESTIGATOR'S MANUAL, *supra* note 28, at 21 (noting "levels of competition" refers to the quality of competition indicated by whether a team competes against other teams at the same division level).

188. *See Roberts*, 814 F. Supp at 1510-11; *Favia*, 812 F. Supp. at 584; *Cohen*, 809 F. Supp. at 989-90.

To determine whether an educational institution affords equal opportunities, courts must apply a three-pronged test:[189]

> (1) Whether the intercollegiate or interscholastic participation opportunities for male and female students are substantially proportionate to their respective enrollments; or
> (2) Where the members of one sex are underrepresented among athletes, whether the institution can show a history and continuing practice of program expansion which is demonstrably responsive to the developing interests and abilities of that sex; or
> (3) Where the members of one sex are underrepresented among athletes, and the institution cannot show a continuing practice of program expansion, whether it can be demonstrated that the interests and abilities of the underrepresented sex have been fully and effectively accommodated.[190]

This section of the article analyzes these three prongs and discusses how the federal courts have applied, and should apply, the test.

a. *Substantially Proportionate*

Recent Title IX litigation has repeatedly addressed the definition and breadth of the "substantially proportionate" prong of the interests and abilities compliance area.[191] The Final Policy Interpretation does not define the term "substantially proportionate,"[192] and the Investigator's Manual provides no additional assistance: "There is no set ratio that constitutes 'substantially proportionate' or that, when not met, results in a disparity or a violation. All factors for this program component [interests and abilities] and any justifications for differences offered by the institution, must be considered before a finding is made."[193] How, then, do courts determine whether an institution's athletic participation rates are "substantially proportionate" to its undergraduate enrollment rates? Recent cases provide some guidance.

In *Roberts v. Colorado State University*, the district court held a 10.5% differential between women's enrollment and women's athletic participation was not in compliance with the "substantially proportionate" requirement of the first prong.[194] The *Roberts* district and circuit courts further agreed that the OCR had found Colorado State in noncompliance with this prong when its differential had been 7.5%, 12.5%, and 12.7%.[195] In *Favia v. Indiana University of Pennsylvania*, the district court found a 17.84% differential not substantially proportionate.[196] Moreover, in *Cohen v. Brown University*, the district court found an 11.6% differential not substan-

189. Final Policy Interpretation, *supra* note 2, at 71,418.
190. *Id.*
191. *See Roberts*, 814 F. Supp. at 1511-13;*Favia*, 812 F. Supp. at 584-85; *Cohen*, 809 F. Supp. at 991.
192. Final Policy Interpretation, *supra* note 2, at 71,418.
193. INVESTIGATOR'S MANUAL, *supra* note 28, at 24.
194. 814 F. Supp. at 1512-13. The Tenth Circuit Court of Appeals upheld this finding. *Roberts*, 998 F.2d at 829-30.
195. *Roberts*, 998 F.2d at 830; *Roberts*, 814 F. Supp. at 1513.
196. 812 F. Supp. at 584-85.

tially proportionate.[197] The smallest differential recognized by a court as violating the substantially proportionate prong was the 7.5% differential noted by the *Roberts* courts where the OCR found Colorado State in noncompliance.[198] Nevertheless, none of these cases provided a method for other courts to analyze the "substantially proportionate" standard in the future.[199]

In developing a workable application of the "substantially proportionate" standard, courts must understand the purpose for using this apparently amorphous benchmark. Neither the Final Policy Interpretation nor the Investigator's Manual attempts to set a rigid, bright-line standard that distinguishes compliance from noncompliance. To do so would require an athletic institution to add or subtract individual athletes every year as undergraduate enrollment rates fluctuated within normal matriculation parameters or recruitment efforts. A rigid rule would be unable to take into account an unexpected dearth in male enrollment or the equally unexpected decision of a top female athlete to accept a scholarship from another school. Accordingly, the Final Policy Interpretation and the Investigator's Manual require "substantial proportionality," a standard which ideally aims to keep the enrollment rates and participation rates in exact proportion,[200] but which can accommodate year-to-year *fluctuations* in the enrollment and athletic participation. Courts should not undermine this purpose by imposing a bright-line standard for compliance.

In practice, educational institutions have used bright-line standards under the "substantially proportionate" analysis to justify continued discrimination.[201] To those educational institutions, "substantially proportionate" has become a glass ceiling with which they attempt to justify their deliberate limitation of athletic opportunities for women. Around the country, institutions have come to set their sights on the 40% female athletic participation demarcation. For example, the Big Ten Athletic Conference set the 40% participation mark as a requirement for its institutional members.[202] While this is a step in the right direction, since few schools are anywhere near the 40% mark, the Big Ten's arbitrary selection of 40% female athletic participation rate actually continues discrimination and ensures women will remain second class athletes.[203] Institutional members of the Big Ten will strive for this 40% demarcation regardless of their female undergraduate enrollment. An institution with

197. 809 F. Supp. at 991.

198. *Roberts*, 998 F.2d at 830; *Roberts*, 814 F. Supp. at 1513.

199. *See, e.g., Roberts*, 998 F.2d at 830 (stating the Tenth Circuit would not further demarcate the line between substantial proportionality and disproportionality other than to agree a 10.5% differential was not substantially proportionate).

200. "[I]f the enrollment is 52% male and 48% female, then, ideally, about 52% of the participants in the athletics program should be male and 48% female . . . ." INVESTIGATOR'S MANUAL, *supra* note 28, at 24.

201. *See* Kelley v. Board of Trustees of Univ. of Ill., 832 F. Supp. 237, 240 (C.D. Ill. 1993).

202. *See id.*

203. *See* Wolff, *supra* note 3, at 54-55; *see also* Wieberg, *supra* note 78 (noting that only 12 of 86 Division I-A schools provide women with as much as a 30% share of total athletic funding).

40% female undergraduate enrollment and one with 55% undergraduate enrollment each strive for the same goal of 40% women athletic participation. Instead of reaching for "substantial proportionality," the educational institutions reach for the arbitrarily determined and, in many cases, inexplicably low standard of 40% female athletic participation.[204]

The "substantially proportionate" prong should not be applied to create a glass ceiling through which female athletes will never pass, nor one at which educational institutions will be found in compliance even though their female enrollment rates are substantially higher than an arbitrarily determined participation rate. To ensure true equality between the sexes, this prong must not be applied like a pre-set quota as urged by some educational institutions.[205] Instead, the method of analyzing "substantial proportionality" must reflect the purpose of the standard to bring about near-exact parity while allowing for year-to-year fluctuations.

In future Title IX athletics litigation, the "substantially proportionate" prong should be applied by using evidentiary presumptions. Under this method of analysis, plaintiffs have the initial burden of proving the rates of undergraduate enrollment and athletic participation are not *exactly* proportionate. Plaintiffs could meet this burden by producing evidence of the number of female undergraduate enrollees as compared with the number of female athletes. Once disproportionality is established, a rebuttable evidentiary presumption should operate in the plaintiffs' favor that the participation rates are not "substantially proportionate" within the meaning of the Final Policy Interpretation.[206]

If the educational institution chooses to challenge the "substantially proportionate" prong of the interests and abilities compliance area, it may produce evidence to rebut the presumption now operating in plaintiffs' favor. To rebut the presumption, the educational institution need merely produce evidence that the disparity between the enrollment rates and participation rates are due to enrollment fluctuations or unexpected transfers of student athletes. Under this analysis, however, an educational institution cannot claim compliance with this prong by demonstrating its differential has always been less than 10% year after year. To allow such evidence to justify participation-enrollment rate disparities would only promote the glass ceiling defense discussed above. If a court determined that any differential below 7% is substantially proportionate, educational institutions will strive to hit the 7% differential and nothing more. Schools will have no greater incentive to raise women's athletic participation higher than the number arbitrarily determined by the courts and will continue to provide proportionately more athletic opportunities to men

204. The Big Ten's ultimate goal of 60/40 male to female participation ratio does not preempt compliance with Title IX's requirements of equal opportunity for male and female athletes because a school's "actions must be measured by Title IX's requirements." *Kelley*, 832 F. Supp. at 242 n.5. "The obligation to comply with [Title IX] is not obviated or alleviated by any rule or regulation of any . . . athletic or other league, or association . . . ." 34 C.F.R. § 106.6(c); *see also* Final Policy Interpretation, *supra* note 2, at 71,422.

205. *See Kelley*, 832 F. Supp. at 240.

206. Final Policy Interpretation, *supra* note 2, at 71,418.

simply because the schools have historically chosen to do so and prefer to continue on that course. Schools will never treat women as equals in athletics; rather, women will always get only as much as the courts require the schools to give.

True nondiscriminatory fluctuations may only be proven by demonstrating unexpected increases or decreases in matriculation for a certain gender, changes in participation rates due to the graduation of a large number of senior athletes, or similar fluctuations. Unless proof of these nondiscriminatory reasons for differences between the enrollment rates and athletic participation rates can be found, educational institutions should be unable to rebut the evidentiary presumption that athletic opportunities are not substantially proportionate.

If, however, an educational institution produces evidence of nondiscriminatory reasons such as unexpected matriculation fluctuations, plaintiffs carry the ultimate burden of proving the participation rates are not substantially proportionate to enrollment rates.[207] Plaintiffs may present statistical evidence that demonstrates the differences between the enrollment and participation rates are statistically significant and demonstrate a persistent pattern over a period of time. Statistical evidence is useful to demonstrate that disparities between enrollment and participation rates are not the result of mere chance or year-to-year fluctuations. Plaintiffs may also introduce evidence that demonstrates women's participation rates have never exceeded their enrollment rates.[208] Arguably, true year-to-year fluctuations would cause the female participation rates to sometimes *exceed* female enrollment rates. Plaintiffs may argue that the educational institution's assertion of nondiscriminatory factors are unjustified, fabricated, or overstated. Of course, the ultimate burden lies with plaintiffs to persuade the trier of fact that the enrollment-participation rates are not substantially proportionate.[209]

The use of presumptions is a fair and reasonable way to apply the "substantially proportionate" prong. The purpose of Title IX is to make athletic opportunities *equal.* It is therefore logical to presume that where the participation rates are not *equal* to the enrollment rates, the educational institution is unfairly favoring one gender. The presumption is easily overcome where the institution shows the differential is not a result of unfair favoritism, but a result of unanticipated changes in enrollment or athletic participation. Furthermore, educational institutions are protected from undue hardship by the fact that plaintiffs bear the ultimate burden

207. Plaintiffs bear the burden of proof on the substantially proportionate prong of the interests and abilities compliance area. *Roberts*, 998 F.2d at 829 n.5; *Favia*, 812 F. Supp. at 584.

208. No co-educational post-secondary institution reported in a 1992 *USA Today* study has female participation rates which exceed female enrollment rates. Wieberg, *supra* note 78, at C10. This demonstrates that educational institutions are not prevented from "substantially proportionate" rates by virtue of mere fluctuations. Instead, schools have made conscious choices to continue to promote their men's athletics programs at the expense of their women's programs.

209. *See supra* note 207.

of proof on this prong,[210] and defendants may still find safe harbor in the two additional prongs of this compliance area: showing a history and continuing practice of program expansion and full and effective accommodation of the interests and abilities.[211]

The Investigator's Manual supports this evidentiary presumptions approach, noting that the ideal result is exact parity between enrollment rates and athletic participation rates.[212] Thus, exact parity is the logical starting place for analyzing whether an institution meets the "substantially proportionate" prong. Nevertheless, the Investigator's Manual directs the OCR to consider "any justifications for differences offered by the institution" before making a finding under this prong.[213] If a plaintiff can demonstrate, however, that the asserted "justifications" are less than genuine or are also discriminatory, the justifications should not be used to validate discrimination.

The desired effect of this evidentiary presumption approach is to stop the existing search for the magical number at which an educational institution will be in compliance with this prong. There should be no pre-set athletic participation number (such as 40%) or set differential between enrollment and athletic participation (such as 10% or less) that establishes compliance. Instead, the courts should focus on the manner in which an institution allocates its limited athletic opportunities to men and women. The manner of distribution of athletic opportunities and whether the distribution is a result of an existing preference for male sports underlies the evidentiary presumption approach. Whether an educational institution complies with this first prong will always remain a question for the trier of fact. The presumption, however, will ensure that plaintiffs may effectively challenge an educational institution's practice of keeping female athletes close but never quite equal. Courts also must bear in mind that a violation of this first prong does not end the inquiry under the interests and abilities compliance area.

b. *History and Continuing Practice of Program Expansion*

Where an institution's enrollment rates are not "substantially proportionate" to its athletic participation rates, the analysis moves to the second prong: whether an educational institution can show a "history and continuing practice of program expansion which is demonstrably responsive to the developing interest and abilities of the members of [the underrepresented] sex."[214] This second prong of the opportunity to compete analysis allows educational institutions to come into compliance with Title IX by demonstrating continued progress toward gender equity in their athletic departments. Where, for example, an institution became co-edu-

210. *See supra* note 207.
211. *See* discussion *infra* parts IV.A.3.b., IV.A.3.c.
212. INVESTIGATOR'S MANUAL, *supra* note 28, at 24.
213. *Id.*
214. *See* Final Policy Interpretation, *supra* note 2, at 71,418; INVESTIGATOR'S MANUAL, *supra* note 28, at 21.

cational within the last decade, the institution's gradual, but continuous, expansion of the newly-admitted gender's athletic opportunities satisfies this prong. Thus, the educational institutions are given the opportunity to reach gender equity at a reasonable pace and are not required to achieve full equality overnight.

Like the "substantially proportionate" prong, however, this prong has been used by long-time co-educational institutions to feign compliance with Title IX.[215] When Congress passed Title IX in 1972, many post-secondary educational institutions offered no athletic opportunities to their female students. Since that time, institutions have expanded women's sports in reaction to Title IX's passage. For example, Colorado State offered no women's sports prior to 1970 and added eleven women's sports during the 1970s in reaction to the mandates of Title IX.[216] Thus, Colorado State expanded its program from 0% female athletic participation in the pre-1970 period to approximately 37.7% female athletic participation in the 1992-93 athletic year.[217] Allowing an educational institution to claim it has a history and continuing practice of program expansion simply because it added sports in the 1970s, however, undermines Title IX's purpose and renders the interests and abilities compliance area meaningless. As Judge Weinshienk succinctly stated in *Roberts*, "[a]cceptance of this argument also would implicitly condone the attitude that female athletes . . . should be satisfied with their current opportunities given the pre-1970 lack of participation opportunities for women in intercollegiate athletics."[218] The history of expansion prong was never intended to allow educational institutions to get off scot-free for present discrimination simply because they can demonstrate they added a few women's teams within the last twenty years.

In analyzing history of program expansion, courts should focus on the "continuing practice" language of this prong, since addition of teams at the inception of women's athletics cannot provide a defense under Title IX.[219] The district court in *Roberts* suggested the relevant period for determining program expansion is the era in which Colorado State was put on notice by the OCR that its participation rates were not substantially proportionate to its enrollment rates.[220] This analysis, although relevant for Colorado State, should not be applied as a standard. Some institutions may not have the record of OCR investigations that Colorado State has.[221] Furthermore, notice to an educational institution is not an element of a

215. *See Roberts*, 814 F. Supp. at 1514.

216. Defendants in *Roberts* used as a defense to plaintiffs' Title IX claim the fact that CSU added 11 women's sports in the 1970s. *Id.*

217. *Id.* at 1512, 1514.

218. *Id.* at 1514.

219. It is important to note that a court would have to look at the entire history of the women's athletic program if the institution had only recently begun accepting women into its undergraduate enrollment and providing women athletic opportunities.

220. *Roberts*, 814 F. Supp. at 1515.

221. The OCR began to evaluate Colorado State for Title IX violations in 1983. *Id.* at 1512-13.

violation of this compliance area.[222] Courts should not limit their inquiry of program expansion to a period wherein the school was put on notice of its potential Title IX violation. The sole inquiry under this prong should focus on whether the educational institutions can demonstrate a *continuing practice* of program expansion that responds to the developing athletic interests of women.[223]

In evaluating the history and continuing practice of program expansion, courts also must look at the number of women's athletic teams recently added and the increase in number of participation opportunities.[224] If the institution shows a continuing practice of program expansion for the underrepresented gender, the educational institution will be found in compliance with this prong, and the interests and abilities compliance area analysis ends.[225]

Faced with economic cuts, however, most schools have been contracting athletic programs for both men and women. Contraction of both sexes' athletic programs, with a smaller contraction of women's programs, cannot satisfy this second prong. The "expansion by contraction" argument was raised by Colorado State in the *Roberts* case, to no avail. Colorado State claimed that since it had increased its participation numbers by virtue of contracting its women's and men's athletic programs (eliminating 18 softball players and 55 baseball players), it showed a history of program expansion.[226] The Tenth Circuit explicitly rejected this argument.

> We recognize that in times of economic hardship, few schools will be able to satisfy Title IX's effective accommodation requirement by continuing to expand their women's athletics programs. Nonetheless, the ordinary meaning of the word "expansion" may not be twisted to find compliance under this prong when schools have increased the relative percentages of women participating in athletics by making cuts in both men's and women's sports programs. Financially strapped institutions may still comply with Title IX by cutting athletic programs such that men's and women's athletic participation rates become substantially proportionate to their representation in the undergraduate population.[227]

The burden of proof for program expansion is placed on the educational institution.[228]

---

222. Although Congress directed that an educational institution must be notified of its violation before federal funds are withdrawn, Congress, the Regulations and the Final Policy Interpretation never directed that a plaintiff must prove an educational institution was on notice in order to prove a violation. *See, e.g.*, 20 U.S.C. § 1682 (1988).

223. *See* Final Policy Interpretation, *supra* note 2, at 71,418; INVESTIGATOR'S MANUAL, *supra* note 28, at 21.

224. *See Roberts*, 814 F. Supp. at 1514; INVESTIGATOR'S MANUAL, *supra* note 28, at 24-25.

225. *See* INVESTIGATOR'S MANUAL, *supra* note 28, at 25.

226. *See Roberts*, 814 F. Supp. at 1514.

227. *Roberts*, 998 F.2d at 830.

228. *See* Final Policy Interpretation, *supra* note 2, at 71,418 (stating the necessary inquiry as "[w]hether *the institution can show* a history and continuing practice of program expansion") (emphasis added); INVESTIGATOR'S MANUAL, *supra* note 28, at 21 (same); *Roberts*, 998 F.2d at 830 n.8 (noting that the language of Policy Interpretation places burden of proof on the institution).

c. *Full and Effective Accommodation of Interests and Abilities*

When athletic participation and undergraduate enrollments are not substantially proportionate and where the institution cannot show a history and continuing practice of program expansion, the courts must analyze the third and final prong of the interests and abilities compliance area's opportunity to compete test: "whether it can be demonstrated that the interests and abilities of the members of that sex have been fully and effectively accommodated by the present program."[229]

This prong allows institutions to provide greater athletic opportunities to one gender where the other gender simply does not wish or is not able to compete. For example, a community college which primarily serves older, part-time students may meet this prong. The school may offer no scholarships or intercollegiate athletics. It may, however, offer intramural athletics for its students. The school may demonstrate that it has no (or very few) female participants because women have shown no interest in competing in the intramural program (because of jobs, parenting, etc.). In this instance, the institution would satisfy the third prong and would not be in violation of the interests and abilities compliance area.

To evaluate this prong, courts must consider a number of factors: any surveys conducted by the institution demonstrating athletic interest or ability in the underrepresented gender; "expressed interests" of the underrepresented gender; club and intramural participation by the underrepresented gender; and, participation levels in feeder schools, community programs, or physical education classes.[230] Moreover, the manner in which the educational institution met the developing interests of men, and whether the institution actively encouraged any male sport, also must be considered.[231] If an institution has cut an existing and viable women's team, this prong will be easily satisfied by testimony by the plaintiff athletes regarding their interests and abilities in continuing play on their former team.[232]

Educational institutions have argued that they need accommodate women only to the extent they accommodate men. Thus, the argument goes, if male interest in baseball is unmet, the school may ignore women's interest in softball.[233] The First and Tenth Circuits have specifically rejected this argument.[234]

Even if unmet interests and abilities are demonstrated by the underrepresented gender, an institution need not upgrade or create a new sport

229. *See* Final Policy Interpretation, *supra* note 2, at 71,418; INVESTIGATOR'S MANUAL, *supra* note 28, at 21.

230. INVESTIGATOR'S MANUAL, *supra* note 28, at 25.

231. *Id.*

232. *See Roberts*, 998 F.2d at 831-32.

233. *Id.* at 831.

234. *See, e.g., Cohen*, 991 F.2d at 898 ("[T]his benchmark sets a high standard: it demands not merely some accommodation, but full and effective accommodation. If there is sufficient interest and ability among members of the statistically underrepresented gender, not slaked by existing programs, an institution necessarily fails this prong of the [interests and abilities compliance area] test.").

where there is no reasonable expectation of competition for that sport within the institution's normal competitive region.[235] If, for example, a sufficient number of women at the University of Florida wish to institute a cross-country ski team, the institution need not accommodate that interest if cross-country skiing is not a sport available within the University of Florida's normal competitive region. Institutions, however, may be required to actively encourage the development of such opportunities within the competitive region where overall athletic opportunities have been historically limited for the members of one sex.[236] So, for example, the entire Pac Ten Conference could not decide to drop all women's volleyball, thus providing themselves the excuse that there is no opportunity for women's competition within the Pac Ten competitive region. Moreover, the historical discriminating practices of a group of educational institutions is not a sufficient reason to justify and continue discrimination. Otherwise, the old boys' network would have little incentive to change.

Plaintiffs carry the burden of proof for the third prong of the interests and abilities compliance area.[237] Although some district courts have held otherwise, circuit courts have held the burden is on plaintiffs since proof of disproportionality alone will not state a claim under Title IX.[238] Thus, plaintiffs state a prima facie claim under the interests and abilities compliance area by alleging and proving (1) the rates of female undergraduate enrollment and the rates of female athletic participation are not substantially proportionate, and (2) there are unmet athletic interests and abilities in women attending the educational institution. A defendant institution can avoid liability under this compliance area by proving a history and continuing practice of athletic program expansion for the underrepresented gender.[239]

4. Discriminatory Intent: Not a Required Element

To state a claim under Title IX, a plaintiff need not prove discriminatory intent.[240] In construing Title VI, the statute on which Title IX was modeled,[241] a majority of the Supreme Court held that a violation of the

235. *See Roberts*, 998 F.2d at 831; *see also* Final Policy Interpretation, *supra* note 2, at 71,418; INVESTIGATOR'S MANUAL, *supra* note 28, at 27.

236. *See Cook*, 802 F. Supp. at 746-47 (rejecting as defenses defendant's claims that women's ice hockey was not sponsored by NCAA, that region had insufficient competition, and that high school women's ice hockey provided inadequate recruitment base for Colgate to develop a women's team).

237. *See, e.g., Roberts*, 998 F.2d at 831 (stating an institution would be hard pressed to prove full and effective accommodation in the abstract).

238. *Id.*

239. *See supra* notes 214-28 and accompanying text.

240. *Haffer v. Temple Univ.*, 678 F. Supp. 517, 539-40 (1987).

241. Title IX was patterned after Title VI, Civil Rights Act of 1964, Pub. L. 88-352, 78 Stat. 252 (codified at 42 U.S.C. § 2000d (1988)), and should be construed similarly. *Grove City*, 465 U.S. at 566; *Cannon*, 441 U.S. at 694. Some courts have held, however, that Title VII, not Title VI, is "the most appropriate analogue when defining Title IX's substantive standards, including the question of whether 'disparate impact' is sufficient to establish discrimination under Title IX." *Roberts*, 998 F.2d at 832-33 (quoting Mabry v. State Bd. of Community Colleges & Occupational Educ., 813 F.2d 311, 316 n.6 (10th Cir.), *cert. denied*, 484 U.S. 849 (1987)).

statute requires proof of discriminatory intent.[242] Nevertheless, a different majority agreed proof of discriminatory effect was all that was required to state a claim under Title VI's *implementing regulations.*[243] The first court to reach this issue in Title IX litigation followed this Title VI analysis and ruled the plaintiffs did not need to prove discriminatory intent on their Title IX claim because they had alleged a violation of the statute and the Regulations.[244] Moreover, the recent athletics cases also hold discriminatory intent is not required to state a claim under Title IX.[245] Proof of intent, however, may have an impact on the remedies available to a plaintiff, as discussed below.[246]

B. *Affirmative Defenses to a Title IX Athletics Case*

Once a plaintiff establishes a prima facie case of violation under one of the three compliance areas, the burden shifts to the defendant educational institution to establish affirmative defenses. This section discusses what legitimate defenses allow an educational institution to maintain an athletic program that violates one of the three compliance areas.

Nondiscriminatory reasons may be used to justify overall disparities in all compliance areas except the interests and abilities compliance area.[247] In the interests and abilities compliance area, the only defense for a prima facie violation is proof of a history and continuing practice of program expansion for the underrepresented gender.[248]

---

242. Guardians Ass'n v. Civil Serv. Comm'n, 463 U.S. 582, 584 (1983); *see also Haffer*, 678 F. Supp. at 539.

243. *Guardians Ass'n*, 463 U.S. at 584 n.2.

244. *Haffer*, 678 F. Supp. at 539-40 (stating Title IX regulations do not explicitly impose an intent requirement).

245. *See Roberts*, 998 F.2d at 832-33; *Roberts*, 814 F. Supp. at 1518; *Favia*, 812 F. Supp. at 584.

246. *See* discussion *infra* part V.B.

247. *See* Final Policy Interpretation, *supra* note 2, at 71,415. The financial assistance compliance area states: "[i]nstitutions may be found in compliance if . . . a resulting disparity can be explained by adjustments to take into account legitimate, nondiscriminatory factors," and the other athletic benefits compliance area states: "a finding of compliance may still be justified if the differences are the result of nondiscriminatory factors." *Id.*

248. *Id.* at 71,418 (setting out a three-pronged test for this compliance area, making the second prong (whether institutions can show history and continuing practice of program expansion) a defense to the failure to provide athletic opportunities substantially proportionate to enrollment rates when interests are not fully and effectively accommodated). The interests and abilities compliance area is the only compliance area in which courts are not allowed to consider legitimate, nondiscriminatory reasons in assessing compliance.

The Investigator's Manual appears to contradict the Final Policy Interpretation on this issue by directing the OCR to consider nondiscriminatory reasons for disparities in all compliance areas. INVESTIGATOR'S MANUAL, *supra* note 28, at 3-4. Nevertheless, a closer reading of the Investigator's Manual demonstrates support for the proposition that the only defense to a violation of the interests and abilities compliance area is the second prong in that three-pronged test; the Investigator's Manual makes no mention of nondiscriminatory justifications that may serve as defenses. *See id.* at 24-27. Instead, the Manual discusses justifications for differences in determining whether the rates are substantially proportionate, the first prong of the test. For a discussion of this first prong, see *supra* part IV.A.3.a. (concerning burden shifting presumptions in establishing whether participation and enrollment rates are substantially proportionate). Nondiscriminatory year-to-year fluctuations are not affirmative defenses to a violation of the overall compliance area, but are evidence to consider in

As for affirmative defenses in the other two compliance areas, the Final Policy Interpretation and Investigator's Manual make repeated references to the use of legitimate, nondiscriminatory reasons by an institution to justify a disparity in one of these compliance areas. Nevertheless, neither the Final Policy Interpretation nor the Investigator's Manual is particularly good at cataloging these nondiscriminatory reasons. Some alleged nondiscriminatory reasons have been specifically allowed, while others have been specifically rejected. This article makes no effort to discuss every possible nondiscriminatory reason an educational institution may raise. Instead, this article discusses below some nondiscriminatory reasons presented in the Final Policy Interpretation and some defenses previously raised by educational institutions. In examining defenses raised by educational institutions, courts should carefully examine the alleged nondiscriminatory reason to ensure the reason is not a result of other discrimination.[249]

1. Team-by-team Differences: Allowed

Differences in teams, such as team-specific rules and regulations, are allowed as an affirmative defense.[250] For example, if one team requires more equipment or needs greater security for stadium events, an educational institution will not violate Title IX by reacting to the specific needs of the particular team.

2. Temporary Annual Fluctuations: Allowed

Temporary annual fluctuations in the other athletics benefits area are specifically allowed as an affirmative defense.[251] This prevents female athletes from suing simply because the male athletes got more recruiting dollars or other benefits in any one academic year.[252] These annual fluctuations are permitted to provide flexibility to the educational institution without exposing it to Title IX liability.

3. Voluntary Affirmative Action: Allowed

Voluntary affirmative action initiated by the educational institution to correct historical conditions that have limited participation in athletics by

---

determining "substantial proportionality" under the first prong of the interests and abilities compliance area test. *Id.*

249. For example, the Final Policy Interpretation allows as a defense to a violation of the financial assistance compliance area that schools have more out-of-state male athletes, thus requiring greater amounts of financial assistance. Final Policy Interpretation, *supra* note 2, at 71,415. Looking behind that reason, however, the Investigator's Manual points out that schools had more out-of-state male student athletes because the schools recruited more for men than for women. INVESTIGATOR'S MANUAL, *supra* note 28, at 20. This discriminatory recruiting practice cannot be upheld as a legitimate nondiscriminatory reason for failure to provide appropriate amounts of financial assistance to women athletes.

250. *See* Final Policy Interpretation, *supra* note 2, at 71,415-16.

251. *Id.*

252. If greater benefits are afforded to men on a regular basis, as opposed to a single academic year, the educational institution may be in violation of the other athletic benefits compliance area.

the members of one sex is an affirmative defense to a Title IX violation.[253] Thus, if an educational institution maintained an athletic department which had men as the overrepresented gender, the educational institution could raise the affirmative action defense to correct past discrimination against women to any suit brought by male athletes.

4. Compliance with NCAA or Other Rules: Rejected

An educational institution cannot hide behind the rules and regulations of the NCAA or other such group as a defense for its discriminatory conduct. The Regulations and the Final Policy Interpretation specifically reject this as a defense.[254]

5. Financial Constraints: Rejected

Many a school has claimed its hands are tied from complying with Title IX because of the economic woes facing most of this country's educational institutions. As difficult as these economic times may be for educational institutions, "a cash crunch is no excuse."[255]

> Otherwise, if schools could use financial concerns as a sole reason for disparity of treatment, Title IX would become meaningless. Under such circumstances, a school could always use a lack of funds as an excuse to deny equality because it costs money to implement equivalent women's programs with long standing men's programs. This cannot be either the spirit or meaning of Title IX.[256]

6. Equal Reductions in Male/Female Athletic Programs: Rejected

An educational institution's protestations that it has not discriminated, but equally cut male and female teams, will likewise not serve as an affirmative defense. An educational institution may not equally cut teams when one sex is already underrepresented. To do so will almost always result in a violation of the interests and abilities compliance area.

> [W]here budget restrictions have led a recipient to eliminate sports previously offered, there is frequently a compliance problem with this program component. The tendency is for institutions to eliminate a sport previously offered to women who are already underrepresented in the institutions' athletic programs. The result has been that women are now more disadvantaged by the elimination of a women's team despite sufficient interest and ability to sustain a viable team. In this situation, the institution may well be in violation of this program component.[257]

---

253. Final Policy Interpretation, *supra* note 2, at 71,416; 34 C.F.R. § 106.3(b) (1993).

254. 34 C.F.R. § 106.6(c) (1993) ("The obligation to comply with [Title IX] is not obviated or alleviated by any rule or regulation of any . . . athletic or other . . . association . . . ."); *see also* Final Policy Interpretation, *supra* note 2, at 71,422.

255. *Favia*, 812 F. Supp. at 583 (no exception to Title IX's requirements because of school's financial difficulties).

256. *Cook*, 802 F. Supp. at 750.

257. INVESTIGATOR'S MANUAL, *supra* note 28, at 27.

Just as ordinary budget reductions will not afford safe harbor, "equal opportunity" budget reductions will not justify the continuation of pre-existing gender inequality.[258]

V. Remedies for Title IX Violations in Athletic Programs

The stated remedy for a violation of Title IX is the withdrawal of federal funds from the noncomplying educational institution.[259] Nevertheless, Congress requires the Department of Education to notify an educational institution of its failure to comply before withdrawing federal funds or taking any other action against the institution.[260] Moreover, the enforcing agency must attempt to secure compliance by voluntary means before taking any further action.[261]

In practice, OCR has never withdrawn federal funds from an educational institution as a result of Title IX noncompliance in athletics.[262] Instead, as in the case of Colorado State University, the OCR makes repeated attempts to secure compliance from the educational institution through voluntary means and conducts periodic compliance reviews to ensure the institution has met its goals.[263] In effect, Title IX's twenty-year standing threat of withdrawing federal funds has done little to bring compliance in educational institutions across the country.[264] Regulatory remedies against the institutions have proved worthless, perhaps because the OCR has failed to take a hard line with noncomplying educational institutions.[265]

---

258. *See Roberts*, 998 F.2d at 831-32.

259. 20 U.S.C. § 1682 (1988).

260. *Id.*

261. *Id.*

262. *See* Katherine Connor & Ellen J. Vargyas, *The Legal Implications of Gender Bias in Standardized Testing*, 7 Berkeley Women's L.J. 13, 77 (1992).

263. *See Roberts*, 814 F. Supp. at 1513, 1515-16 (discussing OCR's 1983 compliance review of Colorado State). Colorado State made several representations to the OCR in its corrective action plan upon which OCR based its finding of Colorado State's "compliance" with Title IX. Nevertheless, OCR specifically conditioned compliance on Colorado State's promise to carry out the corrective action plan. Colorado State, however, did not meet the goals laid out in the corrective action plan, failing to attain the female participation numbers it pledged it would meet. Indeed, in some instances, Colorado State misrepresented participation numbers to the OCR. *Id.* at 1515-16.

264. *See id.*, 814 F. Supp. at 1513 (noting Colorado State's evidence that every educational institution in this country would be in violation of Title IX if 10.6% point differential is not "substantially proportionate" and rejecting same as a defense for Colorado State); *see also* Wieberg, *supra* note 78, at C10 (noting the disparity in 86 Division I-A schools of men's and women's athletic funding); Wolff, *supra* note 3, at 54-55 (noting lack of spending in most aspects of women's sports).

265. OCR's action with respect to Colorado State provides a good example of agency inaction. In OCR's compliance review of Colorado State, the agency did little to truly enforce Title IX's provisions as it applied to the athletic department. At trial, it became known that Colorado State had misrepresented its female participation numbers to the OCR. Trial Transcript at 477, 483-85, Roberts v. Colorado State Univ., 814 F. Supp. 1507 (D. Colo.) (No. 92-Z-1310) (testimony of Rosalyn Cutler, Colorado State Title IX Coordinator), *aff'd in part, rev'd in part sub nom.* Roberts v. Colorado State Bd. of Agric., 998 F.2d 824 (10th Cir.), *cert. denied*, 114 S. Ct. 580 (1993).

Furthermore, as part of its compliance plan, Colorado State promised to build a softball facility. *Id.* at 491-92. OCR based its original finding of compliance, in part, on this promise to build a softball facility. *See Roberts*, 814 F. Supp. at 1516. Colorado State, however, never

Private litigation, however, now poses a greater threat to educational institutions if they fail to comply with Title IX's gender equity mandates. Since the Supreme Court's decision in *Franklin v. Gwinnett County Public Schools*,[266] which allowed a private litigant to recover money damages for a Title IX violation, educational institutions are more likely to pay closer attention to their Title IX obligations than ever before.[267] Broad injunctive relief also gives schools an incentive to avoid Title IX litigation. This portion of the article discusses the remedies available to successful plaintiffs under Title IX.

A. *Injunctive Relief*

The right to injunctive relief under Title IX seems to be beyond question. In *Franklin*, the Supreme Court impliedly accepted without discussion that a plaintiff could recover equitable relief for a violation of Title IX.[268] Furthermore, in *Guardians Association v. Civil Service Commission*,[269] a majority of the Supreme Court agreed injunctive and other equitable remedies were appropriate for violations of Title VI,[270] the statute on which Title IX was modeled. Courts have followed these cases in imposing injunctive remedies in Title IX athletics cases.[271]

The question then is not whether injunctive relief is an available remedy, but instead, what type of injunctive relief is available. In keeping with the plain dictates of the Regulations and Final Policy Interpretation, a court may order an educational institution in violation of Title IX to create a compliance plan demonstrating the school's intended actions to remedy the violation.[272] Under such an equitable remedy, the court would function much like the OCR by monitoring the educational institution's progress toward meeting the terms of the compliance plan. Finally, the *Roberts* court suggested that an appropriate remedy in a class action case would be to enjoin the violating institution's conduct of its men's

---

built the facility and never informed OCR of its failure to build the facility. Trial Transcript at 492-93, *Roberts* (No. 92-Z-1310). OCR failed to learn of Colorado State's noncompliance, never took any corrective actions against Colorado State, and never took any steps to withdraw federal funds. *See Roberts*, 814 F. Supp. at 1515-16.

Moreover, the Supreme Court in *Franklin v. Gwinnett County Public Schools* discussed the OCR's response to Ms. Franklin's administrative complaint: since the offending teacher had resigned and the school had instituted a grievance procedure, the OCR found the district had come into compliance and terminated its investigation. 112 S. Ct. 1028, 1031 n.3 (1992).

266. 112 S. Ct. 1028 (1992).

267. *Id.* at 1038 n.8.

268. *See id.* at 1032, 1036.

269. 463 U.S. 582 (1983).

270. *Id.* at 584, 607 (White, J., writing the opinion for the Court in which four Justices concurred in the judgment and four Justices dissented).

271. *See Favia*, 7 F.3d at 344; *Roberts*, 998 F.2d at 833-35; *Cohen*, 991 F.2d at 906; *Roberts*, 814 F. Supp. at 1518-19; *Favia*, 812 F. Supp. at 584-85; *Cohen*, 809 F. Supp. at 1001; *Cook*, 802 F. Supp. at 751.

272. *See Cohen*, 991 F.2d at 906 (stating the district court intended to require Brown University to propose a compliance plan rather than mandate the creation or deletion of particular athletic teams if Brown was found in violation of Title IX).

varsity competition until the institution presented a plan which would bring it into compliance with Title IX.[273]

Plaintiffs, however, generally seek more specific injunctive relief to remedy their injuries, such as the creation or reinstatement of a specific women's athletics team. In every federal district court case thus far, plaintiff athletes have been granted the requested injunctive relief.[274] Defendant institutions, however, have challenged the courts' ability to impose such specific remedies as the reinstatement of a specific team, claiming the courts are inappropriately intruding on the institution's discretion.[275] Since Title IX does not require institutions to fund any particular number or type of athletic opportunities, courts have been somewhat receptive to this argument from the defendant institutions.[276]

A distinction is developing based on whether the Title IX action is brought by an individual plaintiff or as a class action. When a plaintiff has brought a claim in her individual capacity, specific injunctive relief is appropriate.[277] When, on the other hand, the action is brought on behalf of a class, the more appropriate remedy may be broad-based injunctive relief that gives educational institutions discretion in the manner in which they will comply with Title IX.[278] In cases brought by athletes in their individual capacities, allowing an educational institution in violation of Title IX to devise its own compliance plan would force plaintiffs to become "unwilling representatives in a class action suit they chose not to bring."[279] When plaintiffs come as a class, however, "the many routes to Title IX compliance make specific relief most useful in situations where the institution . . . demonstrates an unwillingness or inability to exercise its discretion in a way that brings it into compliance with Title IX."[280]

District courts, however, are limited in the reach of their injunctive relief, even when awarding specific relief. In *Roberts*, the district court ordered the reinstatement of the women's softball team.[281] While monitoring Colorado State's compliance with the court's order, the district court ordered Colorado State to hire a coach promptly, recruit new members, and organize a fall season of softball, all in response to Colorado State's

---

273. *Roberts*, 998 F.2d at 833.

274. *See Roberts*, 814 F. Supp. at 1519; *Favia*, 812 F. Supp. at 583-85; *Cohen*, 809 F. Supp. at 1001; *Cook*, 802 F. Supp. at 751.

275. *Cohen*, 991 F.2d at 906; *see also Roberts*, 998 F.2d at 833 (arguing that institution should have been given an opportunity to present a compliance plan).

276. *See id.; see also* Swann v. Charlotte-Mecklenburg Bd. of Educ., 402 U.S. 1, 16 (1971) ("As with any equity case, the nature of the violation determines the scope of the remedy.")

277. *Roberts*, 998 F.2d at 833-34.

278. *Id.* at 833. "Were this a class action, there might be some power to defendant's argument that an order specifically requiring an institution to maintain a softball team goes further than is necessary to correct a violation of Title IX." *Id.*; *see also Cohen*, 991 F.2d at 906-07 (court will initially require university to propose compliance plan rather than court mandating teams added or deleted).

279. *Roberts*, 998 F.2d at 833-34.

280. *Cohen*, 991 F.2d at 907; *see also Swann*, 402 U.S. at 16 ("In default by the school authorities of their obligation to proffer acceptable remedies, a district court has broad power to fashion a remedy that will assure [compliance with federal law].")

281. *Roberts*, 814 F. Supp. at 1519.

apparent "foot-dragging."[282] Colorado State complained that this amplified order from the district court amounted to "micromanaging" of Colorado State's athletics program.[283]

The court of appeals accepted this argument to some extent. Although the district court could order Colorado State to provide the softball team with all incidental benefits of varsity status, it could not go beyond the mandates of Title IX. The Tenth Circuit ruled, "[n]othing in Title IX requires an institution to create a 'top flight' varsity team."[284] The Tenth Circuit further determined that the district court did not have the power to ensure that the reinstated softball team had a good season.[285] Where the district court's order was directed at eliminating a condition which did not violate federal law, the court exceeded the appropriate limits of its power.[286] Since Title IX mandates equal opportunity in the other athletics benefits compliance area,[287] the district court's order regarding coaching, recruiting, equipment, field schedules, and uniforms was affirmed.[288] These types of specific relief were appropriate to remedy claims brought by plaintiffs in their individual capacity.

Choosing to bring an action in an individual capacity gives the plaintiff the advantage of obtaining specific, individual relief. Nevertheless, certain trade-offs accompany the decision to bring the suit in an individual capacity. In *Cook*, the plaintiffs brought the action in their individual capacity rather than as a class action.[289] In the lower court, the federal magistrate judge ordered Colgate University to grant varsity status to the women's ice hockey team.[290] When Colgate appealed this decision, however, the appellate court found the issue moot because all plaintiffs had or would have graduated from Colgate by the time the injunctive relief was to take effect.[291] Plaintiffs' action did not fall within the exception to the mootness doctrine for situations capable of repetition, yet evading review.[292]

> [T]hese plaintiffs may not litigate the claims of students unnamed and unrepresented in this action. We have suggested that a student's claim may not be rendered moot by graduation if he or she sued in a "representational capacity" as the leader of . . . a[n] organization. . . . [H]owever, the complaint herein sought

282. *Roberts*, 998 F.2d at 826.
283. *Id.* at 834.
284. *Id.* at 835.
285. *Id.*
286. *Id.* (citing Milliken v. Bradley, 433 U.S. 267, 282 (1977), for the proposition that federal court decrees exceed appropriate limits if aimed at eliminating conditions that do not violate the Constitution or do not flow from the violation).
287. *See supra* part IV.A.2.
288. *Roberts*, 998 F.2d at 834. The district court did not require Colorado State to provide additional scholarships to its softball players. Nevertheless, Colorado State will have to comply with the requirements of the financial assistance compliance area. *See supra* part IV.A.1. Otherwise, women athletes may bring a new suit for violation of that compliance area.
289. *Cook*, 802 F. Supp. at 739.
290. *Id.* at 751.
291. *Cook*, 992 F.2d at 19.
292. *Id.* at 19-20 (citing Southern Pac. Terminal Co. v. Interstate Commerce Comm'n, 219 U.S. 498, 515 (1911)).

> damages and injunctive relief solely on behalf of the plaintiffs individually, not as representatives of the women's ice hockey club team or other "similarly situated" individuals.[293]

In a class action suit, however, graduation of representative plaintiffs will not render the case moot.[294]

Furthermore, injunctive relief granted to individuals may have force and effect only as long as those individuals remain students at the institution. *Roberts* suggested Colorado State could return to court and seek to have the injunction dissolved once all individual plaintiffs had transferred or graduated.[295] *Favia*, on the other hand, held that modification of an injunction is inappropriate in a class action simply because some of the named plaintiffs have graduated.[296] Given the remedial ramifications of the status chosen, plaintiffs and their counsel must carefully weigh the relative advantages and disadvantages of proceeding as a class or as individuals. Plaintiffs must have in mind the injunctive remedies they wish to recover before they file suit. Plaintiffs may attempt to find the best of both worlds by bringing a class action suit on behalf of a *specific athletic team*, its members, and all future members of the team. This approach could prevent the mootness problem while preserving the ability to get specific injunctive relief, such as the reinstatement of the specific team.

Moreover, to obtain injunctive relief, plaintiffs must meet the standard test set out for injunctions:

(1) Success on the merits (or likely success for preliminary injunction);

(2) irreparable harm to the plaintiffs;

(3) balance of interests between the parties, that is, whether the harm to the plaintiffs outweighs the potential harm to a defendant if the injunction is granted; and

(4) the public interest favors granting an injunction.[297]

---

293. *Cook*, 992 F.2d at 20 (citations omitted).

294. *See Favia*, 7 F.3d at 344. *Favia* presented a case where the court ordered *specific relief* for the representative plaintiffs in a class action suit. The complaint alleged the plaintiffs represented all women students or potential students who participated or sought to participate in the intercollegiate athletic program sponsored by IUP. *Favia*, 812 F. Supp. at 579. There are three possible reasons for *Favia*'s apparent divergence from the class/individual distinction drawn by *Roberts* and *Cohen*. First, reinstatement of those specific teams would be an appropriate remedy for the affected subclass of plaintiffs if the named plaintiffs were construed to be representatives not only of all women students but also of all actual and potential members of the individual teams they represented. *Favia*, 7 F.3d at 342-43 (noting that reinstatement would benefit the class, though to differing degrees as individuals, and that women currently attending and planning to attend IUP would be interested in participating in gymnastics). Second, *Favia* notes that IUP did not have a specific overall plan to achieve total compliance with Title IX. *Id.* at 344. Finally, *Favia* reviewed a preliminary, not a permanent, injunction. To preserve the status quo, the district court ordered reinstatement of the cut teams. *Id.* at 342. The court implied that final injunctive relief should be of the more general, broad-based type suggested by *Cohen* for class actions. *See id.* at 344.

295. *Roberts*, 998 F.2d at 834.

296. *Favia*, 7 F.3d at 342, 344.

297. *Cohen*, 991 F.2d at 902 (citations omitted); *Favia*, 812 F. Supp. at 583 (citation omitted).

1. Success on the Merits

Necessary elements of proof to establish success or likely success on the merits have been discussed fully above. The other elements are discussed here.

2. Irreparable Harm

Plaintiffs may demonstrate irreparable harm in several ways. The methods discussed and recognized by the courts have related to individual and team harms. In *Cohen,* for example, the court found that the potential harm from Brown University's decision to cut women's teams would be irreparable to the teams cut: "[P]laintiffs would suffer irremediable injury in at least three respects: competitive posture, recruitment, and loss of coaching."[298] *Favia* recognized the irreparable harm to the individual athletes:

> By cutting the women's gymnastics and field hockey teams, IUP has denied plaintiffs the benefits to women athletes who compete interscholastically: they develop skill, self-confidence, learn team cohesion and a sense of accomplishment, increase their physical and mental well-being, and develop a lifelong healthy attitude. The opportunity to compete in undergraduate interscholastic athletics vanishes quickly, but the benefits do not. We believe that the harm emanating from lost opportunities for the plaintiffs are likely to be irreparable.[299]

3. Balancing the Interests

The only harm raised by a defendant institution thus far has been a financial one. Courts have soundly rejected financial constraints as an excuse for discrimination.[300] Thus, the financial harm to defendants is outweighed by the irreparable harm plaintiffs will suffer as a result of defendant's unlawful gender discrimination.[301] Defendants can always reallocate resources which have been historically designated for men's athletics or a school may cut back in other areas.[302]

4. Public Interest

As stated succinctly by the district court in *Favia,* "[t]he public has a strong interest in the prevention of any violation of constitutional rights."[303]

If a plaintiff can demonstrate success or likely success on the merits, other athletics cases under Title IX support the imposition of injunctive relief. Indeed, the elements other than success on the merits are nearly

298. *Cohen,* 991 F.2d at 904.
299. *Favia,* 812 F. Supp. at 583.
300. *E.g., Haffer,* 678 F. Supp. at 530 ("[F]inancial concerns alone cannot justify gender discrimination.").
301. *See Cohen,* 991 F.2d at 905; *Favia,* 812 F. Supp. at 584.
302. *Favia,* 812 F. Supp. at 584.
303. *Id.* at 585.

presumed given the nature of the violation in an athletics claim. Thus, injunctive relief is a remedy plaintiffs can anticipate with certainty, should plaintiffs clear the success on the merits hurdle. As discussed above, however, the nature of that injunctive relief will be determined based on the status of the plaintiffs before the court.

B. *Monetary Relief*

The Supreme Court's decision in *Franklin* laid to rest the question whether a plaintiff could recover monetary damages in a private right of action under Title IX.[304] The unanimous Court gave real leverage to victims of educational institutions' gender discrimination by holding that Title IX relief includes, among other things, compensatory damages.[305] *Franklin* was a sexual harassment case wherein a high school teacher/coach sexually harassed and assaulted a female high school student. Despite lower court decisions, the Supreme Court held that the student was entitled to compensatory damages under Title IX for the emotional distress she had suffered as a result of the sexual harassment.[306] Relying on precedent, the Court presumed Congress intended all appropriate remedies, including monetary relief, to be available under Title IX since Congress did not express a contrary intent.[307] Indeed, Justice Scalia noted that Congress has twice addressed Title IX since a private right of action was implied by *Cannon* and implicitly acknowledged that damages are available.[308]

The question left unresolved is whether a plaintiff must prove intentional discrimination to recover compensatory damages under Title IX. The Supreme Court, however, distinguished other cases that held compensatory damages were not recoverable by stating those cases dealt with unintentional discrimination. Since the facts before the *Franklin* court involved intentional discrimination, the Court did not address the application of those cases to Title IX[309] and did not decide whether proof of intent is a prerequisite to the recovery of monetary damages.[310]

Those who argue that damages are not recoverable for unintentional discrimination incorrectly point to the Supreme Court's decision in *Guard-*

304. *Franklin*, 112 S. Ct. at 1036.

305. *Id.* at 1032.

306. For a discussion of *Franklin* and its history, see Note, *supra* note 9.

307. *Franklin*, 112 S. Ct. at 1032 (citing Bell v. Hood, 327 U.S. 678, 684 (1946)).

308. *Id.* at 1039 (Scalia, J., concurring) (citing Civil Rights Remedies Equalization Amendment of 1986, 42 U.S.C. § 2000d-7(a)(1) (1988) (withdrawing States' Eleventh Amendment immunity), and 42 U.S.C. § 2000d-7(a)(2) (1988) (providing that, in suits against States, "remedies (including remedies *both at law and in equity*) are available for [violations of Title IX] to the same extent as such remedies are available for such a violation in a suit against any public or private entity other than a State")) (emphasis added).

309. *Id.* at 1037. The main case distinguished by the *Franklin* court was Pennhurst State Sch. & Hosp. v. Halderman, 451 U.S. 1, 17 (1981), because it involved a case of unintentional discrimination under the Spending Clause. The Court never determined to what extent, if any, *Pennhurst* would apply in cases of unintentional discrimination under Title IX.

310. Some courts have misconstrued *Franklin*'s limited holding and interpreted it as precluding the recovery of monetary damages except in cases of intentional discrimination. *E.g.*, Houston v. Mile High Adventist Academy, 846 F. Supp. 1449, 1457 (D. Colo. 1994).

ians Association v. Civil Service Commission.[311] In *Guardians*, the Court discussed the award of compensatory damages under Title VI.[312] The plaintiffs sought monetary compensation for unintentional discrimination as a result of the disparate impact the New York City Police Department's testing policy had on racial minorities. The Supreme Court reversed the court of appeals' award of compensatory damages for the disparate impact discrimination, but there was no majority opinion discussing the reason for the reversal. Given the confusion caused by the multiple opinions, some commentators have incorrectly construed *Guardians* as having ruled that compensatory damages are not recoverable for a claim of unintentional discrimination.[313]

In fact, there was no majority on whether compensatory damages are recoverable for unintentional discrimination. Justice White, joined by Justice Rehnquist, was the only Justice in *Guardians* to express such an opinion.[314] Justice Powell, joined by Chief Justice Burger, opined that Title VI did not afford a private right of action at all, and, if it did, it only allowed such private actions for intentional discrimination.[315] He did not reach any further issues. Justice O'Connor asserted that discriminatory intent was a necessary element of Title VI claims and that the implementing regulations allowing otherwise were invalid.[316] She did not reach any further issues. Justices Marshall, Brennan, Blackmun and Stevens all agreed that retroactive and prospective relief are available for intentional and unintentional discrimination under Title VI.[317] Justice Stevens opined that monetary relief was available for violations of the implementing regulations through disparate impact discrimination.[318] Thus, a majority of the Supreme Court has not ruled, under Title VI or Title IX, whether compensatory damages are recoverable for unintentional discrimination.[319]

---

311. 463 U.S. 582 (1983).

312. 42 U.S.C. § 2000d (1988). Title VI, dealing with racial discrimination, is the statute upon which Title IX was modeled. *See supra* note 241.

313. *See, e.g.*, Note, *supra* note 9, at 1372 (stating *Guardians* court determined Title VI did not authorize award of compensatory damages for unintentional discrimination). In construing similar anti-discrimination provisions under the Rehabilitation Act, 29 U.S.C. § 794(a) (1988), some courts have relied on *Guardians* to deny monetary damages in instances of disparate impact discrimination. *See, e.g.*, Carter v. Orleans Parish Pub. Schs., 725 F.2d 261, 264 (5th Cir. 1984).

314. *Guardians*, 463 U.S. at 591-603.

315. *Id.* at 609-11. Justice Powell's opinion as to no private right of action was rejected soundly by a majority of the Court in *Cannon*. Moreover, five justices in *Guardians* specifically agreed that a claim could be stated under Title VI for unintentional discrimination. Lower courts have followed this reasoning in Title IX cases and rejected Justice Powell's opinion in *Guardians*. *See, e.g.*, *Roberts*, 998 F.2d at 832.

316. *Guardians*, 463 U.S. at 612-15. A majority of the *Guardians* court rejected this opinion. *See supra* note 315.

317. *Guardians*, 463 U.S. at 615 (Marshall, J., dissenting); *id.* at 645 (Stevens, J., dissenting, joined by Brennan and Blackmun, JJ.)

318. *Id.* at 645 (Stevens, J., dissenting).

319. *But see* Consolidated Rail Corp. v. Darrone, 465 U.S. 624, 631 n.9 (1984) (noting a majority of the *Guardians* court agreed that "retroactive relief is available . . . for all discrimination, *whether intentional or unintentional*, that is actionable under Title VI") (emphasis added); *Franklin*, 112 S. Ct. at 1035 (quoting same).

A closer analysis of Title IX, the Regulations, the Final Policy Interpretation, and the scant cases construing them, demonstrates that compensatory damages are recoverable even in instances of unintentional discrimination. Courts must start with the presumption set forth in *Bell v. Hood*: "where legal rights have been invaded, and a federal statute provides for a general right to sue for such invasion, federal courts may use any available remedy to make good the wrong done."[320] Absent congressional limitation on the remedies available under a statute, the general principle set forth in *Bell v. Hood* must control.[321]

The only previously suggested limitation on the general *Bell v. Hood* principle is the limitation typically placed on Spending Clause legislation as set forth in *Pennhurst State School & Hospital v. Halderman*.[322] In holding in *Guardians* that compensatory damages were not recoverable under Title VI, Justice White resorted to this limitation on Spending Clause legislation, which he determined Title VI to be.[323] Relying on *Pennhurst*, Justice White held that "make whole" remedies are typically not appropriate under Spending Clause legislation.[324] Spending Clause legislation is much like a contract: in return for federal funds, recipients agree to comply with federally imposed conditions.[325] To be valid, those conditions must be knowingly accepted by the institution.[326] The Court held that where a recipient is in violation of the legislation, the recipient should be given the option of complying or withdrawing and terminating the receipt of federal funds.[327]

However, Justice Marshall, in his dissenting opinion, distinguished *Pennhurst* and held that even though Title VI was Spending Clause legislation, compensatory damages need not be limited solely to cases of intentional discrimination.[328] Justice Marshall persuasively argued that *Pennhurst* was inapplicable to the facts before the Court in *Guardians* since *Pennhurst* found the violation urged by plaintiffs was not an actual "condition" imposed on recipients of federal funding.[329] *Pennhurst* merely discussed what obligations were imposed as a result of the statute at issue in that case. It never discussed what consequences flow from the failure to comply with a known condition.[330] Unlike the statute at issue in *Pennhurst*, Title VI clearly imposes the condition on recipients of federal funds that no person shall be denied benefits or subjected to discrimination on the basis of race.[331] Therefore, argued Justice Marshall, recipients

320. *Bell*, 327 U.S. at 684; *Franklin*, 112 S. Ct. at 1033.
321. *Franklin*, 112 S. Ct. at 1034-35.
322. 451 U.S. 1, 17; *see also Guardians*, 463 U.S. at 596 (White, J.) (construing limitations on damages under Title VI, which was passed pursuant to the Spending Clause).
323. *Guardians*, 463 U.S. at 598.
324. *Id.* at 596.
325. *Pennhurst*, 451 U.S. at 17.
326. *Id.*
327. *Guardians*, 463 U.S. at 596-98.
328. *Id.* at 628-34 (Marshall, J., dissenting).
329. *Id.* at 628-29 (Marshall, J., dissenting).
330. *See id.* at 628-30 (Marshall, J., dissenting).
331. *Id.* at 629 (Marshall, J., dissenting).

of funds under Title VI have ample notice that their programs must not have any discriminatory *effect.*[332]

Similarly, providing an environment free from discriminatory effect is also a clear requirement of Title IX. The statute states that no person shall, on the basis of sex, be denied the benefits of or be subjected to discrimination by an educational institution that receives financial assistance.[333] Title IX does not specify that the denial of benefits or discrimination be intentional. Indeed, given the strong directive that no person shall be *subjected* to discrimination on the basis of sex, the statute encompasses any form of discrimination, intentional or disparate impact. The statute's clear language prohibits *all* discrimination. Thus, by accepting federal financial assistance, educational institutions agree to comply with the condition that no person in their programs shall be subjected to any form of discrimination. The existence of disparate impact discrimination in a federally funded educational institution is a violation of the plain language of Title IX.[334]

Furthermore, the Regulations and Final Policy Interpretation also support the argument that avoiding all forms of discrimination is a condition of receiving federal funds. In providing athletics, the Regulations require educational institutions to provide "equal athletic opportunity" to members of both sexes.[335] No distinction is drawn for unequal opportunities that result from unintentional discrimination. Moreover, the Final Policy Interpretation directs a finding of compliance or noncompliance to be based on "whether the policies of an institution are discriminatory in language *or effect.*"[336] Like the clear language of Congress expressed in Title IX, the clear language of the implementing agency demonstrates that Title IX covers intentional and disparate impact discrimination based on sex. Educational institutions, therefore, have clear notice that, by accepting federal funds, they are under an obligation to maintain an environment in which no person is subjected to any form of discrimination.

The *Pennhurst* limitation is inapplicable to Title IX for another reason. The Supreme Court impliedly rejected the idea that educational institutions should be protected from unknown "conditions" when it implied a private right of action under Title IX in *Cannon.*[337] As a result of the Supreme Court's decision in *Cannon,* additional "conditions" were imposed on recipients that were not known before the *Cannon* decision. These conditions include costs of hiring counsel and defending a private action, monetary liability for attorneys' fees, and costs of complying with

---

332. *Id.* at 629-30 (Marshall, J., dissenting).

333. 20 U.S.C. § 1681(a) (1988).

334. *See Guardians,* 463 U.S. at 628-32 (Marshall, J. dissenting) (arguing educational institutions are fully aware of the nondiscriminatory "conditions" placed on them by Title VI and that disparate impact discrimination violates these "conditions").

335. 34 C.F.R. § 106.41(c) (1993).

336. Final Policy Interpretation, *supra* note 2 at 71,417, 71,418 (emphasis added) (defining standards under other athletic benefits and interests and abilities compliance areas).

337. *See Cannon,* 441 U.S. at 709-10.

injunctive relief.[338] Likewise, courts should reject the argument that federally funded educational institutions should be protected from monetary damages for disparate impact discrimination because this is a new "condition" not previously known to them. The purpose of Title IX was to protect individuals from discrimination, not to protect schools.[339]

Finally, it is not altogether clear that Title IX was passed solely pursuant to the Spending Clause. Petitioners in *Franklin* argued that Title IX had also been passed under section five of the Fourteenth Amendment to the United States Constitution.[340] Since the Supreme Court found that *Pennhurst* did not prevent the award of compensatory damages in cases of intentional discrimination, the Court never decided under what authority Congress passed Title IX.[341] The legislative history of Title IX, however, supports the argument that the Fifth and Fourteenth Amendments' guarantees of equal protection are sources of authority for Title IX: "Title IX is an anti-discrimination law. It prohibits discrimination based on sex in educational institutions that are recipients of Federal assistance. The legislative history of Title IX clearly shows that it was enacted because of discrimination that . . . was being practiced against women in educational institutions."[342] "Millions of women pay taxes into the Federal treasury and we collectively resent that these funds should be used for the support of institutions to which we are denied equal access."[343] Title IX was passed "[i]n view of the scope and depth of the discrimination" against

338. *See* Lieberman v. University of Chicago, 660 F.2d 1185, 1191 (7th Cir. 1981) (Swygert, J., dissenting) (criticizing majority opinion based on *Pennhurst* and fully chronicling added "conditions" as a result of *Cannon*), *cert. denied*, 456 U.S. 937 (1982). These "conditions" imposed by virtue of creating a private right of action can have a higher monetary impact than a compensatory damages award. Counsel for Colorado State's softball players, for example, sought over $400,000 as an interim award for attorneys' fees not including time spent on the Tenth Circuit oral argument or Colorado State's petition for certiorari to the U.S. Supreme Court. *See* Plaintiffs' Motion for an Interim Allowance of Attorneys' Fees and Costs Pursuant to the Civil Rights Attorney's Fees Awards Act of 1976, 42 U.S.C. § 1988, at 26, Roberts v. Colorado State Univ., 814 F. Supp. 1507 (D. Colo.) (No. 92-Z-1310), *aff'd in part, rev'd in part sub nom.*, Roberts v. Colorado State Bd. of Agric., 998 F.2d 824 (10th Cir.), *cert. denied*, 114 S. Ct. 580 (1993).

339. *Lieberman*, 660 F.2d at 1191 n.3 (Swygert, J., dissenting).

340. *See Franklin*, 112 S. Ct. at 1038 n.8.

341. *Id.* One federal court has held Title IX is Spending Clause legislation and denied the recovery of any compensatory damages under Title IX. *Lieberman*, 660 F.2d at 1187. Title IX was classified as spending power legislation because the expenditure of federal funds for educational institutions justified the imposition of the prohibition against sex discrimination. *Id.* The *Lieberman* court concluded, prior to the Supreme Court's ruling in *Franklin*, that compensatory damages were not recoverable for a claim of discrimination under Title IX. Only injunctive relief and attorneys' fees were allowed. *Id.* at 1188. Nevertheless, *Lieberman*'s ultimate decision, that no damages are recoverable in any instance, has been clearly rejected by the Supreme Court in *Franklin*. Moveover, the dissent in *Lieberman* persuasively argued that *Pennhurst* was inapplicable to Title IX in the same way Justice Marshall argued it was inapplicable to Title VI. Finally, *Lieberman* never considered whether Title IX had been passed pursuant to another source of congressional authority such as the Fourteenth Amendment.

342. Final Policy Interpretation, *supra* note 2, at 71,423.

343. Comments of Representative Mink, 117 CONG. REC. 39,252 (1971); *see also Cannon*, 441 U.S. at 704 n.36 (discussing legislative intent of Title IX).

women in educational institutions.[344] The main purpose is not to aid schools through funding, but to prevent the use of federal funds to support gender discrimination and to remedy past discrimination.[345] Thus, the Fifth Amendment, which gives Congress broad authority to remedy past discrimination,[346] is a source of congressional authority for Title IX.

In applying the *Bell v. Hood* principle of awarding all appropriate remedies, courts should award compensatory damages even in instances of disparate impact discrimination. Congress has addressed Title IX twice since the Supreme Court's decision in *Cannon* implying a private right of action.[347] In both instances, fully aware of the *Cannon* ruling, Congress declined to limit the private right of action in any manner, choosing instead to expand plaintiffs' rights under Title IX.[348] In the Civil Rights Restoration Act, Congress broadened the Supreme Court's decision in *Grove City College v. Bell*,[349] which limited the application of Title IX to specific programs or activities within schools receiving federal funds, by defining program or activity to mean all the operations of a college or university.[350] More telling is the fact that in the Civil Rights Remedies Equalization Act, Congress waived states' Eleventh Amendment immunity to litigation under Title IX and under other civil rights legislation.[351] In doing so, Congress directed that all remedies otherwise available, both *legal* and equitable, are available against a state. Apparently, Congress anticipated that legal remedies, including damages, would be available in Title IX litigation.[352]

Congress has made no attempt to limit the application of the traditional presumption in favor of all appropriate remedies for a violation of a federal right to Title IX.[353] Certainly, an analysis of the remedial provisions under Title IX confirms that Congress has not limited the recovery of compensatory damages under Title IX to instances of intentional discrimination.[354] In Title VII, Congress specifically limited the recovery of monetary damages to instances of "intentional discrimination (not an em-

344. H.R. REP. No. 554, 92d Cong., 2d Sess. (1972), *reprinted in* 1972 U.S.C.C.A.N. 2462, 2512.

345. *See Cannon*, 441 U.S. at 704.

346. *Cohen*, 991 F.2d at 901.

347. Civil Rights Restoration Act of 1987, Pub. L. No. 100-259, § 1, 102 Stat. 28 (1988) (codified at 20 U.S.C. § 1687 (1988)); Civil Rights Remedies Equalization Act, Pub. L. No. 99-506, t. X, § 1003, 100 Stat. 1845 (1986) (codified at 42 U.S.C. § 2000d-7 (1988)).

348. Unlike its silence under Title IX, Congress has specifically addressed, and limited, the remedies available under Title VII. *See* 42 U.S.C. § 1981a(a) (1) (Supp. V 1993). Congress provided that a complaining party under Title VII may obtain compensatory and punitive damages where a respondent/defendant engaged in *intentional* discrimination. *Id.*

349. 465 U.S. 555 (1984).

350. 20 U.S.C. § 1687 ("[T]he term 'program or activity' and 'program' mean[s] all the operations of . . . a college, university, ...").

351. 42 U.S.C. § 2000d-7.

352. *See Franklin*, 112 S. Ct. at 1039 (Scalia, J., concurring) (stating that Congress had implied a damages remedy under Title IX).

353. *Id.* at 1036.

354. *Id.* at 1035-36.

ployment practice that is unlawful because of its disparate impact)."[355] Having thus demonstrated its ability to limit the recovery of monetary damages when it chooses, Congress refrained from so limiting Title IX's remedies. Since Congress conspicuously failed to place remedial limitations on Title IX similar to those placed on Title VII, *Bell v. Hood* mandates that all remedies apply for claims of intentional and unintentional discrimination.[356] Whether compensatory damages is an appropriate remedy is an issue left to the sound discretion of the court.[357]

Despite the multiple opinions in *Guardians*, a later Supreme Court decision, *Consolidated Rail Corp. v. Darrone*,[358] found a majority of the *Guardians* court agreed that retrospective relief, or back pay, was a form of compensation to which a Title VI plaintiff was allowed whether the discrimination was intentional or not.[359] The Court so held because retrospective compensation was deemed a form of equitable relief.[360] Future damages are also recoverable under Title IX.[361] Failure to provide prospective monetary relief would deny some plaintiffs the benefit of any remedy. For example, if an athlete were in her senior year when a school decided to cut her sport in violation of Title IX, retrospective compensation would be inapplicable. Moreover, future injunctive relief would be of no value to the graduating plaintiff. Only future money damages can remedy that plaintiff. In such instances, courts should avoid limiting remedies where doing so will deny a plaintiff any remedy at all.[362] "If no relief were available to [plaintiff] here, it would be pointless to remand to the district court. An exercise to find liability without finding a remedy would be an exercise in futility."[363]

In order to assure that a plaintiff who has suffered a wrong has a remedy, courts should not limit *Franklin* to cases of intentional discrimination. Victims of discrimination must also be allowed to seek remedies for the discriminatory effect of school actions and policies. In many instances of discrimination within federally funded athletic departments, however, proving intentional discrimination may not be a very difficult task. Each institution has specifically chosen to offer different participation opportu-

355. 42 U.S.C. § 1981a(a)(1) (Supp. V 1993). Although Title VII and Title IX are similar in that they prohibit discrimination, they differ in their focus and purpose. Title VII is a deterrence-based statute, focusing on the discriminatory actions of employers. "It shall be an unlawful employment practice for an employer — [enumeration of prohibited acts]." 42 U.S.C. § 2000e-2(a) (1988). Title IX, on the other hand, is an impact-oriented statute, focusing on the discriminatory effect on victims of gender discrimination. "No person . . . shall, on the basis of sex, be excluded from . . . or be subjected to discrimination . . . ." 20 U.S.C. § 1681(a) (1988).

356. *Franklin*, 112 S. Ct. at 1033, 1035.

357. *See Guardians*, 463 U.S. at 634 (Marshall, J., dissenting) (stating that issue of compensatory damages in intentional and disparate impact discrimination is best decided by federal district courts as triers of fact).

358. 465 U.S. 624 (1984).

359. *Id.* at 630 & n.9.

360. *See id.* at 630 n.9, 631 n.10.

361. *Franklin*, 112 S. Ct. at 1038 (holding that future damages and emotional distress damages were recoverable in Title IX litigation).

362. *Id.*

363. Pfeiffer v. School Bd. for Marion Ctr. Area, 917 F.2d 779, 786 (3d Cir. 1990).

nities to athletes based on their gender (for example, men's basketball v. women's basketball). Moreover, each institution has full control over the number of athletic opportunities, the financial aid, and the other athletic benefits provided to members of each sex. Therefore, the educational institution's deliberate acts of distributing athletic opportunities and resources in a manner which continually favors men is intentional, not accidental or unintentional, discrimination. Although there is probably no intent to discriminate against the individual herself, there is intent to discriminate against women as a group when favoring and preferring men as a group.

One can easily contrast intentional discrimination by an athletic department with disparate impact discrimination by looking at the facts in *Guardians.*[364] The plaintiffs in *Guardians* presented a claim of unintentional discrimination resulting from the disparate impact a facially neutral police entrance exam had on racial minorities.[365] In giving the entrance exam, the police department was not purposely trying to eliminate minorities as police department applicants. Instead, the test *indirectly* limited the number of racial minorities in hiring and limited minorities' promotion and job retention because the test contained a racial bias.[366] Likewise, one could argue that educational institutions may indirectly discriminate against women or minorities because they require college entrance exams that contain a gender or racial bias. This would be an instance of disparate impact or unintentional discrimination.

Unlike the disparate impact of a race or gender biased test, failure to provide proportionate athletic opportunities to women is direct, intentional discrimination by the institution. The institution's conscious acts to offer only a discrete number of positions to women is a deliberate and direct limitation on their ability to participate in athletics. Although schools may provide separate teams to men and women athletes,[367] schools may not provide unequal athletic opportunity. A school would be hard pressed to argue that it accidently failed to provide women with more athletic opportunities. It is difficult to imagine a school accidentally, yet consistently, maintaining participation numbers at a level 10% to 15% lower than female enrollment rates. Likewise, a school intentionally determines the amount of financial aid it will award to women athletes, the number of new uniforms they will have, and whether women athletes will travel to away games by bus while their male counterparts travel to the same location by airplane.

Intention, under tort law, is not necessarily a hostile intent or a desire to do any harm to the injured party. Instead, intent is defined as taking *deliberate actions designed to bring about a result* that will invade the interests of another in a way that the law does not sanction.[368] It is the deliberate

364. *Guardians,* 463 U.S. at 582.
365. *Id.* at 585.
366. *Id.* at 585-86.
367. Final Policy Interpretation, *supra* note 2, at 71,417-18.
368. WILLIAM L. PROSSER, HANDBOOK OF THE LAW OF TORTS 31 (4th ed. 1971).

act which is the focus, not the concomitant effect on others. Applying the same analysis to Title IX, the school's deliberate act of providing fewer participation slots, operating dollars, and financial assistance to women must be the focus, not whether the school, by taking these actions, intended harm to any individual female athlete.

Thus far, no court has reached the issue of monetary damages in a Title IX athletics suit.[369] Based on the Supreme Court's ruling in *Franklin* and the general principles stated therein, however, all monetary relief should be available to female athletes who prove a violation of Title IX whether or not they prove discriminatory intent. The question of damages is one of proof and not one of substantive law. Courts should allow female athletes to present evidence of all damages and award all appropriate relief[370] regardless of the educational institution's intent. Discriminatory *effect* is the issue.

C. *Attorneys' Fees*

As prevailing parties, plaintiffs[371] are entitled to attorneys' fees incurred in pursuing their Title IX claims.[372] "In any action or proceeding to enforce a provision of . . . [Title IX] . . . the court, in its discretion, may allow the prevailing party, . . . a reasonable attorney's fee as part of the costs."[373] Although the award of attorneys' fees is discretionary,[374] a prevailing plaintiff should recover fees unless special circumstances would render such an award unjust.[375] The purpose of awarding attorneys' fees is to "ensure effective access to the judicial process for persons with civil rights grievances."[376]

Fee awards are designed to give attorneys incentive to prosecute civil rights actions which they would otherwise be unable or unwilling to prosecute.[377] Many Title IX athlete plaintiffs will be college-aged (or younger) students with limited financial resources. Moreover, the potential legal fees for these actions far outweigh what a student-plaintiff might win as compensatory damages at trial, making a contingency fee structure unap-

369. Colorado State settled the damages portion of its suit with the softball player plaintiffs. *See Roberts*, 998 F.2d at 833. In *Cook*, the district court denied plaintiffs monetary damages because plaintiffs were fully aware before matriculating that Colgate did not sponsor a varsity women's ice hockey team. *Cook*, 802 F. Supp at 751. That case, however, did not discuss whether damages were generally available to remedy a Title IX violation or whether a plaintiff must prove intentional discrimination to recover such a remedy.

370. *See Franklin*, 112 S. Ct. at 1032.

371. If defendants are found to be prevailing parties, they too may be entitled to attorneys' fees. 42 U.S.C. § 1988(b) (Supp. V 1993). Nevertheless, the standard for the recovery of attorneys' fees by a defendant is far more stringent than by a civil rights plaintiff. To recover attorneys' fees, a defendant must prove that plaintiff's claims were frivolous, unreasonable, or groundless. *See, e.g.*, Christiansburg Garment Co. v. EEOC, 434 U.S. 412, 422 (1978) (discussing fees under Title VII).

372. 42 U.S.C. § 1988(b).

373. *Id.*

374. *Id.*

375. Hensley v. Eckerhart, 461 U.S. 424, 429 (1983) (citing H.R. REP. No. 1558, 94th Cong., 2d Sess. 1 (1976)).

376. *Id.*

377. *See* Ramos v. Lamm, 713 F.2d 546, 552 (10th Cir. 1983).

pealing to potential counsel. Thus, the award of fees is necessary to ensure that these claims will be prosecuted by competent counsel.

Finally, the award or the potential award of attorneys' fees provides a strong incentive to educational institutions to comply with Title IX. Counsel for Colorado State's softball players, for example, sought an *interim* award of attorneys' fees of over $400,000.[378] This request did not include fees for appellate arguments to the Tenth Circuit or briefing in response to Colorado State's petition for writ of certiorari to the United States Supreme Court. An award of attorneys' fees of this size will cause educational institutions to re-think their priorities within their athletic departments, providing those institutions with the incentive to come to early settlements with plaintiffs should litigation be initiated. Accordingly, attorneys' fees themselves become an integral remedy necessary to the proper enforcement of Title IX and other civil rights statutes.

CONCLUSION

After twenty years of virtual paralysis in its application to athletics, Title IX is becoming the vehicle for gender equity that Congress intended it to be. Indeed, in the last few years, courts have recognized Title IX claims concerning unequal athletic benefits and expanded the remedial arsenal available to female athlete litigants. Moreover, female athletes are using litigation as an effective weapon to vindicate women's rights to athletic equality that historically have been denied by schools. After two decades of struggling to implement a law that on its face prohibits gender discrimination, women have finally made courts recognize that "[e]qual athletic treatment is not a luxury. . . . Equality and justice are not luxuries. . . . They are essential elements now codified under Title IX."[379]

Gender discrimination in school-sponsored athletic departments is still rampant. Nevertheless, Title IX litigation, with its potential for money damages and wide-ranging injunctive relief, provides the first real hope for bringing true gender equity to athletics. In the Title IX litigation arena, women athletes have just begun to level the athletic playing fields. Despite admirable initial litigation victories by women athletes, Title IX athletics litigation is in its infancy. Many issues have yet to be reached or resolved by federal courts. This article discussed many of those unresolved issues and presented substantive and procedural methods for analyzing claims of Title IX violations in athletics in hopes of continued pursuit of expanded athletic opportunities for women. Courts should implement these substantive and procedural guidelines to permit the continued effective use of Title IX litigation to bring gender equity to athletics.

---

378. *See* Plaintiffs' Motion for an Interim Allowance of Attorneys' Fees and Costs Pursuant to the Civil Rights Attorney's Fees Awards Act of 1976, 42 U.S.C. § 1988, at 26, Roberts v. Colorado State Univ., 814 F. Supp. 1507 (D. Colo.) (No. 92-Z-1310), *aff'd in part, rev'd in part sub nom.* Roberts v. Colorado State Bd. of Agric., 998 F.2d 824 (10th Cir.), *cert. denied*, 114 S. Ct. 580 (1993).

379. *Cook*, 802 F. Supp. at 750.

Title IX Compliance Areas in Athletics
Title IX
Financial Assistance (Scholarships)
Other Program Areas
Interests and Abilities
Determination of Interests
Selection of Sports
Levels of Competition
Opportunity to Compete
Competitive Team Schedules
Substantial Proportionality
Effective Accommodation of Interests and Abilities
History of Expansion
Proportionate Competitive Opportunities
History of Upgrading Competitive Opportunities

YARDSTICK DAMAGES IN LOST PROFIT CASES: AN ECONOMETRIC APPROACH

ROGER D. BLAIR* AND AMANDA KAY ESQUIBEL**

INTRODUCTION

A business, which is a profit generating asset of its owners, can be damaged or destroyed in a variety of ways: franchise terminations, refusals to deal, tortious interference with contractual relationships, other business torts, contract breaches, and so on.[1] When destroyed, the business's owners have lost their asset and, consequently, the stream of profits generated by it.[2] Not surprisingly, the victimized owners often seek to recover their lost profits through litigation. Two traditional techniques exist for determining these lost profits: the "before-and-after" approach and the "yardstick" approach.[3] The before-and-after approach compares the plaintiff's business to itself during different time periods, before and after the wrongful act.[4] The yardstick approach typically compares the plaintiff's

---

* Huber Hurst Professor of Business and Legal Studies, Department of Economics, University of Florida. B.A. 1964, M.A. 1966, Ph.D. 1968, Michigan State University. Financial support provided by the College of Business Administration is gratefully acknowledged. A good deal of this paper was completed while I enjoyed the hospitality of the Department of Economics at the University of Hawaii.

** Attorney, Amanda K. Esquibel, P.A., Miami, Florida. B.A. 1983, Jacksonville University; M.B.A 1989, J.D. 1989, University of Florida. We have received useful suggestions from James A. Burt, Jeffrey L. Harrison, Lawrence Kenny, Scott D. Makar, Mark Rush, and Deborah Sampieri, who must not be blamed for any remaining errors.

1. Allen S. Joslyn, *Measures of Damages for the Destruction of a Business*, 48 BROOK. L. REV. 431, 431 (1982). With the limitations discussed, the analysis presented in this paper may apply to any legal theory.

2. It is analytically useful to focus on a business's income stream rather than its tangible assets. *See infra* notes 11-13 and accompanying text.

3. *See* Bigelow v. RKO Radio Pictures, 327 U.S. 251, 266 (1946) (upholding a damage award calculated by means of the before-and-after approach without reaching the question of the validity of the yardstick approach); *see also* Lehrman v. Gulf Oil Corp., 500 F.2d 659, 671-72 (5th Cir. 1974) (holding that future profits are an appropriate measure of damages and that a treble damage instruction is unnecessary); Richfield Oil Corp. v. Karseal Corp., 271 F.2d 709, 711 (9th Cir. 1959) (holding that under antitrust laws it is necessary to prove both unreasonable restraint of trade and a causal connection between the defendant's act and the plaintiff's lost revenue in a private treble damages action). For a discussion of the assessment of lost profits in antitrust actions, see Richard C. Hoyt et al., *Comprehensive Models for Assessing Lost Profits to Antitrust Plaintiffs*, 60 MINN. L. REV. 1233, 1233 (1976). Hoyt et al. identify a third approach known as the market share approach, which compares the relative market shares of the plaintiff and the defendant. *Id.* at 1239. The market share approach attempts to compensate for deficiencies in the before-and-after and yardstick approaches. *Id.* at 1233. The Supreme Court recognized the validity of the market share approach in Zenith Radio Corp. v. Hazeltine Research, Inc., 401 U.S. 321 (1971). *See* Hoyt et al., *supra*, at 1241. To employ the market share approach, considerable data on relevant market definition, market sales, market economics, and market entry conditions are required. *Id.* at 1243. *See generally* E. COMPTON TIMBERLAKE, FEDERAL TREBLE DAMAGE ANTITRUST ACTIONS 302-57 (1965) (discussing the damage measures available in antitrust cases).

4. Hoyt et al., *supra* note 3, at 1233-36.

business to another business that is substantially similar.[5] The underlying assumption of the yardstick approach is that "but for" its destruction, the plaintiff's business would have performed as the one to which it is being compared, thereby providing a measure (or yardstick) of the plaintiff's injury.[6]

This article provides a detailed examination of an econometric method of estimating damages by the "yardstick" approach. Generally when employing the "yardstick" approach, a plaintiff attempts to find a clone of its destroyed business.[7] These attempts are usually unsuccessful because clones do not exist in real world markets.[8] This article strives to overcome this obstacle by using econometric methods[9] to construct a "composite business" that is a "statistical clone" of the plaintiff's business. Using this composite model, the plaintiff can project the profits that his business would have earned had it not been destroyed. Although this measure of damages has not received much attention and is not as frequently employed as the before-and-after approach, it nevertheless has great potential.

This article identifies some of the problems associated with the traditional yardstick approach to measuring damages. It then illustrates the use of econometric modelling to construct a "statistical clone" of a hypothetical plaintiff's business with some data obtained from a restaurant chain.

I. THE YARDSTICK APPROACH TO DAMAGES

A. *The Millers' Tale*

A hypothetical example is helpful in illustrating the application of econometrics to the yardstick approach. The B&E Development Corporation ("B&E") owns a chain of highly successful restaurants in the casual theme segment of the restaurant industry. These restaurants, which are known as Jolly Rogers, serve seafood and other family fare in a casual setting. In an effort to expand its culinary empire faster, B&E decided in

---

5. *Id.* at 1236-39.

6. *Id.*

7. For example, in Autowest, Inc. v. Peugeot, Inc., 434 F.2d 556, 564-65 (2d Cir. 1970), the plaintiff, a wrongfully terminated automobile dealer, offered proof of sales projections of other dealers in different markets. Similarly, in Smith Dev. Corp. v. Bilow Enterprises, Inc., 308 A.2d 477, 483 (R.I. 1973), the Rhode Island Supreme Court found evidence of McDonald's' extensive experience elsewhere to be admissible to prove lost profits at a specific location. While it is nearly impossible to find an identical business, the clone approach requires only that the two businesses be comparable and similarly situated in their respective areas of competition. TIMBERLAKE, *supra* note 3, at 330. If a single firm is used for comparison purposes, the variations between the two will be taken into account by the trier of fact. *Id.* Our approach accounts for these variations in a systematic way.

8. *See, e.g.*, Home Placement Serv. Inc. v. Providence Journal Co., 819 F.2d 1199, 1206 (1st Cir. 1987) (holding that the plaintiff must show "product, firm and market comparability" and that the yardstick firm is "unaffected by the defendant's antitrust violation").

9. Econometrics is the application of statistical methods to economic models of markets or firms. *See generally* JAN KMENTA, ELEMENTS OF ECONOMETRICS (1971).

1987 to begin franchising some of its new locations.[10] But things did not go smoothly for B&E's franchising efforts and in 1990, as a result of continuing quality control problems, B&E decided to abandon its franchise operations and terminate its franchisees. Following the terminations, B&E began operating the formerly franchised locations.

Not surprisingly, the franchisees were negatively affected by this decision. Mark and Mary Miller, in particular, were outraged because of the numerous personal and professional sacrifices they had made to become franchisees, only to be terminated after a scant ten months of operation. By terminating their franchise, B&E destroyed the Millers' business.

The Millers were determined to sue for their damages suffered from the destruction of their business. In determining damages for the destruction of a business, the general rule is that the plaintiff may recover an award of lost profits.[11] This rule treats a business as an asset that produces a stream of profits for its owners.[12] The plaintiff's damages are the present value of these lost profits, which is expressed by the following equation:

$$PV(L) = \frac{L_1}{1+i} + \frac{L_2}{(1+i)^2} + \cdots \frac{L_t}{(1+i)^t} + \cdots$$

where $PV(\bullet)$ signifies present value, L represents the lost profits, i is the discount rate, and the subscripts on L denote the year in which the lost profits would have been earned.

Assuming that the discount rate (i) is constant, a compact way of expressing this summation is

---

10. For an examination of the problem of quality assurance for the franchisor and the conflicts that arise between a franchisor and its franchisees, see Scott Makar, *In Defense of Franchisors: The Law and Economics of Franchise Quality Assurance Mechanisms*, 33 VILL. L. REV. 721 (1988).

11. Lost profits are widely accepted as a form of consequential damages for a wrongful breach of contract under common law principles. *See* Blanton v. Mobil Oil Corp., 721 F.2d 1207, 1217 (9th Cir. 1983), *cert. denied*, 471 U.S. 1007 (1985); Cecil Corley Motor Co. v. General Motors Corp., 380 F. Supp. 819, 853-54 (M.D. Tenn. 1974). Similarly, many franchise termination statutes provide for recovery of lost profits. *See, e.g.*, Busch v. National Sch. Studios, Inc., 389 N.W.2d 49, 53 (Wis. Ct. App. 1986), *aff'd*, 407 N.W.2d 883 (Wis. 1987) (applying Wisconsin Fair Dealership Law).

In a case of the complete destruction of a business, the present value of the lost profits, which represent the return on the business as an asset, equals the going concern value but for the business's destruction. The profit stream must be reduced to present value by discounting at a rate that reflects the time value of money as well as business risk. Present value is the value of a dollar, which will not be received until sometime in the future, adjusted to determine its equivalent value if it were received today. In Jones & Laughlin Steel Corp. v. Pfeifer, 462 U.S. 523, 536 (1983), the United States Supreme Court endorsed the present value concept, stating that "the damages award is paid in a lump sum . . . and when . . . invested will earn additional money." For an exposition of discounting methods, see EUGENE F. BRIGHAM, FINANCIAL MANAGEMENT: THEORY AND PRACTICE 89-127, 90 (4th ed. 1985) ("[O]f all the techniques used in finance, none is more important than the time value of money."). *See also* Lehrman v. Gulf Oil Corp., 500 F.2d 659, 663-64 & n.14 (5th Cir. 1974) (finding that "because future profit potential is a principal element of a firm's going concern value an award should not include both").

12. In the case of a partially destroyed business, the lost profits equal the decrease in the business's going-concern value due to the partial destruction.

$$PV(L) = \sum_{t=1}^{T} \frac{L_t}{(1+i)^t}$$

where Σ denotes summation and T is the length of the damage period. When a business has been destroyed and cannot be reestablished, the appropriate value for T arguably is infinity because businesses do not necessarily have finite lives. If this is true, a further simplification is possible if the future losses are assumed to be the same in each year. In those circumstances,

$$PV(L) = \frac{L}{i}$$

For example, if the projected loss were $100,000 per year forever and the discount rate was 12.5%, then the present value of the lost profits would be

$$PV(\mathrm{L}) = \frac{\$100{,}000}{0.125} = \$800{,}000$$

In the Millers' case, the lost profits approach represents the only proper conceptual measure of damages.[13] Since the Millers had only 10 months of actual experience operating a Jolly Rogers restaurant, the before-and-after approach to damage estimation is not a viable option.[14] This leaves the yardstick approach as the most appropriate technique to determine damages.

---

13. From the standpoint of economic theory, the destruction of a plaintiff's business represents a lost opportunity to make profits on this investment. Occasionally courts refer to and permit an award of loss of going concern value in lieu of, but not in addition to, the lost profits. However, the lost profits measure and the going concern value measure are not so much alternative valuations of damages as they are different names for the same thing. The going concern value of a business is the amount received when a business is sold as an operating business. In essence, the going concern value of this profit producing asset is the present value of the profit stream associated with it. *See* Richard R. Rulon, *Proof of Damages for Terminated or Precluded Plaintiffs*, 49 ANTITRUST L. J. 153 (1980). "[T]he resulting value should be a reasonable approximation of what a willing buyer would pay and a willing seller would accept for the business." *Id.* at 155 (citation omitted). For a comprehensive analysis of lost profits in franchise termination cases, see Roger D. Blair, *Measuring Damages for Lost Profits in Franchise Termination Cases*, 8 FRANCHISE L. J. 3 (1988) (demonstrating lost profits method, discounting principles, alternative measures of business valuation and the mitigation principle).

Agreement on the conceptual measure, however, is simpler than empirical measurement itself. The parties will dispute each element of the present value calculation: (1) the estimate of lost profit, each period, (2) the appropriate discount rate, and (3) the duration of the damage period. While each of these is important, our focus is on ways of estimating the lost profits. In these calculations, the interest rate (i) need not be held constant and may vary from year to year. Similarly, the annual profit (L_t) may change over time due to growth in the business, anticipated business cycle influences, the introduction of new competition, and similar influences.

14. Some courts have permitted damage recoveries on the basis of very short operating histories. *See* Mechanical Wholesale, Inc. v. Universal-Rundle Corp., 432 F.2d 228 (5th Cir. 1970) (one month); Larsen v. Walton Plywood Co., 390 P.2d 677 (Wash. 1964) (six months); Edwards v. Container Kraft Carton & Paper Supply Co., 327 P.2d 622 (Cal. Ct. App. 1958) (12 days).

The ideal yardstick would be an identical business functioning in an independent yet identical market. In other words, the ideal yardstick would be a clone of the plaintiff's business. If a clone existed, its performance would reflect the "but for destruction" performance of the plaintiff's business. Clones, however, rarely exist in the real world. Characteristics of demand and supply vary in myriad ways: population size, wealth, income, age distribution of the population, racial and/or ethnic composition, female labor force participation rates, quality of operations, hours of operation, prices charged, presence of competitors and locational convenience can vary widely across different geographic markets. Finding an identical twin of Mark and Mary Miller's restaurant would be exceedingly difficult, if not impossible.

Case law, however, does not demand that the Millers find an identical twin. Instead, it imposes the less stringent requirement of substantial similarity, permitting the use of the yardstick approach.[15] The yardstick approach is nonetheless vulnerable to arguments that the yardstick candidate differs from the plaintiff's business in ways which make the comparison meaningless and render the plaintiff's damage estimate speculative.[16]

B. *The Case Law*

In *Bigelow v. RKO Radio Pictures*,[17] the United States Supreme Court reviewed evidence of damage estimates using both the before-and-after approach and the yardstick approach.[18] The plaintiffs were independent movie exhibitors who claimed rival exhibitors received films before the independent exhibitors as a result of an illegal conspiracy with the defendants.[19] Because of the continuous delays in distribution, plaintiffs lost profits.[20] The plaintiffs introduced two damage estimates at trial.[21] The first estimate used a comparison of the plaintiffs' operations with those of another competitor during the conspiracy period.[22] The competitor, which obtained movie releases in advance of the plaintiffs, was comparable in size to the plaintiffs.[23] The competitor's equipment and location made it less attractive to movie theater patrons.[24] Despite the differences, the

15. *See* TIMBERLAKE, *supra* note 3, at 302-57.

16. *See* Hoyt et al., *supra* note 3, at 1237-39. In some cases, lost profits are too speculative to form the basis for a damage award, but are accepted anyway because the alternative would be the denial of recovery to a wronged party. *See* 2 PHILLIP AREEDA & DONALD F. TURNER, ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION ¶ 344(c) (1978); *see also* Roger D. Blair & William H. Page, *Speculative Antitrust Damages*, 70 WASH. L. REV. (forthcoming 1995).

17. 327 U.S. 251 (1946).

18. *Id.* at 257-58.

19. *Id.* at 253-54.

20. *Id.* at 254.

21. *Id.* at 257. Although damages should be an award of lost profits, the jury in *Bigelow* may have based its award on lost receipts or lost earnings. *Id.* at 258.

22. *Id.* at 257-58.

23. *Id.*

24. *Id.*

yardstick estimate showed a decline in the plaintiffs' business.[25] The second estimate compared the performance of the plaintiffs' business before and after the violation.[26] This before-and-after approach also demonstrated that the plaintiffs' business had declined during the conspiracy period.[27]

On the issue of damages, the Court addressed the defendant's argument that because the plaintiffs' business and the yardstick business competed in the *same* market, the conspiracy actually enhanced the financial performance of the selected yardstick.[28] The yardstick evidence, therefore, amounted only to speculation regarding the impact of the conspiracy upon the plaintiffs' performance.[29] Consequently, the defendant argued that both measures were invalid.[30] The Court rejected this argument and upheld the lower court's damage award.[31] The Court carefully affirmed on the basis of the before-and-after evidence alone, leaving open the question of whether the yardstick evidence would have provided sufficient support for the verdict.[32]

By negative implication, one may draw the inference from *Bigelow* that the yardstick approach is a viable technique.[33] The yardstick business in *Bigelow,* however, was not in an independent market; therefore, its financial success could be attributed at least in part to the violation.

Cases such as *Farmington Dowel Products v. Forster Manufacturing Co.*[34] also illustrate what factors a court considers significant in determining whether a yardstick business is substantially similar.[35] In *Farmington,* the defendants violated the antitrust laws by discriminating in price.[36] The district court refused to admit evidence offered by the plaintiff's expert that used one defendant's business as a yardstick.[37] The district court identified several ways in which the two businesses were not comparable.[38] First, the product lines were different.[39] While the plaintiff had a single product, the defendant's business was more diversified.[40] Second, the methods of distribution differed considerably.[41] The plaintiff's business

25. *Id.* at 258, 260. On the issue of fact of damage, the Court considered the two estimates to be cumulative evidence tending to show that the plaintiff was harmed. *Id.* at 260.

26. *Id.* at 258.

27. *Id.*

28. *Id.* at 260-63.

29. *Id.* at 263.

30. *Id.*

31. *Id.* at 266.

32. *Id.*

33. Another inference to be drawn from *Bigelow* is that the yardstick approach is inferior to the before-and-after approach because of the difficulties associated with finding a substantially similar business in an independent market.

34. 421 F.2d 61 (1st Cir. 1970).

35. *Id.* at 82.

36. *Id.* at 65. Price discrimination is illegal if it may "substantially . . . lessen competition or tend to create a monopoly." 15 U.S.C. § 13(a) (1988).

37. *Farmington,* 421 F.2d at 82.

38. *Id.* at 82 n.48.

39. *Id.*

40. *Id.*

41. *Id.*

had little sales organization compared to the defendant's national system.[42] Third, significant differences existed between the financial structures of the two firms.[43] The plaintiff's business was minimally capitalized, while the defendant's business was adequately capitalized.[44] Finally, the defendant operated in the same market as the plaintiff, not an independent one, meaning the yardstick's profits included the illegal profit gained from the price discrimination.[45] On appeal, the court of appeals affirmed the exclusion of the yardstick evidence.[46]

In an effort to overcome similar problems, the plaintiff's expert in *Admiral Theatre Corp. v. Douglas Theatre Co.*[47] introduced a composite yardstick—a rudimentary version of that to be discussed later in this article.[48] Because of the geographic proximity of two competitor theatres to the plaintiff's theatre, and because the seating capacity of the plaintiff's theatre was about one half that of the two competitors' theatres combined, a major assumption of the damage evidence was that the plaintiff's business was comparable to a composite, or average, of the two competitor theatres "taken together and halved."[49] The court was required, therefore, to examine both the independent market issue and the substantial similarity issue. Echoing *Bigelow* and *Farmington Dowel*, the court pointed out flaws in using firms within the same market as yardsticks.[50] The profits reflected the illegal activity because the yardsticks operated in the tainted market.[51] The court criticized the damage estimate for failing to accommodate differences between the plaintiff and the yardstick's film supply, bidding practices, and revenues and expenses associated with films not affected by the conspiracy.[52] Based on these foundational weaknesses, the court rejected the composite technique.[53]

Although the court in *Admiral Theatre* rejected the composite yardstick damages as too speculative, the idea of using a composite when there are no substantially similar businesses is nonetheless a good one. A useful composite would contain all the characteristics of the plaintiff's business but would not suffer from the infirmities associated with operating in the same market. This composite approach is within the spirit of the yardstick

---

42. *Id.*
43. *Id.*
44. *Id.*
45. *Id.* This sentiment was echoed in the *Todorov* decision regarding a physician's claim for a share of the monopoly profit, which he lost through a hospital's denial of staff privileges. Todorov v. DCH Healthcare Authority, 921 F.2d 1438, 1452-53 n.23 (11th Cir. 1991).
46. *Farmington*, 421 F.2d at 83.
47. 437 F. Supp. 1268 (D. Neb. 1977).
48. *Id.* at 1297. In *Admiral Theatre*, the plaintiff alleged a conspiracy in film distribution. *Id.* at 1272-73. Before delineating its reasons for rejecting this unconventional approach, the court noted that the foundation upon which the plaintiff built the case was a contributing factor in not overcoming a directed verdict. *Id.* at 1274.
49. *Id.* at 1297.
50. *Id.* The court also questioned the validity of plaintiff's assumption that better film quality for the plaintiff would have had no effect on the yardstick's patronage. *Id.* at 1298.
51. *Id.* at 1298.
52. *Id.*
53. *Id.* In some instances, industry averages have been used successfully. *See, e.g.*, Tucson Fed. Sav. & Loan Ass'n v. Aetna Inv. Corp., 245 P.2d 423, 429 (Ariz. 1952).

method because it is an attempt to build a "clone" where a suitable substitute does not actually exist.[54]

The difficulty with the composite approach was aptly illustrated in *Admiral Theatre.* The composite did not withstand a rigorous analysis. For the composite yardstick approach to be effective, it must be more scientific. Econometric or statistical techniques can be used to develop composite yardsticks that will provide a better foundation for comparison with the plaintiff's business and withstand judicial scrutiny. Techniques like multiple regression analysis allow an expert to systematically incorporate more relevant data in a composite yardstick and consequently produce a more reliable and precise damage estimate. In the next section, a yardstick is constructed through the use of statistical techniques that are simple to follow.

II. AN ECONOMETRIC APPROACH TO YARDSTICK DAMAGES

Econometrics is a blend of mathematical and statistical methods that is used to analyze economic data.[55] A basic tool of econometrics is multiple regression analysis, a statistical technique that has proved to be extremely useful in estimating the effects of one or more economic variables on some item one is interested in measuring. In the language of statistics, this item is referred to as the "variable of interest."[56] In the context of a suit for commercial damages, the variable of interest is lost profits, because that is what a plaintiff is entitled to recover as a result of a defendant's wrongful conduct.[57] In simple terms, if one is interested in estimating profits, multiple regression analysis is a statistical method of identifying the set of factors that impact upon or determine profits. In addition, multiple regression analysis provides a way to estimate the size of each factor's impact on profits. Therefore, multiple regression analysis, by isolating and estimating the importance of the determinants of profits,

54. For other cases discussing yardstick damages, see Autowest, Inc. v. Peugeot, Inc., 434 F.2d 556, 563-67 (2d Cir. 1970) (termination of an automobile distribution franchise); Union Carbide & Carbon Corp. v. Nisley, 300 F.2d 561, 575-78 (10th Cir. 1961) (government purchase of ore allowed as a yardstick measure of plaintiff's damages); Richfield Oil Corp. v. Karseal Corp., 271 F.2d 709, 714 (9th Cir. 1959) (to prove lost sales, plaintiff introduced the sales of a comparable product); Flintkote Co. v. Lysfjord, 246 F.2d 368, 391 (9th Cir. 1957) (plaintiffs based their lost profits estimate on the assumption that they would have made as much working for themselves as they and their previous employer made combined); Delaware Valley Marine Supply Co. v. American Tobacco Co., 184 F. Supp. 440, 444-50 (E.D. Pa. 1960) (lack of business history and lack of yardstick made it necessary to construct a hypothetical profit projection); Homewood Theater, Inc. v. Loew's Inc., 110 F. Supp. 398, 412 (D. Minn. 1952) (classic yardstick case where plaintiff introduced evidence showing the gross receipts of comparable theatres); William Goldman Theatres, Inc. v. Loew's, Inc., 69 F. Supp. 103, 107 (E.D. Pa. 1946) (another classic yardstick case where plaintiff introduced evidence of the receipts of a comparable theater operated by the defendant).

55. For thorough treatments of econometrics, see JAN KMENTA, ELEMENTS OF ECONOMETRICS (2d ed. 1986). *See also* G.S. MADDALA, INTRODUCTION TO ECONOMETRICS (2d ed. 1992).

56. For a compact treatment of multiple regression analysis, see Franklin M. Fisher, *Multiple Regression in Legal Proceedings*, 80 COLUM. L. REV. 702 (1980). For more extended developments, see KMENTA, *supra* note 55. *See also* MADDALA, *supra* note 55.

57. In an antitrust context, see Jeffrey L. Harrison, *The Lost Profits Measure of Damages in Price Enhancement Cases*, 64 MINN. L. REV. 751 (1980).

makes possible the prediction of profits "but for" the wrongful action. Absolute certainty about the size of the loss cannot be attained, but the available data can be used to make the estimate as scientific as possible. Furthermore, multiple regression analysis can scientifically indicate the level of certainty associated with a damage estimate. For plaintiffs desiring to use multiple regression analysis or any other evidence of damages, proof of damages does not require absolute certainty. On the contrary, the case law provides that "reasonable certainty" or a "reasonable inference" is enough.[58] Because multiple regression analysis is a scientific method, it offers a plaintiff the potential for obtaining damage evidence that is more likely to satisfy this standard than other types of damage evidence that may be presented anecdotally or on an *ad hoc* basis. Multiple regression analysis works by using a random sample drawn from the population and making a statistical inference or conclusion about the population from the observed characteristics of the sample. The analysis expresses that inference or conclusion in terms of the level of certainty or degree of confidence associated with the conclusion.[59]

A. *Fundamental Steps in Multiple Regression Analysis*

Any sound multiple regression analysis proceeds through several orderly steps.

Step 1. First, one builds an economic model and thereby identifies the most important economic and demographic factors that influence the variable of interest. Economic theory determines which variables should be included and which should be excluded.[60] Practical problems of implementation relating to the availability or reliability of data may require subsequent modification of the model, but the initial focus should be on the *theoretical* construction of the model.

At this stage, the model is in a general, abstract form:

$$P = f(X,Y,Z)$$

In this equation, P represents profit and is the dependent variable of interest to be explained by the model. The independent variables—those that determine or explain profits—are X, Y and Z. Finally, the f(•) notation denotes the mathematical relationship, whatever it may be, between the independent variables and the dependent variable. In other words, the

58. *See* Bigelow v. RKO Radio Pictures, 327 U.S. 251, 264 (1946) ("the jury may make a just and reasonable estimate of the damage based on relevant data, and render its verdict accordingly"); Story Parchment Co. v. Paterson Parchment Paper Co., 282 U.S. 555, 562 (1931) ("The rule which precludes the recovery of uncertain damages applies to such as are not the certain result of the wrong, not to those damages which are definitely attributable to the wrong and only uncertain in respect of their amount."); Malcolm v. Marathon Oil Co., 642 F.2d 845, 858 (5th Cir. Unit B Apr. 1981) ("averages may be used to show that the plaintiff generally lost money over time").

59. Degrees of confidence have a very precise statistical definition; simply put, however, they are measures of one's confidence in a model or the reliability of that model.

60. In particular, it is inappropriate to let the data dictate the specification of the model. If this is done, the results will be *ad hoc* and may be of no use in predicting sales or profits in general.

value of P is explained by or dependent on the values of X, Y and Z, and conversely, X, Y, and Z determine or explain the value of P.

Step 2. The second step in this process is to select a specific functional or mathematical form that will accurately express the relationship between the independent variables and the dependent variable. For most practical purposes, a linear relationship is assumed, at least initially. This means the relationship between the dependent and independent variables would appear to be a straight line if plotted on a graph with one of the axes representing P and the other representing X or Y or Z. In other words, a linear relationship means that one would expect that there would be some ratio (e.g., 3:1 or 2:1) that would *always* explain the relationship between the dependent and independent variables. One may assume that the economic model can be expressed as follows:

$$P = a + bX + cY + dZ$$

Of course, we do not live in a linear world. It is unrealistic to expect that the relationship between P and X, Y or Z is precisely linear or does not change depending on the value of variables. But in some instances, a linear approximation, although not precisely accurate, may provide a sufficient level of certainty to be used for purposes of legal proof—a determination which should be jointly made by the lawyer and expert. When linear approximation is not accurate, however, nonlinear estimation methods can be used. They require a more complicated multiple regression model that is conceptually the same as a linear model but involves higher-level mathematics.

Step 3. Having specified a clear mathematical relationship between the dependent variable (P) and a set of independent variables (X,Y,Z), the third step involves gathering data, because one must have observations or sample data on all of the variables P, X, Y, and Z. Normally, a sample drawn from the population is examined instead of the entire population. Statistical inference permits us to make judgments about the population on the basis of sample information. The sample data are drawn from appropriate data sources, which will vary from case to case. In many instances, company records can be tapped for important economic variables such as sales, prices, output, employment, production capacity, salaries and wages, and the like. Data on population, age distribution, ethnicity, wealth and other economic and demographic factors can be obtained from public sources or from private data collection services. In some cases, the data must be compiled from market transactions.

In any event, care must be exercised to insure that the data are recorded accurately and that common measurements are used. For example, if *annual* advertising expenditures are recorded from one restaurant location, then *annual* advertising expenditures must be recorded for all locations. It would distort the analysis if *quarterly* data were inadvertently included for several locations. Moreover, it is helpful if the data span the same time period; otherwise, exogenous economic factors could distort

the analysis.[61] Suppose, for example, that one firm is on a July 1 through June 30 fiscal year and another is on a September 1 through August 31 fiscal year. Hurricane Andrew devastated the Miami area in August 1992. The influence of this devastation appears in the 1992 fiscal year for one firm but not for the other. This is a potential source of distortion that may require adjusting the data.

Step 4. The fourth step is to use the multiple regression model constructed to analyze the data for the particular case at hand. This will be examined in greater detail in the next section.

Step 5. Finally, one must interpret the results of the analysis and apply them to the problem of damage estimation. This too shall be examined below.

B. *Application to Jolly Rogers*

We will use the Jolly Rogers hypothetical to develop the concept of multiple regression analysis and its application in proving yardstick damages. Following termination, the Millers sued the B&E Development Corporation and won a resounding victory. Among the remedies they requested were damages for lost profits. To make a damage award to the Millers, one must estimate the lost profits during the 1990-1992 period.[62]

One approach to estimating lost profits relies upon the historical experience of the business in question. As discussed before, this approach may be a problem if the franchisee does not have a long financial history, which is precisely the Millers' situation. The yardstick approach addresses this problem by examining the economic performance of other similar franchisees and drawing inferences about the Millers' performance "but for" the wrongful termination. In this regard, multiple regression analysis is a useful tool because it assists in estimating profits for the terminated franchisee using the factors that determine profits for other franchisees. To this end, we will estimate the relationship between profits and certain economic and demographic variables for the outlets in the franchise system. We can then employ the specific characteristics of the Millers' location and the estimated relationship for franchisees generally to statistically infer the Millers' lost profits.

A sample of seventy-five restaurants in the Jolly Rogers system was selected for analysis. B&E's records provided information on profits by location in 1990. These data displayed a wide range of values: the lowest profit figure was $80,300, the highest was $266,600, and the average was $153,500. The distribution within this range is provided in Table 1. From these data alone, it would be difficult to predict the profit level that the

61. It is possible to control for these exogenous factors, and one must be sure to do so if data are drawn from different time periods. In addition, this makes the analysis more complicated.

62. This may be a bit too simplistic and unfair to the Millers. To some extent, momentum is critical to business success. This interruption could have adverse consequences that extend into the future despite the injunction. In this event, future losses will have to be estimated as well.

Millers could reasonably expect. One might consider using an average, but multiple regression analysis provides a more scientific and accurate prediction.

TABLE 1
DISTRIBUTION OF FRANCHISEE PROFITS 1990

Profit*	Number of Locations
$75-100	7
100-125	12
125-150	23
150-175	15
175-200	19
200-225	3
225-250	2
250-275	4

* Measured in thousands of dollars.

Experience has shown that patrons of this franchise system are drawn largely from those who live within a five-mile radius of the restaurant. As a result, we are interested in the economic and demographic characteristics of the population within a five-mile ring around each location. The initial model, which is very simple, has only one independent variable. Specifically, this model expresses restaurant profits as a function of the total population within a five-mile ring surrounding each restaurant location.[63] In Figure 1, each restaurant location is represented by a point that reflects the population within a five-mile ring and the profits at that location. An examination of Figure 1 does not reveal any precise mathematical relationship between population and profits. The horizontal line represents the average profit of $153,500 across the entire sample of seventy-five locations. Those restaurants located in less populous areas tended to have below average profits. Conversely, Figure 1 indicates that restaurants located in more populous areas *generally* had higher profits. This suggests that we do not want to rely upon the sample mean for an estimate of profit at any particular location.

Multiple regression analysis allows us to uncover the general tendency of profit to increase with population and to approximate it with a straight line. Thus, the initial regression equation is specified as

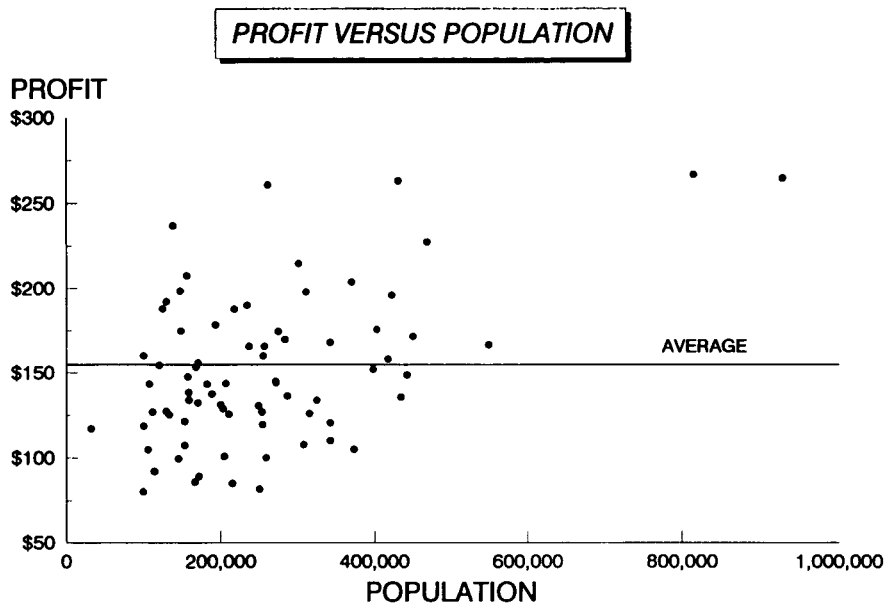
$$P = a + bPOP + e$$

where P represents profit, POP represents population within a five-mile ring, and e accounts for random disturbances or factors other than POP. By including the random disturbance term, we recognize that the profit figure equals the sum of a deterministic component, a + bPOP, and a probabilistic component, e.[64] The deterministic portion is composed of

63. A private data collection service obtained the population data.

64. Adding the random error term in the regression equation makes the model probabilistic rather than deterministic. As a result, there is a whole range of values for P corresponding to each value of the independent variable. For instance, one restaurant with

FIGURE 1



the identified independent variables (in this case, just POP) that we expect to explain the franchisee's profit (P). The probabilistic component of the equation is the unexplained portion of P, those factors other than population within a five-mile ring influencing the value of P, which is referred to as the random error. The inclusion of e is an explicit recognition that a precise relationship between profit and population is unlikely in real life. While population may help to explain profits, it does not completely explain or determine them. In other words, it is unlikely that we can completely explain P. While we believe that our model captures the major determinant of sales, population within a five-mile ring, there are other variables that have relatively minor effects on the dependent variable.[65]

200,000 people within a five-mile ring may have profits of $95,000 per year, while another restaurant in a different location that also has a population of 200,000 within a five-mile ring may have profits of $125,000. Naturally, this means that P cannot be predicted exactly, but we can make certain probabilistic statements with measurable degrees of confidence.

In statistical generalities, this set of values for P is referred to as a probability distribution. A probability distribution is the range of values and frequency of values of P. In other words, P may range between 0 and $1,000,000 with 0 and $1,000,000 being very uncommon and $153,500 being the most common value. A graph of all the values of P around the average value of P would produce a probability distribution.

65. The cumulative influence of these minor effects, however, may not be trivial and, therefore, must be captured. The random error incorporates these minor factors in the regression model. It is assumed that the cumulative impact of these minor factors is small relative to the influence of the independent variable(s). Moreover, it is assumed that these influences can be treated as though they occurred by chance, i.e., that on average they will cancel each other out and have no net effect on our damage estimate.

There are three standard assumptions regarding statistical properties of the error term. First, we assume that the expected value of the error term is zero. This means that if we take repeated samples, the random error term for any particular value of POP is on average

C. *Estimation of the Regression Parameters*

The model presented in the initial regression equation (1) is a linear one. Since the expected value of e is zero, on average it will wash out. The next task is to estimate a and b, which are known as the regression coefficients or the parameters of the regression model. The goal is to obtain estimates of a and b that will provide a straight line that best fits the data in Figure 1. The notion of a "best fit" implies some standard. The statistical standard employed is that of a measure known as "least squares."[66] A statistical computer program is most often used to calculate least squares. This technique finds estimates of a and b such that the sum of the squared deviations of the observed values of P from the values that lie on the regression line are minimized. In other words, for any given a and b, one calculates the difference between the actual value of P and the predicted value, squares these differences, and then totals. Then different values for a and b are chosen until the lowest value for the sum of the squared deviations is found. In that sense, the least squares regression line "fits" the data "better" than any other line. Using these "best fit" estimates of a and b, we construct the straight line through the data plotted in Figure 1.

For the observations plotted in Figure 1, least squares regression yields the following estimates:[67]

$$\hat{a} = \$114{,}125$$

---

zero—it is neither positive nor negative and, therefore, it will wash out on average. This is extremely important, because if the expected value of the error term were not zero, then we could not estimate the effect of POP on P simply by looking at different values for POP. In this event, changes in POP would be related *systematically* to changes in e and our model would not provide a good estimator because it is really not explaining variations in P.

The second assumption about the random error term is that its probability distribution is *normal*, i.e., it follows the classic bell-shaped curve. A normal probability distribution has the majority of the values distributed close to the mean or average value; as you move further from the mean the frequency of the values drops off. Normality of the error term means that for every value of POP, the random error is distributed normally around its mean or average value, which we just assumed equals zero. This seems like an extremely restrictive assumption, but it is less so than one would think. Recall that the error term is due to the presence of a large number of small influences that are not explicitly included in the model. Each of these minor influences produces a small deviation of P from the value that it would assume if the relationship between POP and P were completely deterministic.

An interesting statistical result is that such random *unsystematic* errors tend to be distributed normally. Normality of the error term is an extremely useful property for two reasons. First, it makes precise probability statements about the regression results possible because we know so much about the normal distribution. Second, it is essential for an important property of the estimators, which we will examine in more detail below. The third assumption is that the error term is homoskedastic, which simply means that every random error has the same (albeit unknown) variance. In our example, this means that the variance of the error term is not greater for areas with larger populations than for areas with smaller populations. In other words, the chance of observing a small error rather than a large error is the same for all values of POP. This assumption is very important because regression analysis is a form of averaging that deals with small random deviations in an efficient way provided that they are independent. If the random errors are not independent, then something should be done explicitly to deal with this. *See* KMENTA, *supra* note 55; *see also* Fisher, *supra* note 56.

66. *See* KMENTA, *supra* note 55 for an especially clear derivation of the least squares estimation formulas. The mathematics is straightforward and uncontroversial.

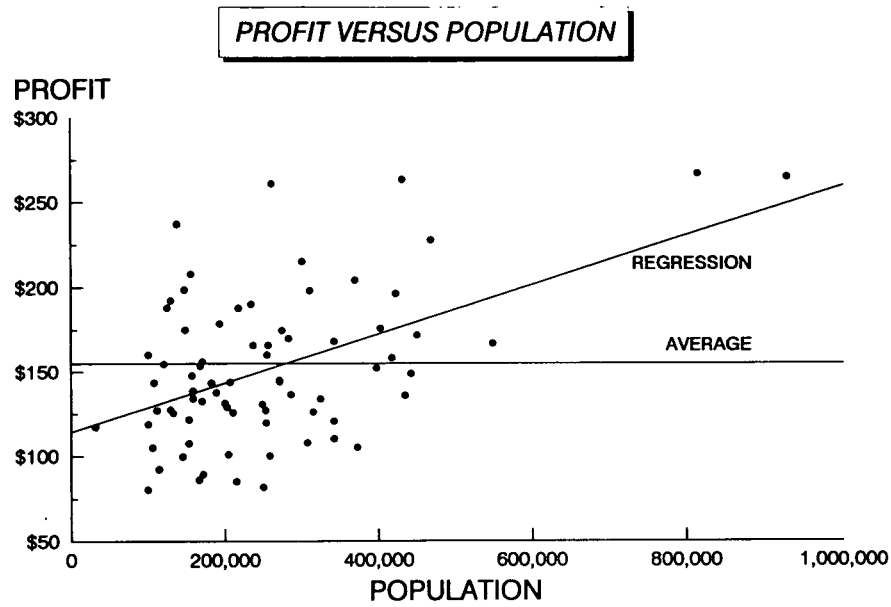
67. The caret or "hat" over a variable is a signal that this is an estimate of the true (unobservable) value of the population parameter.

and

$$\hat{b} = \$.15$$

where $114,125 is the vertical intercept of our regression line and $.15 is the slope. In Figure 2, we have reproduced the observations displayed in Figure 1. In addition, we have used $114,125 as the intercept on the vertical (i.e., profit) axis and $.15 as the slope to plot the regression line.

FIGURE 2



The estimates indicate that each restaurant enjoys average annual profits of some $114,125 irrespective of the population base. In addition, average profits will increase by $.15 per person within the five-mile ring. If, for example, there were 400,000 people within a five-mile ring around a particular site, this simple model would predict average annual profits of $114,125 plus $.15 times 400,000 people or, $60,000, for a total of $174,125.

When using least squares regression analysis, one must be aware that the results are sensitive to extreme values in the sample. Because extreme values represent relatively large deviations, they can distort our perception of the true relationship. Extreme values are known as outliers, because they tend to fall outside the expected range of values for the data.[68] When one plots the data as we have done in Figure 1, one can identify observations that do not seem to fit. With the least squares method, outliers must be handled with care. Any temptation to discard outliers must be resisted

68. For clear examples of how an outlier can distort the regression analysis, see MADDALA, *supra* note 55, at 88-96. *See also* Daniel L. Rubinfeld & Peter O. Steiner, *Quantitative Methods in Antitrust Litigation*, LAW AND CONTEMP. PROBS., Autumn 1983, at 69, 93.

because these observations may provide important information and actually improve the reliability of the model.

The first step in a prudent analysis of outliers is to determine if the potential outlier is really a problem by dropping these observations and allowing the computer to "refit" the line using the modified data set. If the new line is not appreciably different from the first one, the outliers are usually not troublesome. Nonetheless, it may be wise to retain these observations, especially if the data set is small.

If the modified data set produces a substantially different regression line, the outliers must be inspected more closely. Sometimes outliers indicate mistakes in either data collection or data entry. In these cases the error should be corrected or the outlier should be removed from the data set and replaced with the corrected value. If the outlier is a reliable data point, it generally should remain in the model because of the information it inherently provides about the variability in the sample data.[69] The analyst must be very careful when he elects to delete any data. In a litigation context, such decisions will be challenged and justification must be provided in order to avoid the impression that the results were contrived.[70]

In terms of our data set, an outlier would be a data point with a value of, say, $500,000. Such a value may be observed if someone mistakenly recorded an extra zero or it might reflect actual profits for an exceptionally successful restaurant. If the former, it should be corrected. If the latter, it may improve the model because it is an indicator that, given certain factors, such a value for profit is possible.

D. *Hypothesis Testing*

For our regression model to be useful in predicting profit, the independent variable, population, must actually contribute some information to explain profit. To test the utility of the regression model, one must test the validity of the hypothesis that population is a determinant of profit. To understand the testing method, it is necessary to understand the meaning of the model's parameters.

As discussed previously, the first parameter, a, is the intercept on the profit axis. Literally interpreted, this parameter indicates that when a res-

---

69. *See* Rubinfeld & Steiner, *supra* note 68, at 93.

70. Michael O. Finkelstein & Hans Levenbach, *Regression Estimates of Damages in Price-Fixing Cases*, LAW AND CONTEMP. PROBS., Autumn 1983, at 145, 150-153, relates an extreme example of massaging the data. In *Chicken Antitrust Litigation*, 1980-2 Trade Cas. (CCH) ¶ 63,485 (N.D. Ga. 1980), the plaintiff's experts found that the prices of chicken were actually *lower* during the conspiracy period than their model would have predicted absent the conspiracy. To conclude that prices were higher, they excluded all observations during the conspiracy period where actual prices were not above the model's predicted price. On a purely statistical basis, one would expect that half of the time the prices actually observed will exceed the predicted value and half of the time they will fall below the predicted value. The experts' justification for deleting those instances where observed prices were below the predicted prices was that the conspiracy was effective only on an intermittent basis. Be that as it may, by looking at only the observations that support the *a priori* suspicion that prices should be higher due to a conspiracy, one will be led to a damage estimate that is unrelated to the conspiracy.

taurant is located in a spot with no population within five miles, it will have $a in profit. In this example, P = $114,125 when POP = 0. This interpretation, however, is unreliable because it is based on values falling far outside the range of the sample data. In the data set, there is not a single business at a site where there are no people within five miles. In fact, the smallest population recorded in the sample data was 32,000 and the next smallest was 93,000. Therefore, predictions outside the data range that rely on the interpretation of the intercept are suspect and should be avoided if at all possible.[71]

The second parameter, b, also discussed previously, is the slope of the regression line. For the purpose of this model, b indicates that for each additional person living within five miles of a restaurant location, the model predicts that the business will experience an additional $b in profit, which in this example equals an additional $.15. Since this parameter is critical for prediction, this parameter must be tested to ensure that it is statistically significantly different from zero.[72] In other words, the test will determine whether the original hypothesis is valid, i.e., whether the independent variable, POP, contributed to the determination of or explained the dependent variable, P.

The model's hypothesis is given below in two parts. The first is the so-called "null hypothesis," which states that the independent variable, POP, does not in any way explain P or provides no information in predicting the dependent variable, P. The "alternative hypothesis" states that there is a functional relationship between the independent and dependent variables:

$$H_0: b = 0 \qquad H_a: b \neq 0.$$

If the null hypothesis is rejected, then the statistical inference is that some statistically significant relationship exists between population and profit. Rejection also indicates that the model has some utility for prediction. A computer program performs this hypothesis testing.[73]

The results of this computer testing indicate that the probability of drawing a sample of seventy-five restaurants that would produce an esti-

---

71. Rubinfeld & Steiner, *supra* note 68, at 95.

72. For an interesting analysis of statistical significance and legal standards of proof, see David H. Kaye, *Statistical Significance and the Burden of Persuasion*, LAW AND CONTEMP. PROBS., Autumn 1983, at 13. *See also* David H. Kaye, *Is Proof of Statistical Significance Relevant?*, 61 WASH. L. REV. 1333 (1986). For the statistical analysis, see KMENTA, *supra* note 55.

73. The program produces what is called a "t-statistic" to evaluate the coefficient's statistical significance. The resulting t-statistic can be compared to critical values that correspond to any predetermined level of statistical significance, which is referred to as the alpha level. The alpha level is a statistical measure of confidence; the degree of confidence is inversely related to the alpha level. For example, customary alpha levels are .05 and .01. An alpha level of .01 signifies greater confidence than an alpha level of .05. The t-statistic tells the analyst whether the slope coefficient, in our case b, is statistically significant (different from zero) and the degree of confidence in rejecting the null hypothesis, which states that the true value of b is zero. In our case, the t-statistic for the estimate of b is the ratio of the coefficient estimate ($\hat{b}$) to the standard error of the estimate $s_{\hat{b}}$:

$$t = \hat{b}/s_{\hat{b}} = .15/.032 = 4.69$$

The standard error of the estimate of b is denoted as $s_{\hat{b}}$. This is a measure of the variation in the estimate of b. Numerically, in this example,

$$s_{\hat{b}} = 0.032$$

mate of b equal to $.15, when in fact the true value of b is zero, is less than one percent. This is the sort of statement that one can make about the certainty associated with the regression results, in statistical language a "probabilistic statement." As is true for any regression result, one cannot say that POP is, in fact, a determinant of P. One can, however, make a probabilistic statement that the statistical result indicates that it is extremely likely that this is the case. The estimate of b is said to be a reliable predictor of profits or, in the language of statistics, "statistically significant." The null hypothesis is, therefore, rejected, and it can reasonably be inferred that POP is a determinant of profit.

The regression model generates a *point* estimate—or a specific numerical value—of b, which in this instance equals .15. This point estimate is our best guess as to the true coefficient. To get a feel for how precise this *point* estimate may be, one may turn to the "confidence interval" for this estimate. First, we start with the degree of confidence or certainty that we want, say 95 per cent, the common degree of confidence used by statisticians. Then we use the estimate of b and its standard error, which is a measure of the range of possible values for b, to construct the confidence interval.[74] As a result, we can make the following sort of probabilistic statement: If one takes repeated samples and calculates confidence intervals in the same fashion, then (i-a) per cent of those confidence intervals will enclose the true value of b.[75] The smaller the confidence interval, the more precise the estimate. In our example, the 95% confidence interval is

$$0.087 \leq b \leq 0.213.$$

While our point estimate was $.15, the bounds on the 95% confidence interval are approximately $.09 and $.21. In other words, we can say that we are 95% sure that the true value of b falls in this range.

There is another way to evaluate how well the estimated regression line fits the observed data. This is the coefficient of determination, which is usually denoted by R^2. This measure is calculated by breaking down the variation in the dependent variable (P) into its component parts and isolating the proportion of the variation in profit accounted for by the independent variable (POP).[76] In this example,

$$R^2 = 0.2447,$$

which means that the variation in population across the seventy-five restaurant locations explained about a quarter of the variation in profits across those locations. For a simple model like ours, one might consider this to

74. As a general matter, the confidence interval for the estimate of b is where $t_{n-2\,\alpha/2}$ is the value of a t-statistic with n-2 degrees of freedom that cuts off $\alpha/2$ of the area under the t distribution at each tail. Notice that in this expression the end points are random variables because b appears. For a clear explanation of the value of confidence intervals in resolving legal disputes, see generally Herbert Solomon, *Confidence Intervals in Legal Settings*, *in* STATISTICS AND THE LAW, 455-73 (M. DeGroot et al. eds., 1986). For further discussion, see generally D. H. Kaye, *Apples and Oranges: Confidence Coefficients and the Burden of Persuasion*, 73 CORNELL L. REV. 54 (1987).

75. *See* KMENTA, *supra* note 55, at 188-89.

76. Since R^2 is a proportion, it must assume values between zero and one. Furthermore, the addition of variables irrespective of their significance will "improve" (i.e., increase) the R^2 measure.

be a rather good performance. On the other hand, the model leaves about seventy-five per cent of the variation unexplained. A word of caution: preoccupation with indiscriminately "improving" the R^2, as some literature may unintentionally encourage, may be inconsistent with objective analysis and may actually decrease the reliability of the model if there is no scientific rationale for the steps taken to increase R^2.[77]

E. *Statistical Properties of Least Squares Estimators*

Under the assumptions of the least squares regression model, the resulting estimates have some desirable statistical properties.[78] First, the least squares regression estimators, i.e., the estimates of a and b, are *unbiased*, which means that the expected value (or the mean) of the estimator is equal to the true value of the population parameter. This is important because sample data rarely encompass the universe of possible values, known in statistics as the population. A sample is a subset of the population. In our case, the expected value of $\hat{a}$ is a and the expected value of $\hat{b}$ is b. This property is important because, on average, the least squares estimators are correct. Biased estimators, on average, will be wrong.

Second, the least squares regression estimators have a smaller variance than any other linear, unbiased estimator. In this sense, these estimators are deemed "best." In combination, these two properties define *efficiency*. If the estimator is unbiased, and the variance of the estimator is smaller than the variance of any other unbiased estimator, then the estimator is said to be efficient.

Third, the least squares regression estimators are *consistent*, which means that as the sample size increases, the bias (if any) approaches zero and the variance approaches zero.

F. *A More General Model*

In our example, it could be argued that the model is too simplistic. It contains only one explanatory variable, population within five miles of the restaurant location. Other variables influencing demand such as income, wealth, racial composition, educational attainment, and the like could also be important. This possibility cannot be ignored at the outset unless economic theory suggests that these variables should not be included. Accordingly, we return to the first step in the regression analysis; the use of economic theory to identify the relevant economic and demographic variables to include in the model.

---

77. *See, e.g.*, Glen A. Stankee, *Econometric Forecasting of Lost Profits: Using High Technology to Compute Commercial Damages*, FLA. BUS. J., June 1987, at 83, 85 (noting that "the higher the R-squared factor, the more reliable the predictions will be"). Stankee adds, however, that the R^2 can be inflated by simply adding more explanatory variables to the regression model, which encourages a "kitchen sink" approach to modelling. *Id.* As our article points out, what should be included in the model is a matter of economic theory.

78. For the technical developments of these properties, see generally KMENTA, *supra* note 55.

As we concluded earlier, profits are dependent upon the size of the local population. We therefore included the population within a five-mile radius of each restaurant site. In addition to the sheer size of the population, other characteristics of the population may be relevant. For example, eating in a restaurant is more expensive than eating at home. Thus, sales also may be influenced by the income and wealth of the local population. Accordingly, we may want to include measures of household income or per capita income and a wealth measure such as the value of housing in the area. Further, since meal preparation at home is time consuming and women may be disproportionately saddled with that responsibility, female work force participation rates may influence sales as well. To the extent that educational attainment is important in the decision to frequent the restaurant chain in question, we may want to include variables for median educational attainment in years and the percentage of the population that graduated from college. If racial composition seems important, one can include a variable indicating the percentage of the total population that is a given race. Finally, in our case, the restaurant is built in two configurations that can be categorized as "large" and "small." Those locations with large versus small restaurants must be identified to see whether the configuration of the restaurant is a determinant of success.

When we go to a more complete model by adding other independent variables, graphically plotting the resulting relationship becomes impossible because its dimensions cannot be captured on a flat surface. The conceptual basis of the techniques, however, remains the same. The expanded model may look like the following:

$$P = b_0 + b_1\,POP + b_2\,L + b_3\,MHV + b_4\,HHI + b_5\,MEA + b_6\,FWP + e$$

P = annual profit
POP = total population within a five-mile ring
L = 1 if restaurant is large and 0 otherwise
MHV = median housing value
HHI = median household income
MEA = median educational attainment in years
FWP = female work force participation rate
e = random error term

In order to estimate the regression parameters, that is, the coefficients b_0 through b_6, one must gather data on each of these variables for every restaurant in the sample. A computer program is then used to calculate the least squares regression estimates of the coefficients.

The results in this case are as follows:

$$P = \$47{,}078 + \$.187\,POP - \$1670\,L + \$2.87\,HHI + \$.60\,MHV - \$588\,MEA + \$98.55\,FWP$$[79]

79. The random error term has been omitted for the purposes of this hypothetical application.

All of the variables, except median educational attainment (MEA) and the dummy variable for large restaurants (L), are statistically significant at conventional alpha levels, i.e., $\alpha = .05$ or $\alpha = .01$. Furthermore, the expansion of the model has greatly improved our ability to explain the profit performance of the restaurants in the Jolly Rogers system: the R^2 is now .741, which means that about 74% of the variation in profit across Jolly Rogers restaurants is accounted for by variation in these variables.

The interpretation of the coefficients is slightly different. One infers that, after controlling for other variables, a one unit change in the variable under consideration will cause a change in profit equal to the coefficient. For example, when population (POP) increases by one person, annual profit (P) of the restaurant rises by $.187. When median household income (HHI) in the area rises by a dollar, restaurant profits increase by $2.87. Notice that increasing median educational attainment (MEA) causes restaurant profits to *decrease* by $588. We can avoid the conclusion that Jolly Rogers appeals mostly to the uneducated by noting that this coefficient is not statistically significant. In the case of the restaurant size, L is a so-called dummy variable because it can take on values of only zero and one. Its coefficient provides information on the differential effect, if any, that can be attributed to having the "large" restaurant. Here, we got a bit of a surprise because it appears that having a large configuration actually caused profits to *decrease* by $1670. This result could also be spurious since the coefficient was not statistically significant. When evaluating a multiple regression analysis and determining what estimates are reliable, one must focus generally only on those that are statistically significant.

G. *Estimation of Lost Sales: The Millers' Tale Resolved*

The regression coefficients were estimated on the basis of a sample of seventy-five restaurants in the Jolly Rogers system. This average relationship may be used to estimate the profits that the Millers lost in 1990. In order to do this, one must have the economic and demographic data for the Millers' location. Once the data are substituted into the estimated regression equation, an estimate of what the profits would have been but for the termination can be calculated. Assume the actual values for the Millers' location were as follows:

POP = 210,000
L = 1
HHI = 32,000
MHV = 84,000
MEA = 12.5
FWP = 45.3

Plugging these numbers into our equation reveals that the estimated profit for 1990 is $224,032.

For the years 1991 and 1992, the demographic data for the Millers' location may be different than those used to estimate the 1990 profit figure. If this is so, the 1991 and 1992 data must be substituted and the

estimated profit re-calculated for 1991 and 1992. To arrive at the Millers' total lost profits, one must add interest to the lost profits to compensate for the fact that they did not have the benefit of this money during 1990-1992.[80]

III. ADMISSIBILITY OF MULTIPLE REGRESSION MODELS

A multiple regression model is of little use to a litigant if it is deemed inadmissible. Generally, all relevant evidence is admissible.[81] Relevant evidence is defined to be evidence tending to make the existence of a material fact more or less probable than without the evidence.[82] Based on this definition of relevance and the previous discussion, it seems obvious that multiple regression analysis is relevant evidence. The amount of harm, if any, suffered by a litigant is often the centerpiece of a trial and almost always an important fact to be determined. Damage estimates that measure that harm are therefore unquestionably a material fact in a litigant's case. Multiple regression analysis, by its inherent nature, tends to prove the material fact of damages. Hence, so far, our multiple regression analysis seems judicially irresistible. Furthermore, multiple regression analysis is really nothing more than a scientific means of projecting lost profits based on an assortment of data. The projection method has long been recognized as a legitimate means of proving damages.[83]

Multiple regression analysis, however, is a statistical technique and in some respects the law has looked askance at statistical "proof." The principal reason for this jaundiced judicial view of statistics is probably predicated on the view that statistics may lie or be used to deceive. People may also lie or deceive. This does not, in and of itself, preclude the admissibility of oral testimony. We contend, therefore, that statistical evidence should be no different and consequently admissible, provided that the litigant seeking to have the statistical evidence admitted is able to conform to the common law rule that proof of damages must be reasonably certain and not speculative.[84]

Generally, the term "speculative" means that the evidence sought to be admitted is founded on illogical assumptions that have no basis in fact or experience. To be admissible, a multiple regression analysis cannot be

---

80. This is consistent with the well-established and generally accepted concept of the time value of money.

81. FED. R. EVID. 402. This article makes reference to the Federal Rules of Evidence on the assumption that most states have evidence codes similar to the Federal Rules.

82. FED. R. EVID. 401.

83. *See, e.g.*, Dolphin Tours, Inc. v. Pacifico Creative Serv., Inc., 773 F.2d 1506, 1511 (9th Cir. 1985).

84. *See* FED. R. EVID. 403 ("Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of issues, or misleading the jury, or by consideration of undue delay, waste of time, or needless presentation of cumulative evidence.")

For an outline of the legal principles governing recovery of lost profits and cases supporting the same, see generally ROBERT L. DUNN, RECOVERY OF DAMAGES FOR LOST PROFITS (4th Ed. 1992).

For a legal and economic analysis of the concept of "speculativenes," see generally Roger Blair & William Page, *supra* note 16.

speculative. In other words, the assumptions underlying both its theoretical construction and its data must be logical and based on generally accepted statistical principles. This is eminently reasonable when one considers that multiple regression analysis is proffered through the testimony of an expert. Since the expert is subject to cross-examination on this and all aspects of his testimony, the expert must be qualified to render a statistical opinion. Weaknesses in the assumptions or logic of a multiple regression analysis can be rooted out by cross-examination. Our judicial system relies heavily on cross-examination to do this in many other contexts—why not with statistical models? Moreover, if any expert's arithmetic or other mathematical computation is part of his testimony, he may have to explain his method and results. Likewise, a multiple regression analysis is nothing but a sophisticated mathematical analysis requiring an expert and a computer to perform it properly. Furthermore, the subject matter of this expert's testimony is discoverable, and should a litigant wish to counter the testimony of his opponent's expert with that of his own expert, he is free to do so. Such is the nature of the adversarial process.

Since juries may be wary of statistics, unless simplified by an expert, multiple regression analysis may appear hopelessly difficult. Therefore, serious consideration should be given to the method of its presentation. The results should only be presented by those intimately familiar with both statistical technique and the data at issue. Otherwise, the effectiveness of multiple regression analysis may be lost in a morass of technical, and perhaps largely irrelevant, difficulties. Furthermore, multiple regression analysis will often produce "common sense" results. The expert should therefore stress the logic of the model in light of everyday experience and impress them with the scientific method behind the model.

As a simple example, let us say using our previous illustration that Iam Bullsheet, a renowned expert, testifies that the Millers' profits should have been lower by $1670.00 because they had a large restaurant. Mr. Bullsheet should be subject to cross-examination both on that point (as it was within the scope of his direct testimony) and with respect to the statistical insignificance of the "large" restaurant coefficient and the apparent illogic of such a statement in light of everyday experience. More advanced cross-examination may relate to the assumptions discussed previously regarding the inherent statistical properties of the model and the sample data.

Another interesting observation of the judicial process is that far less reliable statistical evidence is admitted every day. For example, experts or other witnesses often are asked about the frequency of certain events occurring, the likelihood of their occurrence, the effect of their occurrence on something else or the existence (or lack thereof) of a causal relationship between events. These types of questions are all asking for a subjective probability assessment of events. They ask for information akin to that produced by our multiple regression analysis. Due to biases of all sorts, including human error in perception, such subjective probabilities may be seriously flawed in comparison with a truly scientific, statistical analysis of the events. Therefore, somewhat ironically, the judicial system that refuses

statistical evidence may forego far more reliable and objective evidence in favor of other forms of evidence that are less reliable and even biased.[85]

Multiple regression analysis is not a dubious statistical technique. Rather, it is well-recognized and widely used by economists and statisticians. The principles upon which it is based are generally accepted by practitioners and academics alike. Furthermore, numerous well-respected statistical treatises and textbooks deal in depth with this type of analysis.[86] Multiple regression analysis is not a suspect bit of hocus pocus designed to inflate (or deflate) figures to mislead a jury. To the contrary, it is a scientific method designed to rationalize the presentation of statistical evidence.

While this argument must be tailored to the facts and circumstances of a given case, it is imperative that the above points be made at trial when proffering statistical evidence of any kind. This will improve the likelihood of admissibility and preserve the issue for appeal.

CONCLUSION

In this paper, we have described a statistical method for creating a composite yardstick that is not plagued by the conceptual flaws of less sophisticated methods.[87] The composite yardstick also provides a way to estimate lost profits where it might otherwise seem too speculative to do so due to lack of historical financial data or a similar, yet independent, business. Thoughtfully employed, multiple regression analysis provides a rational technique for reliably estimating commercial harm to a business. It represents a technological advancement founded upon a long accepted method of calculating damages.

---

85. For a general text on the use of statistics in legal proceedings that includes a chapter on cases involving statistical evidence and demonstrates the wide breadth of applications, see DAVID W. BARNES & JOHN M. CONLEY, STATISTICAL EVIDENCE IN LITIGATION (1986) (especially Chapter Ten, *Cases Involving Statistical Evidence*, 545-95).

86. *But cf.* Fischer Black, *The Trouble with Econometric Models*, 35 FIN. ANALYSTS J., Mar.-Apr. 1982, at 3 (explaining the problem of confusing correlation with causation); David F. Hendry, *Econometrics—Alchemy or Science?*, 47 ECONOMICA 387 (1980) (noting nonsense conclusions drawn from simple correlations, unreliable data, and various errors of linear regression models); Edward E. Leamer, *Let's Take the Con out of Econometrics*, AM. ECON. REV., Mar. 1983, at 31, 33 (noting the limitations and somewhat arbitrary character of statistical inferences).

87. For a seminal work on the subject of composites, see MARY SHELLEY, FRANKENSTEIN (1818). Baron Frankenstein's innovative composite was unprecedented, but it proved troublesomely uncontrollable and was ungratefully determined to turn against its creator. Undoubtedly, Shelley intended her work as an allegory about brilliant but short-sighted expert witnesses under cross-examination.

FREEDOM AND CULTURE: WHY WE SHOULD NOT BUY COMMERCIAL SPEECH

R. GEORGE WRIGHT*

INTRODUCTION

It is not surprising that a commercial culture would grant special constitutional protection to commercial speech. There is, however, no reason to suppose that special protection for commercial speech in our cultural context promotes the overall freedom and well-being of the public. A strong case can be made that in our culture, reducing the constitutional protection of commercial speech would actually tend to promote freedom and well-being in the long term. This is not because some commercial speech is false, fraudulent, or deceptive. The focus of attention should instead be on non-deceptive commercial speech, framed in the broader context of our culture.

Because this Article's argument is broad in scope, it begins by establishing some perspective before focusing on the current case law of commercial speech. Thus Section I below briefly surveys some of the classic discussions of commodity consumption and well-being. The Article then turns to the contemporary social science literature. This literature suggests that for most contemporary Americans, there is actually only a minimal relationship, if any, between consumption and well-being.

Section II discusses further the effects of commercial speech—in particular, the dominance, within its sphere, of commercialism in our cultural context. Commodification and commercial speech are pervasive in our society. Section II traces the implications of this state of affairs for freedom of speech generally.

Most important for free speech doctrine is the absence, in our culture, of any meaningful institutional challenge to the influence of commercialism and commercial speech. No cultural institution is able, or inclined, to provide significant "counterspeech" to the broad "message," intended or unintended, of commercial speech. Any reasonable regulation of commercial speech, whatever its more particular justification, actually tends, at least minimally and indirectly, to contribute to freedom and well-being. Such regulation has this effect by implicitly raising the issue of the proper role of commercialism and commercial speech in our society. There is, at a minimum, nothing in the Free Speech Clause that should bar a society from democratically acting on these beliefs, along with any other appropriate grounds for reasonably regulating commercial speech.

---

* Professor of Law, Cumberland School of Law, Samford University. The author thanks Richard Delgado, Jay Mootz, and Steve Smith for their thoughtful reactions to previous drafts.

This does not mean that government itself can be expected to counterbalance the influence of commercialism, or that implicitly questioning the cultural sovereignty of commercialism by itself establishes viable alternatives to commercialism. Reasonable regulation of commercial speech is merely one step in the process of legitimizing and democratically facilitating free choices of less commercial styles of life, in ways consistent with the Free Speech Clause and the rest of the Constitution.

With this basic argument in place, Section III surveys the most important discussions of the constitutional status of commercial speech. The commentators have been hopelessly split on whether or how to protect specially commercial speech. This dispute stems largely from commentators' inclination to insist on either a broad or a narrow range of values or purposes underlying the Free Speech Clause, and to define each of those values or purposes in ways friendly or unfriendly to commercial speech. This Article seeks to avoid this trap. One way of bypassing this interminable conflict is to adopt the broader cultural institutional focus I have introduced above. Thus this Article does not rest on a narrow, controversial view of why we value free speech in the first place.

To this point, I have relied on intuitive, uncontroversial ideas of the meaning of commercial speech. Classifying speech as either commercial or non-commercial often will be easy. But what if the idea of commercial speech turns out to be so complex or contestable as to be unusable in practice? Section IV is intended to allay such fears and offers a method for minimizing the cost of judicial errors in misclassifying commercial and non-commercial speech in borderline cases.

Section V then considers in some detail the case law establishing the degree of special protection currently accorded commercial speech. As matters stand, the constitutional tests imposed on regulations of commercial speech are unreasonably demanding or, at best, so inescapably subjective and indeterminate as virtually to invite an unsympathetic court to strike the regulation down. This state of the law does not serve the public interest.

Finally, the Conclusion expands the theme of the preceding section, warning that the level of protection currently given to commercial speech jeopardizes reasonable regulation of that speech in a wide variety of subject matter areas. Whether the benefits of striking down reasonable regulation of commercial speech in these areas outweigh the social costs of doing so is, in our cultural context, doubtful in the extreme.

I. WELL-BEING AND COMMERCIAL SPEECH

The relationship between acquiring consumer goods and genuine happiness or well-being has long been doubted. Classically, for example, Epictetus argues that "if you have not the want of riches, know that you possess more than this [rich] man possesses and what is worth much

more."[1] Epictetus raises the possibility that the desire for riches and the objects acquired from riches may be self-defeating and pathological.[2] Epictetus argues as well that the costs of consumption may take the form not merely of the loss of leisure or some other consumer good or service, but of an unintended, unforeseen change for the worse of one's character.[3] A lifestyle emphasizing consumption is said to transform one's character adversely, in ways that the consumer does not recognize even after the fact.[4] Thus Epictetus denies that a lifestyle emphasizing consumption is typically rational, in the sense of being chosen with conscious awareness of its most important costs.

Skepticism as to the relation between consumption and happiness is not unexpected in classical or medieval writers, but similar sentiments are expressed from surprising quarters. Consider, for example, the observation of Adam Smith, a writer not insensitive to the power and virtues of economic markets:

> [W]ealth and greatness are mere trinkets of frivolous utility, no more adapted for procuring ease of body or tranquility of mind than the tweezer-cases of the lover of toys; and like them too, more troublesome to the person who carries them about with him than all the advantages they can afford him are commodious.[5]

Rousseau develops this theme in discussing the conversion of luxuries into "negative" necessities that have largely lost the power to please, but whose absence creates unhappiness.[6] Following Rousseau's lead, Immanuel Kant argues that "with growing wealth we acquire fresh wants, and the more we satisfy them the keener becomes our appetites for more."[7] Henry David Thoreau amplifies this theme in positing that "[m]ost of the luxuries, and many of the so-called comforts of life, are not only not indispensable, but positive hindrances to the elevation of mankind."[8]

Thoreau suggests two additional relevant points. First, his reference to "so-called comforts" suggests with Rousseau that our inevitable adjustment to a good means that we derive only diminishing satisfaction from that good, while becoming more vulnerable to its loss and to whatever

---

1. THE DISCOURSES OF EPICTETUS book IV, ch. 9, at 399 (George Long trans., A.L. Burt ed. 1885); *see also* THE ANALECTS OF CONFUCIUS book VII, § 15, at 126 (Arthur Waley trans. 1938) ("The Master said, He who seeks only coarse food to eat, water to drink and bent arm for pillow, will without looking for it find happiness to boot.").

2. *See* THE DISCOURSES OF EPICTETUS, *supra* note 1, at 399.

3. *See id.* at 400.

4. *See id.*

5. Adam Smith, *Theory of the Moral Sentiments*, *in* 1 BRITISH MORALISTS 309 (part IV, ch. 1) (L.A. Selby-Bigge ed. 1897) (Dover 1965).

6. JEAN J. ROUSSEAU, A DISCOURSE ON INEQUALITY part II, at 113 (Maurice W. Cranston trans., Penguin 1984) (as commodities "degenerated into actual needs, being deprived of them became much more cruel than the possession of them was sweet; and people were unhappy in losing them without being happy in possessing them").

7. IMMANUEL KANT, LECTURES ON ETHICS 7 (Louis Infield trans., Hackett 1992) (1930) (citing Rousseau); *see also* STANLEY LEBERGOTT, PURSUING HAPPINESS 69 (1993) (quoting EMILE DURKHEIM, SUICIDE 249 (1952)).

8. HENRY D. THOREAU, WALDEN AND OTHER WRITINGS 13 (Brooks Atkinson ed., Modern Library 1950).

disutility may attach to our fear for its loss. Thus Thoreau, Rousseau and Kant hint at a loosely "addictive" quality of consumption. But Thoreau does not see addiction as inevitable. Thoreau argues instead that once we have obtained what is necessary for life, "there is another alternative than to obtain the superfluities; and that is, to adventure on in life now, [one's] vacation from humbler toil having commenced."[9] For Thoreau, the alternative to endless acquisition and consumption is not asceticism, self-denial, or nirvana, but an affirmative, active use of one's liberation.

Early Marx was struck by the almost miraculous transformative powers of wealth and the acquisitions made possible by wealth.[10] Marx was equally interested in the ideas of authenticity and alienation in its various forms.[11] Authentic, non-commodified bases and standards of human relationships and human standards may, in Marx's view, become historically possible. Some of these themes have been developed by later writers under the ambiguous[12] rubric of "commodity fetishism."[13] In an informal sense, commodity fetishism refers to the tendency for the sphere of market exchange of goods and services to expand in ways destructive to the fullest and highest development of personality.

It is possible that market transactions tend, at least among the immediate parties, to in some sense maximize wealth.[14] The theorists of commodity fetishism may respond, however, that it is a further and fallacious step to infer from this that human well-being in the broadest sense is most fully realized when social relationships are converted into market-based relationships, or when the logic of the commercial market affects the nature of those social relationships.

A society's attempt to satisfy the full range of its needs through commodity exchange has been described as an aspect of the process of reifica-

---

9. *Id.* at 14; *cf.* Ralph W. Emerson, *Thoreau, in* THE OXFORD AUTHORS: RALPH WALDO EMERSON 475, 477 (Richard Poirier ed. 1990) ("He chose to be rich by making his wants few."); WILLIAM WORDSWORTH, 1 THE POEMS 568 (John O. Hayden ed., Penguin 1990) ("Getting and spending, we lay waste our powers.").

10. KARL MARX, EARLY WRITINGS 191-93 (T.B. Bottomore trans., 1964) (Third Economic and Philosophical Manuscript).

11. *See, e.g., id.* at 192-94; *cf.* JACQUES ELLUL, MONEY AND POWER 76-79 (LaVonne Neff trans., 2d ed. 1984).

12. *See* Margaret J. Radin, *Market-Inalienability*, 100 HARV. L. REV. 1849, 1872 n.85 (1987) (referring to a non-technical sense of commodity fetishism in which one is said "to have one's identity too tied to possessions, to be too dependent upon thing-ownership for pleasure and a sense of self-worth"). We should, of course, recognize the increasingly important role of commercial services as opposed to commercial goods in our culture. Consider purchases of a wide range of advertised commercial services that in other cultures would be unnecessary because of the depth of genuine friendships. In our culture, friendship may itself be commercialized, and thereby impaired or tainted, or left undernourished.

13. Commodity fetishism in a non-technical sense has been identified simply as "the continuing emphasis on 'transactions' under capitalism." William J. Wagner, *The Contractual Reallocation of Procreative Resources and Parental Rights: The Natural Endowment Critique*, 41 CASE W. RES. L. REV. 1, 148 n.654 (1990).

14. *See, e.g.*, Martin H. Redish, *The First Amendment in the Marketplace: Commercial Speech and the Values of Free Expression*, 39 GEO. WASH. L. REV. 429, 433 (1971) (maximum information regarding competing products as necessary in order to maximize "material satisfaction").

tion.[15] The reification of commodities, importantly, does not reflect the conscious, knowledgeable decision or intention of individuals or of the public collectively. It has been said that "[t]he commodity form . . . necessarily functions independent of, or autonomously from, the will of the subjects who set it in motion."[16] On this view, the commodity form, and the role of commodity exchange, tend to take on lives of their own.[17] To the extent that this is so, persons can hardly be said to be sovereign, free, and autonomous with regard to the scale and scope of commodity exchange.

Thus pervasive commodification is not simply the embodiment of free choices. Equally importantly, though, commodification tends to be self-justifying. The very pervasiveness of commodity exchange tends to make any departure from commodity exchange as a way of fulfilling human needs and potentials seem unnatural, utopian, or simply inconceivable.[18] Thus the free market exchange of goods and services is not the unequivocal embodiment of human development, human freedom, and fulfilled intention.[19]

One should consider as well how the early capitalists described by Max Weber, driven by other-worldly considerations, would have reacted to the historical shift in focus from capital accumulation to consumption, however inevitable such a shift may have been.[20] By the early part of the twentieth century, writers such as Thorstein Veblen[21] and R.H. Tawney[22] deemed a significant portion of the production and consumption of goods as "waste," not in the sense that such goods were unwanted,[23] but that consumers had unnaturally learned to want such goods,[24] even though those goods did not genuinely serve human well-being.[25]

---

15. *See* GEORGE LUKÁCS, HISTORY AND CLASS CONSCIOUSNESS 91 (Rodney Livingstone trans., 1971); *see also* HERBERT MARCUSE, NEGATIONS: ESSAYS IN CRITICAL THEORY 172-73 (Jeremy J. Shapiro trans., 1968).

16. Isaac D. Balbus, *Commodity Form and Legal Form: An Essay on the "Relative Autonomy" of the Law*, 11 LAW. & SOC'Y REV. 571, 574 (1977) (emphasis deleted).

17. *See id.*; Marlin H. Smith, *The Limits of Copyright: Property, Parody, and the Public Domain*, 41 DUKE L.J. 1233, 1272 n.202 (quoting A DICTIONARY OF MARXIST THOUGHT 86-87 (Tom Bottomore et al. eds., 1983)).

18. *See* Duncan Kennedy, *The Role of Law in Economic Thought: Essays on the Fetishism of Commodities*, 34 AM. U. L. REV. 939, 991-92 (1985).

19. *Cf.* MAX WEBER, 1 ECONOMY AND SOCIETY 351 (Guenther Roth & Claus Wittich eds., 1978) ("[t]he various modes of want satisfaction, always the result of struggles between different interests, often exert a far-reaching influence beyond their direct purpose").

20. *See* MAX WEBER, THE PROTESTANT ETHIC AND THE SPIRIT OF CAPITALISM 172 (Talcott Parsons trans., Scribner's 1958); *see also* THOMAS S. ROBERTSON, CONSUMER BEHAVIOR 102 (1970).

21. *See* THORSTEIN VEBLEN, THE THEORY OF THE LEISURE CLASS (Mentor 1953) (1899).

22. *See* R.H. TAWNEY, THE ACQUISITIVE SOCIETY (1920).

23. *See* VEBLEN, *supra* note 21, at 78.

24. *See id.* at 69.

25. *See id.* at 78; TAWNEY, *supra* note 22, at 37-38.

These themes were in turn developed by John Maynard Keynes[26] and at length by John Kenneth Galbraith.[27] Nevertheless, it is still standard for anyone trained in economics to think of consumer choices, in the absence of force and fraud, as revealed free preferences, and to think of non-fraudulent commercial speech as contributing to the efficiency of markets and to wealth maximization.

The contemporary discussion of the relationships among consumption, commercial speech, and well-being is addressed below, but it is useful to bear in mind two preliminary points. First, it is not inconsistent for someone to both wish to lead a life emphasizing consumption, and to be chagrined or embarrassed if too many other people make the same choice. By analogy, confining one's charitable donations exclusively to, say, Amnesty International does not imply that one should be pleased if everyone else does so as well. And second, we must not take the language itself of free markets, freedom of commercial speech, and revealed preferences to mean that we need not examine the actual effects of those institutions on human freedom, happiness, and well-being. Whether contemporary commercial speech actually promotes freedom and well-being is a partly normative and partly empirical question, and not simply a matter of the definitions of free markets and commercial speech.

The relationships among commercial speech, consumption, happiness, and freedom are doubtless complex. We may be tempted to think of them as simply mutually supporting, without conflict. But we may also suspect that matters may be more complicated than this, at least in today's highly developed economies. The novelist and philosopher Iris Murdoch argues, for example, that "modern industrial society, with all its vast diversity of entertainments and mass of incoherent information (of which television may serve as image and example), has radically changed people's lives and mode of well-being, bringing some benefits and doing much damage."[28]

An important link in the argument for the protection of commercial speech, and of advertising in particular, is that such speech promotes wealth maximization, and thus satisfaction and well-being. But the linkage between wealth maximization and well-being or happiness, at least for our society, in our time, should not be established merely by definition. Let us then consider, for our culture, the evidence bearing upon the assumed linkage between wealth, consumption, and well-being.

---

26. *See* John M. Keynes, *Economic Possibilities for Our Grandchildren*, *in* ESSAYS IN PERSUASION 358, 365 (1963) (referring to needs "which are relative in the sense that we feel them only if their satisfaction lifts us above, makes us feel superior to, our fellows").

27. *See* JOHN K. GALBRAITH, THE AFFLUENT SOCIETY 120 (1958) ("Nothing in economics so quickly marks an individual as incompetently trained as a disposition to remark on the legitimacy of the desire for more food and the frivolity of the desire for a more elaborate automobile."); *see also* JOHN K. GALBRAITH, THE CULTURE OF CONTENTMENT 8 (1992) ("One may marvel at the attraction of often frivolous and dispensable consumer artifacts and entertainments in our time, but their ultimately controlling appeal cannot be doubted.").

28. IRIS MURDOCH, METAPHYSICS AS A GUIDE TO MORALS 372 (1992).

In some respects, it might be improper to treat wealth or income as an adequate proxy for consumer spending. But given the remarkably low contemporary American private savings rate[29] and the evidence that over 90% of American incomes are devoted to consumption,[30] it seems reasonable to draw upon the literature discussing the relation between income or wealth and reported happiness.

For many economists, voluntary consumption of non-defective goods is not the beginning, but the end of any inquiry into the relationship between consumption and happiness.[31] This approach in turn reinforces the process by which "the market culture teaches us that money is the source of well-being."[32] Many of us suppose that we would be significantly happier, over the long term, if we were able to afford to buy some particular set of commodities. Based on the available social science evidence, however, this belief is likely wrong.

It has been observed, for example, that "[w]hile GNP and material standards of life have advanced substantially in the post-World War II decades, the extent of happiness and the rate of subjective satisfactions have not."[33] In particular, real income rose quite substantially in the United States in the period from 1957 to 1973, but reported levels of satisfaction declined slightly during this period.[34] During this same period, interestingly, the percentage of those describing themselves as "very happy" steadily decreased, and this decrease was most evident among the most affluent.[35]

It is tempting to dismiss these findings as largely a reflection of unique historical events. Naturally, it may be thought, a pre-Sputnik America would, all else equal, be happier than a post-Vietnam era America. To this, several replies may be made. First, the trends referred to above have not reversed themselves in the years following 1973.[36] As the data is extended in time, it becomes increasingly difficult to dismiss as aberrant or artifactual. Second, trying to explain the data on reported happiness in terms of broader cultural and historical events comes perilously close to undermining, rather than rescuing, the notion that wealth and consumption lead to happiness. And finally, the weak, limited relationship between income or expenditures and happiness is supported from a number of angles.

---

29. *See generally* THE U.S. SAVINGS CHALLENGE: POLICY OPTIONS FOR PRODUCTIVITY AND GROWTH (Charles E. Walker et al. eds., 1990); LAWRENCE H. SUMMERS & CHRIS CARROLL, WHY IS U.S. NATIONAL SAVING SO LOW? (1987).

30. *See* LEBERGOTT, *supra* note 7, at 65.

31. *See* Robert E. Lane, *Does Money Buy Happiness?*, 113 PUB. INT. 56, 61 (1993).

32. *Id.*

33. GEORGE KATONA, PSYCHOLOGICAL ECONOMICS 363 (1975).

34. *See* RONALD INGLEHART, THE SILENT REVOLUTION 116 (1977).

35. *See* Angus Campbell, *Subjective Measures of Well-Being*, 31 AM. PSYCHOLOGIST 117, 118 (1976).

36. *See* Alan T. Durning, *Are We Happy Yet? How the Pursuit of Happiness is Failing*, 27 FUTURIST 20, 21 (1993) (percentage of those reporting themselves as "very happy" fluctuating but stable at roughly one-third since the mid-1950s).

In fact, there is evidence suggesting that the linkage between income and happiness is growing weaker over time in the United States.[37] In any event, a number of surveys illustrate the weak relationship between money and happiness.[38] The evidence extends far beyond the American context. Robert Lane concludes, for example, that "[i]n almost all developed countries there is no substantial relation between income and well-being."[39] And the evidence in the American context can be particularized. For example, a 1973 study indicated that despite regional differences in income, when Southern state residents are "asked to evaluate the quality of their lives, they are modestly but consistently more positive than people living in the other major regions of the country."[40] A survey of 22 large lottery winners "found no clear difference between their happiness and that of controls."[41] Nor does wealth tend to be associated with reduction in the amount of worrying; instead, "it simply changes the subject"[42] of the worrying.

As a general rule, then, with some qualifications, happiness does not appear to be strongly linked with income and wealth, or, presumably, consumer expenditures guided, facilitated, or prompted by commercial speech. Turning from survey data to introspection confirms key elements of this conclusion. The philosopher James Griffin observes that "[i]t is depressingly common that when even some of our strongest and most central desires are fulfilled, we are no better, even worse, off."[43]

As we have seen, the typical inability of consumption to generate lasting satisfaction has been variously diagnosed. When discussing consumption-based lifestyles, contemporary writers tend, again consistent with introspection, to recur to concepts such as self-defeatingness[44] and the "hedonic treadmill,"[45] to the phenomenon of being trapped,[46] and of being "addicted."[47] This literature thus helps quantify or make more precise the observations of some of the historical writers discussed above. Thus, for example, based upon his important survey work, the political scientist Ronald Inglehart concludes that "[w]hile increased prosperity may produce a short-term sense of gratification, an individual gradually adjusts his aspiration level to his external circumstances; after a certain time lag, one takes a given level of prosperity for granted and aspires to more."[48] This result should not be surprising.

---

37. *See* MICHAEL ARGYLE, THE PSYCHOLOGY OF HAPPINESS 94 (1987).

38. *See id.* at 93 (citing several studies).

39. Lane, *supra* note 31, at 57.

40. Campbell, *supra* note 35, at 118.

41. ARGYLE, *supra* note 37, at 97.

42. Lane, *supra* note 31, at 60.

43. JAMES GRIFFIN, WELL-BEING 10 (1986).

44. *See, e.g.*, Richard A. Easterlin, *Does Money Buy Happiness?*, 30 PUB. INTEREST 3, 10 (1973).

45. *Id.*; Lane, *supra* note 31, at 63.

46. Easterlin, *supra* note 44, at 10.

47. *Id.*; TIBOR SCITOVSKY, HUMAN DESIRE AND ECONOMIC SATISFACTION 118 (1986).

48. INGLEHART, *supra* note 34, at 147.

If getting and spending, perhaps via commercial speech, does not significantly affect happiness, there are stronger influences on happiness. The evidence suggests that happiness is more crucially affected, generally, by "social relations, work, and leisure."[49] These sources of satisfaction often do not depend upon market-based consumption of goods and services.[50] Nor, for that matter, do we tend to find genuine friendship, appropriately stimulating work, and leisure in general to parallel consumption as an entrapping, self-defeating "treadmill."[51] If wealth is subject to the law of diminishing returns,[52] genuine friendship does not seem to be similarly vulnerable.[53]

Admittedly, it is not easy to measure attributes like the "quality" of work or friendship in a rigorous way. But measurement problems do not obscure the basic message. Incidentally, the evidence seems to suggest that "informal visits between neighbors and friends, family conversation, and time spent at family meals have all diminished in the United States since mid-century,"[54] though this may have been offset by more interesting work for some persons.

Why we remain on the hedonic treadmill need not be fully answered here. It may be that at least some relationships, creative work, and productive, potentially enjoyable leisure activity require investments, initial sacrifices, or training which we are now unable or unwilling to undertake.[55] But a more important explanation, it would seem, for remaining on the hedonic treadmill may be more broadly institutional, with commercial speech at the heart of that explanation.

Before taking up the broader role of commercial speech in our cultural context, though, one should acknowledge the possibility that the genuinely poor, whether conceived of as a relatively poor nation-state, or as a segment of a particular society, would be significantly better off at higher consumption levels, or would, more particularly, be happier as a result of uninhibited non-deceptive commercial speech.

As it turns out, there is some controversy over whether poor societies are significantly unhappier than rich societies.[56] In any event, as the United States presumably does not fall into the former category, broadly increasing our collective happiness through additional consumer spending seems unlikely. More interesting is the relationship between income and happiness among the poor within our own society. This precise ques-

---

49. Durning, *supra* note 36, at 20; *see also* Lane, *supra* note 31, at 63.

50. *See* SCITOVSKY, *supra* note 47, at 119; *see also* Sut Jhally, *Commercial Culture, Collective Values and the Future*, 71 TEX. L. REV. 805, 809 (1993) ("[A] market-based society has a tendency to push people towards those things that it can provide—goods and services—while the real sources of satisfaction are outside the capability of the marketplace to provide.").

51. *See* Lane, *supra* note 31, at 63.

52. *See, e.g.*, Ruut Veenhoven, *National Wealth and Individual Happiness*, *in* UNDERSTANDING ECONOMIC BEHAVIOR 9, 19 (Klaus G. Grunert & Folke Ölander eds., 1989).

53. *See* Lane, *supra* note 31, at 63.

54. Durning, *supra* note 36, at 22.

55. *See* SCITOVSKY, *supra* note 47, at 123.

56. *Compare, e.g.*, Lane, *supra* note 31, at 56-57 and sources cited therein *with* ARGYLE, *supra* note 37, at 94 *and* Easterlin, *supra* note 44 (work is re-examined in Lane, *supra* note 31).

tion must be kept in perspective. We plainly ought to take relief of poverty far more seriously than we do, whether this would be politically popular or not. This is a matter of justice and basic moral principle quite independent of survey evidence of subjective satisfaction among the poor.[57]

With this understanding, it is certainly plausible to argue that the declining marginal utility of wealth[58] is of scant relevance to the poor, and that money can at least reduce some forms of brute distress, such as relieving cold or hunger.[59] It is thus not surprising that people in poverty within our society tend to report greater unhappiness.[60]

Assuming, then, some negative correlation between poverty within the United States and happiness, the question for our purposes then becomes whether the poor can be made significantly happier through a regime of specially constitutionally protected commercial speech. In some respects, this is a difficult question. But we must not lose sight of the obvious. A poor person benefits far more from a warm coat than from commercial speech about coats, even when the indirect benefits of unrestricted commercial speech in enhancing quality and driving down prices are considered.

Doubtless the poor may, to take one example, benefit from prescription drug price advertising, directly or indirectly, at least under some health insurance and welfare policies. But there are costs for the poor as well. It is possible that prescription drug price advertising may affect the commercial viability of small, independent pharmacies owned by persons with a long-term stake in a poor community. There may be some tradeoff between price and the availability of a pharmacist with ties to the poor community who is able and inclined to treat the local poor as individuals, who knows their broader needs, and who can establish the kind of personal, caring relationship of the sort that existing evidence suggests is important to happiness.[61]

As well, we must consider that as constitutional protection for commercial speech increases, so may the potential for commercial speech to manipulate and distract the poor, without violating whatever restrictions may exist on fraudulent, deceptive, or misleading commercial speech. Along with useful information about accessible goods in accessible places comes the broader, phantasmagorical effects of contemporary advertising as well.

---

57. *See, e.g.*, KANT, *supra* note 7, at 42-43.

58. *See* Veenhoven, *supra* note 52, at 19.

59. *See* ARGYLE, *supra* note 37, at 93.

60. *See id.* at 94; Lane, *supra* note 31, at 57; Campbell, *supra* note 35, at 121.

61. *See generally* Durning, *supra* note 36 and Lane, *supra* note 31 and accompanying text. For discussion of similar issues without special reference to the poor, see the several opinions in Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S. 748 (1976).

II. The Effects of Commercial Speech

One need not deny the beneficial effects of commercial speech generally or of advertising in particular. However maldistributed wealth may be, there is plainly more of it, in some obvious sense, in a materialist-oriented consumer economy.[62] Commercial speech, including various forms of advertising, obviously plays a significant role in this process. This discussion does not minimize the plight of the poor, who plainly are better off with consumer goods such as clothing and food.[63] More broadly, the society as a whole clearly benefits from rapid dissemination of accurate information concerning, for example, a safe and uniquely effective drug treatment for a serious illness.[64]

As well, it is difficult to deny the adverse consequences on wages and employment if many of us tomorrow abandoned familiar consumption levels in favor of communing with nature or verse composition.[65] Of course, my proposal to accord no special constitutional protection to commercial speech does not seem likely to lead to any such consequences. Those persons who eventually discover themselves in possession of excess income and wealth may reasonably be advised to enhance the effective market demand of the poor.

Nor should we deny that the very activity of searching for, selecting, and acquiring commercial goods and services may itself be utility-enhancing.[66] On the contrary, "it is a cliche to say that the best way to deal with being depressed is to go shopping."[67] Whether this is actually the best path to well-being seems, however, given the evidence and analyses discussed above,[68] rather doubtful. Or so a reasonable government might come to believe.

Neither is it necessary for us to endorse any distinction between natural, healthy, and authentic consumer needs and artificial, unhealthy, or contrived consumer needs.[69] To some degree, as Jean Baudrillard has argued, modern advertising often tends to break down any traditional distinction between authentic and contrived, artificial responses to advertisements.[70] It might be noted, though, that if any distinction between natural and contrived responses to advertising tends to dissolve, it

---

62. *See* Katona, *supra* note 33, at 363.

63. *See, e.g.*, Erich Fromm, The Revolution of Hope: Toward a Humanized Technology 120 (Ruth N. Anshen ed., 1968).

64. *See* Keith B. Leffler, *Persuasion or Information? The Economics of Prescription Drug Advertising*, 24 J.L. & Econ. 45, 74 (1981).

65. *See* Fromm, *supra* note 63, at 131.

66. *See* Christopher Lasch, The True and Only Heaven: Progress and Its Critics 521-22 (1991) [hereinafter Lasch, The True and Only Heaven]. *But see* Robert E. Lane, The Market Experience 469 (1991).

67. Stuart Ewen, *Advertising and the Development of Consumer Society*, *in* Cultural Politics in Contemporary America 82, 85 (Ian Angus & Sut Jhally eds., 1989).

68. *See supra* text accompanying notes 1-61.

69. *See, e.g.*, Fromm, *supra* note 63, at 124; Jerry Mander, Four Arguments For the Elimination of Television 125 (1978) (citing, perhaps controversially, an electric hair dryer as a contrived, artificial, implanted need).

70. *See* Jean Baudrillard, *The Masses: The Implosion of the Social in the Media*, 16 New Literary History 577, 578-79 (Marie McLean trans., 1985).

may become more difficult to specify how commercial speech distinctively serves any of the values typically thought to underlie the free speech clause.[71] If, for example, our responses to ads tend to be neither classically manipulated nor authentic, it may be hard to say why that murky state of affairs is most consistent with genuine autonomy.

Finally, we need not ascribe to commercial speech, and to commercial advertising in particular, powers they do not possess. Plainly, neither particular advertisements nor broader ad campaigns are invariably effective. It is reported, for example, that "about 70% of test-market brands are not expanded nationally and can therefore be classed as failures."[72] The ability of advertising to influence choice among competing brands often does not translate into effective influence over whether one buys that general kind of good at all.[73] And even when advertising affects behavior, the effect may well be temporary.[74]

None of this implies, however, that commercial advertising is without significant, long-term, intended or unintended effects on American culture and decisionmaking. It is admittedly true that advertisements often compete against each other,[75] with one product's gain being another product's loss. But the proliferation of ads does not amount simply to a process of mutual annihilation. Importantly, advertisements that conflict or compete at one level may, at another level, mutually reinforce one another. Such effects need not be intended or even recognized. Consider, by way of loose analogy, that in the natural world, two or more separate waves may mutually interfere or tend to cancel each other if they arrive out of phase, or they may tend to reinforce one another,[76] thereby in some respect heightening their potential impact.

At a fairly specific level, an ad for a particular drug, as "reinforced" by other ads for competing and non-competing drugs, may tend in our cultural context to promote drug ingestion more generally as a response to medical, psychological, and even social problems. As one former participant in the process has written, "[w]hile it might matter to Upjohn or Cutter Laboratories which drug a consumer buys, both are in agreement that they benefit whenever people seek any drug rather than a nondrug solution to a problem."[77] More precisely, though, this mutual reinforce-

71. *See infra* note 78; *see also* Ronald K.L. Collins & David M. Skover, *Commerce & Communication*, 71 TEX. L. REV. 697, 712 (1993) (noting commercial ads have "helped devalue the coin of communication by developing a massive, unthinking tolerance for nonsense" (quoting LEO BOGART, STRATEGY IN ADVERTISING 7 (2d ed. NTC Business Books 1984) (1967))).

72. J. Hugh Davidson, *Why Most New Consumer Brands Fail*, HARV. BUS. REV., Mar.-Apr. 1976, at 117; *see also* MICHAEL NOVAK, THE SPIRIT OF DEMOCRATIC CAPITALISM 108 (1982) ("The history of advertising is full of quirks and failures.").

73. TIBOR SCITOVSKY, THE JOYLESS ECONOMY 205 (1976).

74. *See* Lester D. Taylor & Daniel Weiserbs, *Advertising and the Aggregate Consumption Function*, 62 AM. ECON. REV. 642, 650 (1972).

75. NOVAK, *supra* note 72, at 108.

76. *E.g.*, NICK HERBERT, QUANTUM REALITY 74 (1985).

77. MANDER, *supra* note 69, at 126.

ment and "generalization" process would seem to operate whether it is intended or even recognized by any party, including the audience.[78]

More broadly, even when commercial ads compete, or fail to sell a particular product, they may at least inadvertently, through their mutually reinforced cumulative impact, legitimize and support commercial consumption as a style of life and a solution to life's problems.[79] Thus it is only a harmless exaggeration to say that "[a]dvertising serves not so much to advertise products as to promote consumption as a style of life."[80]

Thus while consumers reject or ignore most of the particularized ad messages they receive, as would inevitably be the case, they may be less able to resist the unintended broader "message" of the superiority of commercial consumption as a basic approach to living.[81] This assumption reflects a realistic assessment of the contemporary balance of cultural forces. Consider that one writer has estimated that by the age of retirement, the typical American will have seen at least two million television commercials.[82] This is a remarkable figure. Children's television programming, where basic preferences might be formed, commonly seeks to merge substantive programming and commercial huckstering.[83] The vector of forces tending to engender commercial consumption is, plainly, powerful.

By itself, the power of commercial speech to shape inadvertently our culture might not be so troubling, were it not for the fact that today, in our cultural context, there is no realistic prospect for effective "counterspeech" tending to promote noncommercial approaches to life's problems and opportunities. In our cultural circumstances, no institution currently devotes any real energy or resources to provide a counterspeech remedy[84] for the implicit message of our commercial culture.[85] Of course, some

---

78. Actually, the case for restricting commercial speech is additionally strengthened to the degree that the typical ad does not reflect an actual intent to promote an ethos of materialism or consumption, or any other broadly social idea. *See, e.g.*, R. GEORGE WRIGHT, THE FUTURE OF FREE SPEECH LAW 1-31 (1990).

79. *See* Durning, *supra* note 36, at 22.

80. CHRISTOPHER LASCH, THE CULTURE OF NARCISSISM 137 (1979) [hereinafter LASCH, THE CULTURE]; *see also* LASCH, THE TRUE AND ONLY HEAVEN, *supra* note 66, at 518.

81. *See* Ewen, *supra* note 67, at 83. Again, it is possible that the expansion of commercialism tends to promote the atrophy of the most non-commercial elements of other cultural institutions.

82. NEIL POSTMAN, AMUSING OURSELVES TO DEATH 126 (1985); *see also* Collins & Skover, *supra* note 71, at 707 (noting twelve billion display ads, 2.5 million radio ads, and 300 thousand television commercials generated each day; average person spends one and a half years watching commercials).

83. *See* JAMES B. TWITCHELL, CARNIVAL CULTURE 246-47 (1992) (arguing that children's television cartoons are essentially a merchandising medium); *see also* Ellen Edwards, *The Children's Half-Hour: Hostage to Toy Makers?*, WASH. POST, June 10, 1994, at A1 (arguing that children's television is driven by toymakers).

84. *See* Whitney v. California, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring) (explaining that prohibition of free speech is dangerous and not to be done to avert merely trivial harms), *overruled by* Brandenburg v. Ohio, 395 U.S. 444, 449 (1969).

85. *See* LASCH, THE CULTURE, *supra* note 80, at 140 (arguing that teachers pacify students by making school painless); LEBERGOTT, *supra* note 7, at 34-35 (arguing that Western culture maximizes pleasure as a goal); Ronald K.L. Collins & David M. Skover, *The Psychology of First Amendment Scholarship: A Reply*, 71 TEX. L. REV. 819, 826 (1993) (quoting LEO BOGART, STRATEGY IN ADVERTISING 107 (2d ed. 1984)).

elements of many cultural institutions speak, albeit softly or occasionally, against consumption-oriented styles of life. But even their message is often mixed, and it is overridden by the broader culture.

This is a thesis about the sustained balance of cultural institutional forces. It is admittedly not a theme the correctness of which can be demonstrated to a skeptic by a few footnote citations. Its plausibility is presumably widely accepted, and the evidence for its truth can be derived from a mere glance at our contemporary culture. As well, some portion of those who have urged special constitutional protection for commercial speech may find my basic thesis regarding the current institutional role of commercial speech to be plausible and important, and thus reconsider their position on the constitutional protection of commercial speech.

This is not to argue the more controversial thesis that genuine freedom of speech requires at least rough equality of resources among contending forces.[86] All one need argue is that in our current cultural circumstances, no single cultural institution, or set of such institutions, is either inclined to, or able to, provide any substantial "countervailing" speech to counteract the broad, reinforced influence, whether intended or not, of commercial advertising speech. A "bias," in the sense of a distinct vector of cultural institutional forces, toward some form of consumption of commercial goods and services clearly characterizes our day.[87] Of course, some persons are less affected by this cultural bias than others, but this would be true of even the most dominant cultural tendencies in any society.

One should take a moment, though, to consider a possible objection to, or extension of, the argument thus far. Perhaps one might argue that a parallel analysis could be developed for the political sphere, and particularly for the case of competing electoral candidates of opposing ideologies or parties. Isn't it possible that even genuinely competing electoral candidates may, like some waves in the natural world, or like most commercial advertising speech, tend perhaps unintentionally to reinforce certain basic political themes, so as to "bias" public thinking, perhaps in ways not responded to by other institutions?

Certainly, a debate continually dominated by one range of ideologies or parties may tend, even unintentionally, to delegitimize other political options. But for our present purposes, there are important disanalogies between contemporary electoral competition and the competition of the marketplace. A casual examination suggests that most commercial ads in

86. *See* Buckley v. Valeo, 424 U.S. 1, 48-49 (1976) (holding that to restrict speech of some elements in society to enhance voice of others is "wholly foreign to the First Amendment"); *see also* Paul Brest, *Further Beyond the Republican Revival: Toward Radical Republicanism*, 97 YALE L.J. 1623, 1627 (1988); J. Skelly Wright, *Politics and the Constitution: Is Money Speech?*, 85 YALE L.J. 1001 (1976) (analyzing *Buckley* and Federal Election Campaign Act Amendments of 1974).

87. The language of "bias" in this respect is drawn from FRED HIRSCH, SOCIAL LIMITS TO GROWTH 84 (1976). Note that this bias need not be toward "conspicuous" consumption or toward the most expensive affordable goods and services. There may well be fluctuations in a society's consumption patterns over time from the exotic to the plain and functional. *See also* Holly Brubach, *Sackcloth and Ashes*, 67 NEW YORKER, Feb. 3, 1992, at 78.

all media do not refer invidiously to any competitor's good or service, let alone to all significant competitors. Where ads are comparative, they are often done with a light touch. There is rarely any implication that any competitor's product or service is, in any absolute sense, shoddy, dangerous, or not of value in the abstract.

In the electoral realm, however, invidious comparisons are common,[88] even increasingly so.[89] While the effects of "negative" or "attack" ads in political campaigns may be complex and difficult to track,[90] such ads often repel the public, contributing to voter disenchantment, apathy, cynicism, and low voter turnout.[91] Certainly, there is a greater sense conveyed that one's electoral opponent is affirmatively unworthy or harmful than is typically conveyed in the commercial context. Thus electoral speech by competing candidates may not be as typically mutually reinforcing as in the case of commercial ads.

Of course, this is not to deny that the constant exposition of mainstream political ideologies tends to make alternative perspectives seem implausible or inconceivable; this is a central problem in democratic theory. This problem, however, need not be resolved here. Instead, the reader is invited to accept either of two views. First, that mainstream political speech tends, at least by omission, to unreasonably stigmatize some or all non-mainstream political speech and that in the name of equality or freedom, the government should do something about this. Or second, that there is a significant difference between the ways in which commercial speech marginalizes non-commercial attitudes and the ways in which mainstream political speech marginalizes non-mainstream political views, such that it is reasonable to regulate commercial speech in ways unacceptable in the realm of political speech. On either of these views, the argument for the reasonable regulation of commercial speech can go forward.

Nor is it correct to think of all commercial ads as propositional. As any investigation suggests, advertising has undergone qualitative change over time.[92] Many, though hardly all, contemporary ads are largely imagistic or atmospheric, or seek to link a product with a mood, a celebrity, or a somehow assumedly appealing non-celebrity, without proposition

---

88. *See, e.g.*, Neal J. Roese & Gerald N. Sande, *Backlash Effects in Attack Politics*, 23 J. APPLIED SOC. PSYCHOL. 632, 651 (1993) (arguing that politicians frequently resort to a discussion of emotional issues).

89. *See* Sharyne Merritt, *Negative Political Advertising: Some Empirical Findings*, 13 J. ADVERTISING 27-28 (1984); Ruth Shalit, *The Oppo Boom: Smearing for Profit Takes Off*, NEW REPUBLIC, Jan. 3, 1994, at 16; *see also Campaigns in the Muck*, USA TODAY, Feb. 18, 1994, at A1 (increased negative campaigning anticipated by political consultants for 1994).

90. *See* Gina M. Garramone, *Voter Responses to Negative Political Ads*, 61 JOURNALISM Q. 250 (1984) (arguing that backlash is most common effect of negative advertising); Merritt, *supra* note 89, at 37 (admitting that study was limited due to focus on only one election); Roese & Sande, *supra* note 88.

91. *See* Merritt, *supra* note 89, at 37; Roese & Sande, *supra* note 88, at 651. For broader discussion, see generally KAREN S. JOHNSON-CARTEE & GARY A. COPELAND, NEGATIVE POLITICAL ADVERTISING: COMING OF AGE (1991).

92. *See* RAYMOND WILLIAMS, TELEVISION: TECHNOLOGY AND CULTURAL FORM 68 (1974) (arguing that television, as it evolves, is allowing people to experience educational processes rather than just being taught about the processes).

or promise.[93] Such ads are neither true nor false, and neither descriptive nor misleading in the traditional sense.[94] While some modern ads may perhaps be thought of as "semihypnotic and irrational"[95] by design, and others as traditionally propositional, many contemporary ads are non-propositional.

To the extent that freedom of speech is a matter of a search for some propositional truth, protecting ads that do not implicate truth or falsity is in that respect misguided. Such ads are neither fraudulent, deceptive, misleading, nor the opposite, thus cohering poorly or not at all with some of the categories and criteria by which the Supreme Court examines such speech.[96]

Of course, not all commercial speech disdains truth and falsity. More broadly, regulation of contemporary commercial speech raises interesting issues of genuine freedom of speech and of freedom in general. The point of reasonable restrictions on commercial speech is not, however, to flatly prohibit underlying transactions between commercial sellers and potential buyers.[97] In our cultural context, the effects of reasonable regulation of commercial speech on freedom are more subtle.

It is important, for example, not to confuse "diversity of product choice with diversity of life-style or thoughts."[98] Having a choice among brands or products would be a poor sort of freedom if one's culture left one unable to take less consumption-oriented alternatives seriously. More particularly, to the extent that we participate in consumption fads on the basis of popularity, or "bandwagon" effects,[99] we are neither choosing autonomously, nor achieving any genuinely worthy community, especially between generations.[100]

More insidiously, over the long term, in a cultural context in which consumption is a dominant and only ineffectively challenged theme, largely unregulated speech on behalf of consumer goods and services may

---

93. *See* POSTMAN, *supra* note 82, at 127 (commercial ads as increasingly non-propositional).

94. *See* JULES HENRY, CULTURE AGAINST MAN 47 (1963); MARK POSTER, THE MODE OF INFORMATION 59 (1990) (discussing the work of Jean Baudrillard); *see also supra* note 70 and accompanying text; Todd F. Simon, *Defining Commercial Speech: A Focus on Process Rather Than Content*, 20 NEW ENG. L. REV. 215, 239 (1985) (arguing that non-objective advertising is "inherently unverifiable").

95. FROMM, *supra* note 63, at 123; Bailey Kuklin, *Self-Paternalism in the Marketplace*, 60 U. CIN. L. REV. 649, 649 (1992).

96. *See, e.g.*, Posadas de Puerto Rico Assocs. v. Tourism Co., 478 U.S. 328, 340 (1986) (citing Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557 (1980)).

97. *See* KATONA, *supra* note 33, at 376 (questioning Fromm's assertion that the consumer is passive and questioning the need for legal restrictions on advertising); NOVAK, *supra* note 72, at 108; SCITOVSKY, *supra* note 47, at 126 (arguing that although consumer stimulation is expensive, the proper remedy does not include denying the stimulus to the consumer).

98. MANDER, *supra* note 69, at 125.

99. *See* HARVEY LEIBENSTEIN, BEYOND ECONOMIC MAN: A NEW FOUNDATION FOR MICROECONOMICS 48-49 (1976) (defining "bandwagon effect" as "the desire to join the crowd").

100. *See generally* Robert D. Putnam, *The Prosperous Community: Social Capital and Public Life*, 13 AM. PROSPECT 35 (1993). *See also* Jhally, *supra* note 50, at 813 (commercial culture as discouraging bonds with future generations).

tend to change our basic tastes, capacities, and judgments[101] in ways we do not anticipate or even recognize. Some of these changes, such as increasingly demanding standards for product warranties, may well be benign. Others, however, such as loss of capacity to value, to enjoy, or even to envision non-market solutions to problems, are not.[102] To claim that a "free market" in non-deceptive commercial speech makes us genuinely freer thus raises, at a minimum, serious problems of measurement and commensurability. Why must a democratic society interpret its constitution to bias the choice in the commercial direction if that society wishes to maximize the most valuable sorts of freedom?

To illustrate this general problem, let us consider, almost at random, one particular scenario. With regard to a typical consumer good, the interests of producers and consumers may both be served, via expansion of market and economies of scale, if consumer tastes regarding a product can be modified in the direction of some homogeneous fairly low common denominator.[103] There need be nothing narrowly or classically deceptive or coercive about this taste modification process, in the sense that no literal force or fraud is applied. Jazz or classical music, say, is still in some sense available, but is widely found to be genuinely unpalatable. Admittedly, in utilitarian terms, consumers may well gain from even rather intensive development of this process, at least in some sense. But the price in freedom may be high.

True, one must grant that freedom in general and freedom of speech are hardly equivalent. But for our immediate purposes, the most relevant commonly cited purpose or value underlying freedom of speech is that of "self-realization," or the development and flourishing of the personality.[104] The dubious effect of largely unregulated commercial speech on consumers' genuine self-realization means that such speech may actually undermine the purposes of free speech as much as it may undermine freedom more generally.

Thus, even setting aside all cases of false or otherwise deceptive commercial speech, it is far from clear that freedom in general, or at least freedom of speech, is unequivocally maximized by special constitutional protection for commercial speech. What tips the balance, again, is a realistic assessment of contemporary institutional forces and the absence of any real institutional capacity for, or interest in, challenging the dominance of the culture of commercial speech.

---

101. *See* SCITOVSKY, *supra* note 73, at 5 (arguing that the theory that consumers know what is best for them is unscientific and inaccurate).

102. *See id.*; *see also* Allan C. Hutchinson, *More Talk: Against Constitutionalizing (Commercial) Speech*, 17 CAN. BUS. L.J. 2, 15 (1990) ("a commercially saturated atmosphere" as tending to "trivialize and impoverish democratic politics").

103. *See* SCITOVSKY, *supra* note 73, at 9 (arguing that sellers cater to the "desires everybody shares").

104. *See, e.g.*, C. Edwin Baker, *Realizing Self-Realization: Corporate Political Expenditures and Redish's The Value of Free Speech*, 130 U. PA. L. REV. 646 (1982); Robert J. Sharpe, *A Comment on Allan Hutchinson's Money Talk: Against Constitutionalizing Commercial Speech*, 17 CAN. BUS. L.J. 35, 39 (1990).

As matters stand, any reasonable governmental restriction on commercial speech thus serves, if only minimally, to reduce the degree of cultural institutional bias[105] in favor of a commercial speech-guided culture of consumption. Whether such restrictions on commercial speech tend, in the aggregate, to promote genuine happiness, freedom, or human dignity by reducing that bias should be left to reasonable democratic decisionmaking. It is then up to the society to take other necessary steps toward legitimizing or encouraging less commercial ways of life, consistent with the Constitution and other democratic policy preferences.

III. COMMERCIAL SPEECH AND THE VALUES UNDERLYING FREEDOM OF SPEECH

The question of whether, or how stringently, to protect constitutionally commercial speech has proved remarkably resistant to consensus. Perhaps symbolically, the leading modern free speech theorist, John Stuart Mill, was obviously ambivalent as to the level of protection properly to be accorded some forms of commercial speech.[106] Contemporary writers are, if not individually ambivalent, at least mutually divided.[107]

One broad camp favors some degree of special First Amendment protection for commercial speech.[108] Writers in this camp recognize that many commercial ads they would protect are less propositional than imagistic or symbolic, but they note that such ads may promote distinctive lifestyles, such as materialism or hedonistic consumption, as might fully protected non-commercial speech.[109] Failure to protect commercial speech may, according to some, even be dangerous. Such failure to protect commercial speech fully may, it is said, provide government with "a powerful weapon to suppress or control speech by classifying it as merely commercial."[110]

In contrast, there is a broad group of writers more skeptical of the need for special protection for commercial speech.[111] In this camp, it is sometimes argued that special protection for commercial speech may en-

---

105. For broader background, see Owen M. Fiss, *Free Speech and Social Structure*, 71 IOWA L. REV. 1405, 1425 (1986) ("Contemporary social structure will, if left to itself, skew public debate.").

106. *See* JOHN STUART MILL, ON LIBERTY 168-69 (Penguin ed. 1974) (1859).

107. *See, e.g.*, Burt Neuborne, *The First Amendment and Government Regulation of Capital Markets*, 55 BROOK. L. REV. 5, 6-9 (1989) (providing a broad typology of approaches to the constitutional status of commercial speech).

108. *See, e.g.*, *Leading Cases of the 1992 Supreme Court Term*, 107 HARV. L. REV. 144, 225 (1993) (stating that the Supreme Court "continues to leave 'commercial speech' with insufficient protection").

109. *See, e.g.*, Rodney A. Smolla, *Information, Imagery, and the First Amendment: A Case for Expansive Protection of Commercial Speech*, 71 TEX. L. REV. 777, 791 (1993).

110. Alex Kozinski & Stuart Banner, *Who's Afraid of Commercial Speech?*, 76 VA. L. REV. 627, 653 (1990).

111. *See, e.g.*, Joshua Cohen, *Freedom of Expression*, 22 PHIL. & PUB. AFF. 207, 237 (1993) (commercial speech as allegedly less closely connected with "expressive" interests than is political speech); Alvin I. Goldman, *Epistemic Paternalism: Communication Control in Law and Society*, 88 J. PHIL. 113, 128 (1991) (asserting that "an unregulated marketplace of ideas seems essential for political speech but not for commercial speech").

danger, rather than strengthen, protection for political speech, on the grounds that such a policy might, over time, erode any public conviction of the core value of freedom of speech.[112] Alternatively, it has been argued that commercial speech generates fewer "external" benefits uncaptured by the speaker than does political speech.[113]

Typically this battle is played out, quite inconclusively, over the relationship between commercial speech and one or more of the values or purposes thought to underlie special protection for free speech. Perhaps the basic problem in this regard is that "[i]n a pluralistic, heterogeneous society, such as ours, people disagree sharply over the needs, interests, and desires that give rise to the evaluative points of the First Amendment taxonomy of worth."[114]

Thus it has been argued that commercial speech is either unrelated,[115] or at best differently linked,[116] to the values thought to underlie the Free Speech Clause. Not surprisingly, the contrary also has been argued, on various grounds. Commercial speech, it is argued, typically expresses "ideas and values."[117] Professor Burt Neuborne has observed that "abandoning protected commercial speech altogether denies consumers access to valuable information."[118] Michael Perry has argued for protection of commercial speech under what he calls the "democratic" and "epistemic" conceptions of free speech.[119] Writers such as Martin Redish[120] and Daniel Farber[121] have noted the artistic value of at least some commercial advertisements.

This sort of persistent, apparently unresolvable controversy is paralleled by the debate over the relationship between commercial speech and the more particular value of self-realization. Some writers have argued

---

112. *See* Vincent Blasi, *The Pathological Perspective and the First Amendment*, 85 COLUM. L. REV. 449, 486 (1985).

113. *See* Richard A. Posner, *Free Speech in an Economic Perspective*, 20 SUFFOLK U. L. REV. 1, 39-40 (1986). Of course, to the extent that commercial speech, or advertising in particular, helps to generate a culture of commodity consumption, such speech may create external effects, both positive and, as we have suggested, negative.

114. Heidi L. Feldman, *Objectivity in Legal Judgment*, 92 MICH. L. REV. 1187, 1241 (1994).

115. *See, e.g.*, Lillian R. BeVier, *The First Amendment and Political Speech: An Inquiry into the Substance and Limits of Principle*, 30 STAN. L. REV. 299, 353 (1978) (noting the public's interest in commercial speech messages as "totally irrelevant to first amendment values").

116. *See* Thomas I. Emerson, *First Amendment Doctrine and the Burger Court*, 68 CAL. L. REV. 422, 460 (1980) ("Commercial speech does not promote the underlying values of the system in the same manner as does other expression.").

117. Ronald D. Rotunda, *The Commercial Speech Doctrine in the Supreme Court*, 1976 U. ILL. L. F. 1080, 1091.

118. Burt Neuborne, *A Rationale For Protecting and Regulating Commercial Speech*, 46 BROOK. L. REV. 437, 440 (1980); *see also* Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557, 563 (1980) ("The First Amendment's concern for commercial speech is based on the informational function of advertising."); First Nat'l Bank v. Bellotti, 435 U.S. 765, 783 (1978).

119. Michael J. Perry, *Freedom of Expression: An Essay on Theory and Doctrine*, 78 Nw. U. L. REV. 1137, 1171-72 (1984).

120. Redish, *supra* note 14, at 431.

121. Daniel A. Farber, *Commercial Speech and First Amendment Theory*, 74 Nw. U. L. REV. 372, 384 (1979).

that commercial speech typically promotes self-realization,[122] while others are more skeptical on this score.[123] This debate remains unresolvable largely because of ambiguity surrounding ideas such as self-expression and autonomy.[124] Roughly, self-realization in the sense of acting as one happens to wish with respect to receiving commercial messages is obviously promoted by freedom of commercial speech. In contrast, if we think of self-realization in terms of human dignity or the highest development of the human personality, the connection between commercial speech and such self-realization will seem much more dubious. Whether unregulated non-deceptive commercial speech promotes self-realization in the sense of happiness or subjective well-being has been controversial, but as we have seen above,[125] available evidence suggests that it does not.

In this context, little is gained by insisting upon a controversially narrow view of free speech values. Perhaps the leading exponent of such an approach is C. Edwin Baker. Professor Baker argues that "interpretations of the free speech clause should focus on the liberty or freedom of the speaker"[126] as opposed to that of any actual or potential audience. As Professor Baker then develops the argument, corporations—or more particularly, non-household-oriented, non-media corporations other than labor unions—fall outside the scope of the Free Speech Clause.[127] This is because such corporations are, in a competitive market, structurally bound to pursue profit in some relevantly constrained, determinate way that sufficiently explains their speech, and cannot otherwise pursue truth, anyone's self-realization, or freedom.[128]

Each of the necessary steps in Professor Baker's subtle and complex argument will be plausible to some. But the combined controversiality of those steps leads one to suspect that the argument is acceptable only to a subset of those already disposed to deny special constitutional protection to commercial speech. Given the overall stringency of Professor Baker's premises, his theory is of only limited use in persuading those who are agnostic on the question of protecting commercial speech.

---

122. *See, e.g.*, David F. McGowan, *A Critical Analysis of Commercial Speech*, 78 CAL. L. REV. 359, 361 (1990); Kenton F. Machina, *Freedom of Expression in Commerce*, 3 L. & PHIL. 375 (1984) (arguing for full protection for commercial speech on "autonomy" grounds).

123. *See, e.g.*, Michael Davis, *The Special Resiliency of Commercial Speech as Deus Ex Machina*, 6 L. & PHIL. 121 (1987) (replying to Machina, *supra* note 122); Thomas H. Jackson & John C. Jeffries, Jr., *Commercial Speech: Economic Due Process and the First Amendment*, 65 VA. L. REV. 1, 14 (1979) (asserting "commercial speech has no apparent connection with the idea of individual self-fulfillment").

124. *See, e.g.*, NICHOLAS WOLFSON, CORPORATE FIRST AMENDMENT RIGHTS AND THE SEC 63 (1990) (linking the ideas of self-expression, human dignity, and autonomy).

125. *See supra* text accompanying notes 1-61.

126. C. Edwin Baker, *Commercial Speech: A Problem in the Theory of Freedom*, 62 IOWA L. REV. 1, 4 (1976).

127. *See id.* at 13, 28, 40.

128. *See id.* For further relevant discussion by Professor Baker, see C. Edwin Baker, *Advertising and a Democratic Press*, 140 U. PA. L. REV. 2097, 2232-33 (1992) (expanded in C. EDWIN BAKER, ADVERTISING AND A DEMOCRATIC PRESS (1994)). Of course, structural constraints may be a sufficient explanation for speech without being necessary to explain the speech, which may have other sufficient motivations.

More generally, theorists' understandable focus on whether, or to what extent, commercial speech promotes the values thought to underlie the Free Speech Clause has been unproductive because of the equivocality and controversial nature of the concepts involved. It is largely because of this methodological dead end that one takes a different tack. First, the widely presumed linkage between typical commercial speech and freedom or happiness is in doubt. Second, the practical irrelevance of "counter-speech" responses to commercial speech in a cultural context in which no institution is either able or inclined to provide any such challenge on a meaningful scale is noted.

IV. COMMERCIAL AND NON-COMMERCIAL SPEECH

To this point, we have set aside the question of a precise, maximally useful definition of commercial speech, relying instead on uncontroversial examples. A precise and widely useful definition, allowing courts or other persons to classify easily and uncontroversially instances of speech into commercial or noncommercial may well not be possible. But one may at least provide some reason for believing that ambiguities in the idea of commercial speech will not commonly generate implausible or seriously harmful results.

In this respect, the Supreme Court has been of limited assistance.[129] The leading commercial speech case, *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council*,[130] referred to "speech which does 'no more than propose a commercial transaction.' "[131] On the other hand, the Court has also, in the influential case of *Central Hudson Gas & Electric Corp. v. Public Service Commission*,[132] referred to "expression related solely to the economic interests of the speaker and its audience."[133]

Apart from their inconsistency, these two definitions of commercial speech are useful only in particular contexts. Consider first the "no more than propose a commercial transaction" approach. Intuitively, commercial speech may involve more than merely a proposal without losing its commercial character. This definition may serve in contexts of commercial advertising, securities offerings, and other areas, as long as nothing hangs on the fact that this definition itself does not specify what is meant by 'commercial.' But other intuitively commercial forms of speech, such as many proxy statements, corporate financial statements, reports to shareholders, commercial contracts, product safety brochures, warranties, product labels, and so on do not fit neatly within the category of "proposals."[134]

---

129. *See* David F. McGowan, *supra* note 122, at 400-02 (discussing and critiquing judicial inconsistency in defining the scope of commercial speech).

130. 425 U.S. 748 (1976).

131. *Id.* at 762 (quoting Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 385 (1973)); *see also* Board of Trustees v. Fox, 492 U.S. 469, 482 (1989).

132. 447 U.S. 557 (1980).

133. *Id.* at 561.

134. *See* Steven Shiffrin, *The First Amendment and Economic Regulation: Away From a General Theory of the First Amendment*, 78 NW. U. L. REV. 1212, 1214 (1983).

On the other hand, the "solely economic interests" approach to defining commercial speech may include some instances of commercial speech beyond mere proposals, but it also leaves out much that is ordinarily considered commercial speech.[135] Consider, for example, the central example of a typical commercial advertisement for a perfume, an elaborate exercise machine, a hair replacement technique, cigarettes, or athletic shoes. Now, these and similar ads may relate in part to the economic interests of producers and consumers. Far more is at stake, however, than "solely"[136] economic interests in the purchase of these and many other products and services.

These definitions are thus in certain respects less than ideal. But no simple, invariably helpful alternative definition seems attainable. It has been declared that "the doctrine of commercial speech rests on a clean distinction between the market for ideas and the market for goods and services."[137] This approach is useful in certain respects. It is helpful, for example, in reminding us of the important truth that often speech about markets, generally or in particular, and especially speech about whether or how such markets should be regulated, transcends commercial speech.

Plainly, though, markets for ideas, however we reasonably define "ideas," and markets for goods and services are not mutually exclusive. It is not difficult to imagine that authors and publishers of all sorts of books propounding ideas may be in part motivated by a quest for profit, as, in turn, may many of the book's purchasers.[138] Inquiring into the degree, or the causal necessity, of profit motivation on a case-by-case basis is obviously problematic.

No simple distinction between the markets for ideas and for goods and services is thus possible. But this does not mean that the distinction is useless for all purposes, or that no serviceable distinction between commercial and noncommercial speech is possible. Commercial speech can be a viable category even if it has no "essence." Professor Christopher Stone has observed that "[w]e may be able to find nothing better than 'family resemblances' among members of the commercial speech set."[139]

One is at least able to distinguish commercial speech well enough to justify treating commercial speech differently from political speech. The most important reasons for this differential treatment have been set forth at some length above.[140] But a bit more may be said to enhance interest in the distinction. Professor Ronald Coase wonders why governments should be deemed generally more competent to regulate commercial than

---

135. *See id.* at 1222.

136. *See supra* text accompanying note 133.

137. Jackson & Jeffries, *supra* note 123, at 2; *see also* Zauderer v. Office of Disciplinary Counsel, 471 U.S. 626, 637 (1985) (declaring that "advertising pure and simple" is clearly within the bounds of commercial speech).

138. *See* FREDERICK SCHAUER, FREE SPEECH: A PHILOSOPHICAL ENQUIRY 159-60 (1982) (declining to include all profit-motivated speech within the category of commercial speech).

139. Christopher D. Stone, *Theorizing Commercial Speech*, 11 GEO. MASON U. L. REV. 95, 113 (1988).

140. *See supra* sections I-III.

political speech.[141] If Professor Coase is right, much of the fuss over distinguishing commercial and political speech would seem misplaced.

But this skepticism is itself misplaced. No doubt some government regulation of commercial speech is really driven by an enterprise's desire to impose restrictions on entry by potential competitors, by a wish to unjustly discriminate, or by a range of other anticompetitive motivations.[142] But this is hardly the whole story, and it is implausible to imagine that a particular government can be equally trusted to regulate fairly its own rival political parties, movements, and ideologies along with commercial enterprises and commercial speech.

The point may be put this way: it is at least plausible in some cases that we are better off protecting unpopular political ideas, perhaps for the sake of promoting an ethos of tolerance.[143] On the other hand, few[144] would argue that we are better off, even in the long run, legally protecting the advertisement of latently dangerous products and relying on the tort system, *Consumer Reports*, and the possibility of criminalizing the production of the goods themselves to minimize the carnage. Banning an obvious carcinogen and banning an allegedly harmful political party require entirely different justifications.

This is not to minimize the potential for abuse of the government's power to regulate commercial speech. But such abuse can be reduced by means other than specially protecting commercial speech. Most, if not all, denunciations of an enterprise's attempt to reduce competition by inducing the government to restrict commercial speech are themselves fully protected political speech. Any government actions reducing competition by restricting commercial speech should be subject to a judicial test of reasonableness under the Free Speech, Equal Protection, and Due Process Clauses.[145] Any restrictions of commercial speech that are, for example, racially invidious[146] should be subjected to the most stringent judicial scrutiny under the Equal Protection Clause.[147]

There is thus no reason to suppose that permitting reasonable government regulation of even non-deceptive commercial speech could not, at least over the long term, provide significant net social benefits. But we must do what we reasonably can to prevent political or other socially valuable kinds of speech from being inadvertently swept into the net of commercial speech regulation.

---

141. *See* R.H. Coase, *Advertising and Free Speech*, 6 J. LEGAL STUD. 1, 2 (1977); *see also* Alex Kozinski & Stuart Banner, *The Anti-History and Pre-History of Commercial Speech*, 71 TEX. L. REV. 747, 751-52 (1993); Fred S. McChesney, *Commercial Speech in the Professions: The Supreme Court's Unanswered Questions and Questionable Answers*, 134 U. PA. L. REV. 45 (1985).

142. *See, e.g.*, Ronald A. Cass, *Commercial Speech, Constitutionalism, Collective Choice*, 56 U. CIN. L. REV. 1317, 1361-62 (1988); David A. Strauss, *Constitutional Protection for Commercial Speech: Some Lessons from the American Experience*, 17 CAN. BUS. L. J. 45, 45 (1990) (emphasizing speech regulation as a possible cover for discrimination).

143. *See generally* LEE BOLLINGER, THE TOLERANT SOCIETY (1986).

144. *But see* HERBERT SPENCER, THE MAN VERSUS THE STATE (Eric Mack ed. 1982) (1884).

145. *Cf.* Williamson v. Lee Optical, 348 U.S. 483, 487-88 (1955).

146. *See* Strauss, *supra* note 142, at 45.

147. *See, e.g.*, Loving v. Virginia, 388 U.S. 1 (1967).

It is not difficult to think of cases that test the classificatory boundary of political and commercial speech: a cigarette manufacturer's corporate support of the Bill of Rights,[148] or a clothing manufacturer's advertised stance on social issues.[149] Ironically, some ads may even play off the public perception of the fraying of many social relationships, offering consumption-oriented solutions.[150] A desperately poor person who is reduced to begging may engage in commercial or political speech, or both.[151] A magazine such as *Consumer Reports* may seek to guide consumer purchases, but with no financial stake in the recommended goods and services.[152] Presumably scientifically based health claims may be made on behalf of particular products, sometimes in the context of an otherwise purely commercial ad.[153]

However, it is important not to overestimate the scope and severity of these sorts of arguable borderline cases. Most typical advertisements will be classified as commercial speech on any reasonable theory. Most speakers, on the other hand, who wish primarily to convey a political message can readily avoid entanglement with the problem of commercial speech on any reasonable definition. Thus the costs of misclassifying political or commercial speech are likely limited.

In particular kinds of cases, special considerations may play a role in classifying an instance of speech as commercial or non-commercial. Consider, for example, a claim that egg yolks are nutritionally beneficial, despite or because of their cholesterol content. Some courts may wish to distinguish between proponents of such a claim based not on the presence or absence of a pecuniary stake, but on a broader consideration of what reactions would logically please the particular speaker. A scientist who sees special nutritional benefit in egg yolks presumably wants people to actually eat, rather than merely buy, the eggs. An egg producer who makes the same claim, on the other hand, may or may not care what purchasers of eggs actually do with them, and may be most pleased by purchasers who buy eggs, throw them away, and then buy more.

There is no guarantee that some acceptable shortcut method for deciding particular close cases will always be available. On the other hand, approaches are available to reduce the costs, if not the risk, of misdeciding genuinely close cases. For example, courts faced with a close classification

148. *See* Leo Bogart, *Freedom to Know or Freedom to Say?*, 71 TEX. L. REV. 815, 816 n.5 (1993).

149. Brubach, *supra* note 87, at 78; *see also* LASCH, THE CULTURE, *supra* note 80, at 139 (discussing the phenomenon of advertisements attempting, reasonably or otherwise, to link themselves to the idea of freedom).

150. *See* Ewen, *supra* note 67, at 93.

151. *See* Loper v. New York City Police Dept., 999 F.2d 699 (2d Cir. 1993); Young v. New York City Transit Auth., 903 F.2d 146 (2d Cir.), *cert. denied*, 498 U.S. 984 (1990).

152. *See* T.M. Scanlon, Jr., *Freedom of Expression and Categories of Expression*, 40 U. PITT. L. REV. 519, 541 (1979) (*Consumer Reports* entitled to full free speech protection).

153. *See* Martin H. Redish, *Product Health Claims and the First Amendment: Scientific Expression and the Twilight Zone of Commercial Speech*, 43 VAND. L. REV. 1433 (1990) (arguing for protection of scientific health claims in the context of commercial advertising); *see also* National Comm'n on Egg Nutrition v. FTC, 570 F.2d 157 (7th Cir. 1977).

problem may notice that speakers may have some antecedent control over their own destiny. At least some speakers are easily able, if they choose, to formulate what they wish to say in a way that makes their speech as clearly non-commercial as reasonably possible under the circumstances.

Courts may thus wish to establish a reasonable and socially desirable incentive for speakers who fear their speech may be classified commercial. A judicial rule might take something approaching the following form: in genuinely close, borderline cases only, the court may wish to protect specially speech where the speaker came as close as reasonably possible, under the circumstances, to presenting the speech in question as clearly non-commercial. What is "reasonably possible" must be considered in light of the speaker's own resources and capacities, as well as the speaker's interests in not sending a distorted or insincere message, or in addressing an undesired audience.[154]

Thus in borderline cases of commercial speech, the courts should look to the range of speech alternatives antecedently available to the speaker, and determine whether more clearly non-commercial speech alternatives were essentially ignored. For example, a court might point out to a store that it is not difficult to express one's views as to the Fourth of July without also describing items for sale on that occasion in lavish detail. Perhaps in cases on the border of commercial and non-commercial speech, more can be realistically asked of major corporations[155] than of the destitute.[156]

V. INTERMEDIATE-LEVEL PROTECTION AS EXCESSIVE PROTECTION FOR COMMERCIAL SPEECH

Currently, the Supreme Court majority seems to accept the idea of according commercial speech, in at least some cases,[157] some lesser degree of constitutional protection than is given to political or other fully protected speech. The basic contours of the current constitutional test for regulating commercial speech were set forth in *Central Hudson Gas & Electric Corp. v. Public Service Commission*,[158] which concerned regulation of the promotion of electricity consumption, for the sake of energy conservation. There, the Court established a four-part test that begins by specifying that the speech at issue:

> must concern lawful activity and not be misleading. Next, we must ask whether the asserted governmental interest is substan-

154. For a more elaborate exposition of this theme in a related context, see generally R. George Wright, *Speech on Matters of Public Interest and Concern*, 37 DEPAUL L. REV. 27 (1987).

155. *See, e.g.*, Bolger v. Youngs Drug Products Corp., 463 U.S. 60, 68 (1983) ("Advertisers should not be permitted to immunize false or misleading product information from government regulation simply by including references to public issues.").

156. *See supra* note 151 and accompanying text.

157. *See, e.g.*, Edenfield v. Fane, 113 S. Ct. 1792, 1798 (1993) (referring to "an intermediate standard of review"); Board of Trustees v. Fox, 492 U.S. 469, 480 (1989); *In re* Primus, 436 U.S. 412, 435 (1978) (toleration of more imprecise regulation of commercial than of political speech); Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447 (1978).

158. 447 U.S. 557 (1980).

> tial. If both inquiries yield positive answers, we must determine whether the regulation directly advances the governmental interest asserted, and whether it is not more extensive than is necessary to serve that interest.[159]

In practice, the *Central Hudson* test is problematic because it erects obstacles to reasonable regulation of purely commercial speech at every turn. The test invites constitutional challenges to such regulations on grounds so vague and open-ended as, in many cases, to encourage the expression of mere subjective judicial preference. Let us consider the elements, and complications, of the *Central Hudson* test in turn.

It is important to bear in mind that the general burden of proof in these cases is on the government.[160] However useful this burden placement may be in political speech cases, in the commercial speech context it tends to undermine reasonable government regulation at every stage of the *Central Hudson* test.

Consider first the test's invitation to the government to prove that the speech at issue is "misleading."[161] However familiar the idea of "misleadingness" may seem, in practice its application is often complex and indeterminate. With regard to commercial speech, typically, "[t]he line between truth and falsity . . . is hardly crystal clear."[162] More broadly, cases suggest that "verification is very difficult in both the political speech and the commercial speech areas."[163]

The problem here is in part that typical commercial claims may be, even if in some sense true, deceptive or misleading to some persons to some degree.[164] The court is thus required to decide what degree of deception, of what number of persons, with what sorts of arguable consequences, is required before the speech will be labeled as misleading.[165] A court may inquire, for example, whether a commercial claim is misleading because "it unduly emphasizes trivial or relatively uninformative facts."[166] We are left to wonder what degree of emphasis upon the relatively trivial is due.

Further muddying the water is that in most deceptive advertising cases, the crucial question is not whether a given advertising claim is misleading, but whether the presumably misleading claim can fairly be

---

159. *Id.* at 566.

160. *See* Ibanez v. Florida Dep't of Business & Professional Regulation, 114 S. Ct. 2084, 2089 n.7 (1994); *Edenfield*, 113 S. Ct. at 1800; Bolger v. Youngs Drug Products Corp., 463 U.S. 60, 71 n.20 (1983).

161. *See Central Hudson*, 447 U.S. at 566.

162. Frederick Schauer, *Commercial Speech and the Architecture of the First Amendment*, 56 U. CIN. L. REV. 1181, 1192 (1988).

163. Robert Pitofsky, *First Amendment Protections and Economic Activity*, 11 GEO. MASON U. L. REV. 89, 91 (1988).

164. *See* Robert B. Reich, *Preventing Deception in Commercial Speech*, 54 N.Y.U. L. REV. 775, 783 (1979).

165. *See* Shiffrin, *supra* note 134, at 1219; *see also* Gammon v. GC Servs. Ltd. Partnership, 27 F.3d 1254 (7th Cir. 1994) (discussing the inherent unworkability and indeterminacy of any "least sophisticated consumer" standard in the consumer protection area); Clomon v. Jackson, 988 F.2d 1314 (2d Cir. 1993).

166. Shapero v. Kentucky Bar Ass'n, 486 U.S. 466, 479 (1988) (citations omitted).

ascribed to the claimant.[167] The speaker may thus simply dispute the government's interpretation of the meaning or meanings of the ad for particular audiences.[168]

These sorts of complications often are played out through judicial subdivision of the concept of misleading commercial speech. For example, the Court has been known to consider merely whether the speech at issue is misleading "in the abstract,"[169] without explaining what is meant by such a qualifier. Courts have varied their degree of scrutiny of commercial speech depending upon whether the speech at issue is variously deemed actually,[170] inherently,[171] potentially,[172] necessarily,[173] demonstrably,[174] or possibly[175] misleading.[176]

None of these characterizations, however, bypasses the need for loosely constrained judgments about the extensiveness, degree, and consequences of misleading speech. In some cases, the misleadingness of the speech is reducible by means such as a disclaimer. But courts must then judge, somehow, whether the phrasing, prominence, and clarity of the disclaimer are sufficient, or whether the disclaimer, if sufficiently conspicuous, tends unduly to create confusion.[177]

Once any issues of misleadingness are somehow resolved, the court must consider whether the government can identify a substantial interest

---

167. *See* Richard Craswell, *Interpreting Deceptive Advertising*, 65 B.U. L. REV. 657, 659 (1985); *see also* ITT Continental Baking Co. v. FTC, 532 F.2d 207, 213 (2d Cir. 1976) ("For the most part the petitioners admitted that Wonder Bread did not have the nutritional qualities alleged to have been claimed for it in the challenged advertisements, but they denied that the advertisements represented that the bread did have those qualities.").

168. *See ITT Continental Baking Co.*, 532 F.2d at 214 ("children below the age of seven would generally tend to accept the 'fantasy growth sequence' in Wonder Bread advertising as literal truth"); *see also id.* at 218 ("it is fair to suppose that the commercials in question were designed to avoid making any specific nutritional misrepresentations while at the same time conveying the general idea that Wonder Bread would somehow significantly contribute to growth in children"). Relying on free speech cases from other contexts, though, an advertiser might wish to argue that its free speech rights must not be held hostage to speech standards appropriate for children. *See, e.g.*, Butler v. Michigan, 352 U.S. 380, 383 (1957); *see also* Sable Communications v. FCC, 492 U.S. 115, 126-27 (1989).

169. *See* Posadas de Puerto Rico Assocs. v. Tourism Co., 478 U.S. 328, 341 (1986).

170. Peel v. Attorney Registration and Disciplinary Comm'n, 496 U.S. 91, 111 (1990).

171. *Id.*

172. *Id.*; *see also* Shapero v. Kentucky Bar Ass'n, 486 U.S. 466, 472 (1988).

173. *See In re* R.M.J., 455 U.S. 191, 202 (1982).

174. *Id.*

175. Friedman v. Rogers, 440 U.S. 1, 13 (1979); SEC v. Wall Street Publishing Inst., Inc., 851 F.2d 365, 374 n.9 (D.C. Cir.), *cert. denied*, 489 U.S. 1066 (1988).

176. In addition, different degrees of support are required by the Federal Trade Commission depending upon whether the commercial claim includes an assertion of support by scientific tests or not. *See* Removatron Int'l Corp. v. FTC, 884 F.2d 1489, 1492 n.3 (1st Cir. 1989). Importantly, the FTC then goes on to subdivide the former category into "specific" or "non-specific" claims, usually requiring "two well-controlled scientific studies" in non-specific claim cases. *Id.*

177. *See id.* at 1497. For a case illustrating the indeterminacy of these sorts of inquiries, see Kraft, Inc. v. FTC, 970 F.2d 311, 315-16 (7th Cir. 1992), *cert. denied*, 113 S. Ct. 1254 (1993) (discussing the proper characterization of ads relating Kraft Singles to milk and calcium content).

underlying the regulation.[178] This is not a trivial inquiry. The government's burden in this respect "is not satisfied by mere speculation or conjecture; rather, a government body . . . must demonstrate that the harms it recites are real."[179] Depending upon their degree of sympathy for the state regulation, courts may identify the interest, or combination of interests, at stake in more and less favorable ways.[180] The idea of "demonstrating" the reality of a harm is daunting; it is not clear that even a criminal case requires the prosecutor to "demonstrate" literally the existence of the charged harm. Whatever the term "demonstrate" is taken to mean, it is available for use in derailing any reasonable regulation of commercial speech.

Taken together, the Supreme Court's decisions "leave little insight as to what criteria the Court used"[181] in specifying the relevant state interests, or in determining their substantiality. The combination of literally rigorous language and the lack of judicial guidance invites litigation of any governmental regulation of commercial speech.[182]

The problem is made worse at the next stage of the inquiry, during which the government must "demonstrate"[183] that the regulation at issue will "in fact" advance the specified substantial interest "in a direct and material way."[184] In this context, directness has been associated with "an immediate connection between the prohibition and the government's asserted end."[185] On the other hand, directness has been contrasted variously with connections that are tenuous,[186] highly speculative,[187] ineffective,[188] only remotely supportive of the government interest,[189]

178. Central Hudson Gas & Electric Corp. v. Public Serv. Comm'n, 447 U.S. 557, 564 (1980).

179. Edenfield v. Fane, 113 S. Ct. 1792, 1800 (1993); *see also* Ibanez v. Florida Dep't of Business & Professional Regulation, 114 S. Ct. 2084, 2089 (1994).

180. For an example of a generous characterization, see Posadas de Puerto Rico Assocs. v. Tourism Co., 478 U.S. 328, 342 (1986) ("the legislature's interest . . . is not necessarily to reduce demand for all games of chance, but to reduce [local] demand for casino gambling"). The Court could, of course, have insisted on something like a "substantial" reduction in casino gambling.

181. Kansas v. United States, 16 F.3d 436, 443 (D.C. Cir. 1994).

182. The Seventh Circuit has noted that "[c]oncrete evidence of past effects is likely to be difficult and expensive, if not impossible, to obtain in most deceptive advertising cases." National Comm'n on Egg Nutrition v. FTC, 570 F.2d 157, 165 (7th Cir. 1977).

183. *Ibanez*, 114 S. Ct. at 2089; *Edenfield*, 113 S. Ct. at 1800; *see also* Philip B. Kurland, Posadas de Puerto Rico v. Tourism Company: " *'Twas Strange; 'Twas Passing Strange; 'Twas Pitiful, 'Twas Wondrous Pitiful'*", 1986 SUP. CT. REV. 1, 7 (1987) (endorsing a requirement of demonstrating, as opposed to merely asserting, the efficacy of promoting the government interest).

184. *Ibanez*, 114 S. Ct. at 2089; *Edenfield*, 113 S. Ct. at 1800.

185. Adolph Coors Co. v. Bentsen, 2 F.3d 355, 357 (10th Cir. 1993), *cert. granted*, 114 S. Ct. 2671 (1994).

186. *Id.*

187. *Id.*

188. Central Hudson Gas & Electric Corp. v. Public Serv. Comm'n, 447 U.S. 557, 564 (1980); Cal-Almond Inc. v. United States Dep't of Agriculture, 14 F.3d 429, 437 (9th Cir. 1993).

189. *Central Hudson*, 447 U.S. at 564.

conditional,[190] or involving only limited incremental,[191] speculative,[192] or marginal[193] support for the government interest.[194]

The potential for these concepts, which are largely just opposing end points on progressive continua, to serve as merely conclusory labels for opposing outcomes is obvious. More interestingly, the focus on "directness" or "immediacy" itself seems misconceived. If the idea is taken literally, why should we care whether the regulation advances the government interest directly or indirectly, as long as the advancement itself is substantial? The Court's choice of the "directness" terminology is surprising, in that the Court has, in a famous line of Commerce Clause cases, learned to focus on the substantiality of relationships, as opposed to their directness or indirectness.[195]

Finally, under *Central Hudson* the government must show that the regulation "is not more extensive than is necessary"[196] to promote the government interest at stake. Literally, this test would allow courts to strike down reasonable regulations of commercial speech on the basis of the existence, real or supposed, of some slightly less restrictive and available alternative regulation. It need not be difficult for courts to envision such alternatives,[197] with or without[198] sufficient evidence of their practicality or cost in other values, including the speech rights of other persons.

In *Board of Trustees v. Fox*, the Court specified that the "not more extensive than necessary" requirement was not to be taken literally, in the sense of a requirement that the government utilize the supposedly least restrictive means to further the state interest at issue.[199] While *Fox* in this respect lightened the burden on government regulation of commercial speech, it also required that the costs of the government regulation be "carefully calculated."[200] The discretion of judges was then enhanced by requiring that the reasonableness[201] of the degree of fit between the gov-

190. *Id.* at 569.
191. Bolger v. Youngs Drug Products Corp., 463 U.S. 60, 73 (1983).
192. United States v. Edge Broadcasting Co., 113 S. Ct. 2696, 2706 (1993).
193. *Id.*
194. *Id.*
195. *See, e.g.*, Perez v. United States, 402 U.S. 146, 152 (1971) (clearly repudiating any reliance in the commerce clause cases upon the direct versus indirect distinction); NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 40-41 (1937) (distancing itself from reliance on such a distinction); Carter v. Carter Coal, 298 U.S. 238, 307-08 (1936) (relying upon the direct versus indirect distinction).
196. *Central Hudson*, 447 U.S. at 566.
197. *See, e.g.*, Shapero v. Kentucky Bar Ass'n, 486 U.S. 466, 476 (1988) (filing of attorneys' client solicitation letters with a state agency); *In re* R.M.J., 455 U.S. 191, 206 (1982) (filing with state of copies of all attorney general mailings); Posadas de Puerto Rico Assocs. v. Tourism Co., 478 U.S. 328, 356-57 (1986) (Brennan, J., dissenting) (listing wide range of unacknowledged alternatives to the state's regulatory scheme); *see also* John M. Blim, Comment, *Free Speech and Health Claims Under the Nutrition Labeling and Education Act of 1990*, 88 Nw. U. L. REV. 733, 766 (1994) (noting the typical judicial assumption that misleading commercial ads can best be dealt with by mandated warnings and disclosures).
198. *See, e.g., In re R.M.J.*, 455 U.S. at 206.
199. *Fox*, 492 U.S. at 480-81.
200. *Id.* at 480. *But see* City of Cincinnati v. Discovery Network, Inc., 113 S. Ct. 1505, 1510 (1993) (placing the burden of proof on the government).
201. *See Fox*, 492 U.S. at 480.

ernment interest and the scope of the restriction be determined through a balancing process. The scope of the restriction must be " 'in proportion to the interest served.' "[202]

Thus the current test at this stage invites a judge to determine whether the scope of the regulation is, through some unspecified measuring process, worth its costs. Whether one restriction is actually more burdensome to freedom of speech than some envisioned alternative, however, may not be clear to begin with. An alternative regulation might affect more speakers, or affect fewer speakers more severely. The simplest calculation will involve determining whether some assumedly slight or moderate loss in fulfilling the government interest at stake outweighs, say, some reduced overall burden on freedom of commercial speech. Trading off the degree of promotion of a partially indeterminate state interest, under two alternative regulatory schemes, against the overall degree of burden on freedom of commercial speech under both regulations simply invites the expression of a court's predisposition.

The story is thus much the same at every stage of the established judicial test of restrictions on commercial speech. Any court unsympathetic with any reasonable regulation of commercial speech may, without undue strain, apply current case law to strike down the regulation at issue.

CONCLUSION

Overall, as a broad cultural institution, commercial speech can more than take care of itself without special constitutional protection under the Free Speech Clause. In our cultural context, the broad, often unintended implications underlying commercial speech are not constrained by any significant vector of cultural forces. Reasonable regulations of commercial speech, whatever their more particular and immediate justifications, would tend at least minimally to reduce the cultural dominance, within its sphere, of commercial speech. There is no reason in the Free Speech Clause not to pursue this course, in which the largely inadvertently accrued power of commercial speech would be fairly reduced for the sake of greater cultural freedom, and of the values underlying freedom of speech itself. A democratic government should at least be allowed to accept such an approach.

Special constitutional protection for commercial speech is thus currently bad enough. Things could, however, get even worse. Special protection of commercial speech may make reasonable state and federal regulation of business enterprises increasingly difficult in a number of areas, environmental[203] and securities[204] regulation being merely two examples. This would simply not be worth the cost in public values. As an

202. *Id.* (quoting *In re R.M.J.*, 455 U.S. at 203); *see also* United States v. Edge Broadcasting Co., 113 S. Ct. 2696, 2708 (1993) (Stevens, J., dissenting); McHenry v. Florida Bar, 21 F.3d 1038, 1041 (11th Cir.) (scope of restriction must be "in proportion to the interest served"), *cert. granted*, 115 S. Ct. 42 (1994).

203. *See* Peter J. Tarsney, Note, *Regulation of Environmental Marketing: Reassessing the Supreme Court's Protection of Commercial Speech*, 69 NOTRE DAME L. REV. 533, 534 (1994) ("The

admittedly extreme example, since preventing a milk producer from selling above or below a specified price obviously restricts indirectly that producer's accurate commercial speech, conceivably a business might seek to relitigate classic economic substantive due process cases, such as *Nebbia v. New York*,[205] under the Free Speech Clause.

The Supreme Court has of late sought to emphasize the value of commercial speech.[206] Without arguing that commercial speech is in all cases to be considered as valuable as political speech, the Court has in some respects impliedly equated their value. The Court has, for example, held that a city may not attack problems of safety or aesthetics by limiting only commercial speech and not political speech if commercial speech is no more related to the harm or to the state interests than is political speech.[207]

Thus the Court has created something of an Equal Protection Clause, metaphorically, for commercial speech. Unless commercial speech is more related to the harm to be regulated than is political speech, it cannot be disproportionately regulated. But this is actually a curious result. Imagine a ship that will sink unless two units of weight are jettisoned. On board are five weight units of commercial speech and five weight units of political speech. The Supreme Court has, in effect, forbidden us from tossing two units of commercial speech overboard, thereby saving most of the commercial speech and all of the political speech.

This result is simply not required by a reasonable view of our reasons for protecting free speech in the first place, or of our current cultural circumstances.[208] It instead indicates the judicial tendency to protect specially commercial speech at the expense not only of the purposes underly-

---

1993 Court raised commercial speech to a level of protection that would likely require courts to strike down many state environmental statutes.").

204. *See* Manuel S. Klausner, *The First Amendment and Commercial Speech*, 11 GEO. MASON U. L. REV. 83, 87 (1988) ("the area of federal securities legislation is ripe for the assertion of first amendment defenses"); *see also* Lowe v. SEC, 472 U.S. 181, 226-27 (1985) (White, J., concurring in result) (casting doubt on some SEC regulation of stock market newsletters by unregistered investment advisors); SEC v. Wall Street Publishing Inst., Inc., 851 F.2d 365, 371 (D.C. Cir.), *cert. denied*, 489 U.S. 1066 (1988).

205. 291 U.S. 502 (1934) (upholding a criminal conviction of Nebbia for selling milk at prices beyond those permitted by statute). These cases admittedly would involve proposals for an illegal transaction, but in which the weight of the state's interest in regulating the underlying transaction could be balanced against the severity of the restriction on commercial speech.

206. *See, e.g.*, City of Cincinnati v. Discovery Network, Inc., 113 S. Ct. 1505, 1511 (1993); *see also id.* at 1521 (Blackmun, J., concurring) ("I hope the Court ultimately will come to abandon *Central Hudson*'s analysis entirely in favor of one that affords full protection for truthful, noncoercive commercial speech about lawful activities.").

207. *See id.* at 1516; *see also* Graff v. City of Chicago, 9 F.3d 1309, 1319 (7th Cir. 1993) (noting that under *Discovery Network*, noise regulation of speech must generally apply equally to commercial and political speech), *cert. denied*, 114 S. Ct. 1837 (1994).

208. The Court may, on this basis, eventually reconsider certain other distinctions, including its refusal to apply the overbreadth doctrine to commercial speech, or to apply de novo appellate review to all cases of commercial speech. On the overbreadth issue, see Waters v. Churchill, 114 S. Ct. 1878, 1885 (1994) (no application of overbreadth doctrine to commercial speech); Board of Trustees v. Fox, 492 U.S. 469, 481 (1989) (same); Shapero v. Kentucky Bar Ass'n, 486 U.S. 466, 478 (1988) (same). For discussion of the possible application of de novo appellate review to commercial advertising cases, see Kraft, Inc. v. FTC, 970

ing freedom of speech in the first place, but of any democratically expressed view of freedom, well-being, and the proper scope and limits of purely commercial values.

F.2d 311, 317 (7th Cir. 1992) (citing Peel v. Attorney Registration & Disciplinary Comm'n, 496 U.S. 91, 108 (1990)), *cert. denied*, 113 S. Ct. 1254 (1993).

J.E.B. v. Alabama ex rel T.B.: Discrimination by any Other Name . . .

Introduction

In 1986, the Supreme Court ruled in *Batson v. Kentucky*[1] that peremptory challenges[2] motivated by race violated the Fourteenth Amendment's Equal Protection Clause.[3] In 1994, in *J.E.B. v. Alabama ex rel T.B.*,[4] the Court extended *Batson* to include gender-based peremptory challenges. Pursuant to *Batson* and *J.E.B.*, once a litigant demonstrates a prima facie case of race or gender discrimination, the party accused of discrimination must come forward with a race or gender-neutral explanation for exercising the peremptory challenge. Lower court implementation of *Batson* and lower court proscription of gender-based peremptory challenges motivated by gender prior to *J.E.B.* indicate, however, that courts readily accept explanations for peremptory challenges that are merely pretexts[5] for discrimination. This history suggests that *J.E.B.* will not, in fact, eliminate gender discrimination in the exercise of peremptory challenges. In order to rid the jury selection process of this and all types of discrimination, the peremptory challenge system should be abolished and replaced with a procedure that will prevent the injection of bigotry into American courtrooms.[6]

---

1. 476 U.S. 79 (1986).

2. A peremptory challenge is the right to challenge a juror without assigning, or being required to assign, a reason for the challenge. In most jurisdictions each party to an action, both civil and criminal, has a specified number of such challenges. After a party has used all her peremptory challenges, she is required to utilize challenges for cause to eliminate potential jurors. Black's Law Dictionary 1136 (6th ed. 1990).

3. *See* U.S. Const. amend. XIV, § 1 (providing that "[n]o state shall make or enforce any law which shall deny to any person within its jurisdiction the equal protection of the laws").

4. 114 S. Ct. 1419 (1994).

5. While trial court acceptance of litigants' pretextual explanations serves as a justification for the elimination of the peremptory challenge system, the argument for elimination is further supported by the fact that the *Batson* and *J.E.B.* decisions failed to address many important issues which will undoubtedly give rise to substantial litigation and judicial inefficiency. These issues include the absence of an adequate standard by which to determine whether a party has established a prima facie case of improper exclusion, the appropriate remedy upon a finding of improper exclusion, and the ability of a litigant to raise a *Batson* challenge where the opposing party excluded members of the panel who were of a particular race or gender, but where the petit jury nevertheless included members of the excluded class. *See generally* Albert W. Alschuler, *The Supreme Court and the Jury: Voir Dire, Peremptory Challenges and the Review of Jury Verdicts*, 56 U. Chi. L. Rev. 153 (1989).

6. While this Comment calls for the abolition of the peremptory challenge system because it engenders discrimination in the jury selection process, at least one commentator has suggested that peremptories should be eliminated because the Supreme Court cannot limit the application of *Batson* to race and gender. Eventually, argues this commentator, the Court will "have to protect everyone," and because the peremptory challenge will no longer be exercised with full freedom, it must be eliminated. J. Christopher Peters, Note, Georgia v. McCollum: *It's Strike Three for Peremptory Challenges, but Is it the Bottom of the Ninth?*, 53 La. L. Rev. 1723, 1757-58 (1993).

Part I of this Comment provides a brief background of the jury selection process and the role of the peremptory challenge. Part I also explores the history of race and gender discrimination in the jury selection process. Part II offers an in-depth discussion of the Supreme Court's decision in *J.E.B.* Part III discusses the purposes supposedly served by the peremptory challenge, as well as the harms suffered by those who are excluded from jury service due to the discriminatory exercise of this device. Additionally, Part III examines the inability of courts to distinguish between legitimate reasons for peremptory strikes and those that serve as mere pretexts for discrimination. Finally, Part IV suggests the replacement of the peremptory challenge system with an affirmative selection system that will serve to rid the jury selection process of discrimination.

I. BACKGROUND

A. *The Jury Selection Process and the Peremptory Challenge*

The concept of trial by jury has its origins in early Roman law.[7] In the Anglo-Saxon tradition, the concept first appeared in 1166 A.D. with the Assize of Clarendon,[8] which required inquiry into robbery and murder by "the twelve most lawful men."[9] In early England the members of the jury were chosen based upon their knowledge of the event at issue. This principle developed into the view that, in order to assure an impartial jury, members of the jury pool should know nothing of the instant litigation.[10] The trial by jury system immigrated to America with the colonists who traveled from England.[11] Subsequently, it became a protected right in the United States under the Sixth Amendment[12] for criminal proceedings and under the Seventh Amendment[13] for many civil proceedings.

In order to implement properly this constitutional safeguard, a fair procedure for selecting a jury is necessary.[14] While the Jury Selection and Service Act of 1968,[15] which governs jury selection, does not dictate a particular method of assembling a venire or jury panel, courts most often utilize voter registration lists to select randomly a cross section of the community. Once the venire or jury panel is selected, the judge, often with

---

7. *See id.* at 1725; *see also Batson*, 476 U.S. at 119 (Burger, C.J., dissenting) (discussing trial procedure in ancient Rome).

8. The "Assize of Clarendon" was a series of ordinances initiated by King Henry II of England in an assembly of lords at the royal hunting lodge of Clarendon. These ordinances attempted to improve procedures in criminal law and established the grand jury system consisting of twelve men. 3 ENCYCLOPEDIA BRITANNICA 348 (1985); *see also* S. Alexandria Jo, Comment, *Reconstruction of the Peremptory Challenge System: A Look at Gender-Based Peremptory Challenges*, 22 PAC. L.J. 1305, 1306 n.7 (1991).

9. Peters, *supra* note 6, at 1725-26.

10. *Id.* at 1726.

11. *Id.*

12. *See* U.S. CONST. amend. VI (providing that "the accused shall enjoy the right to a speedy and public trial, by an impartial jury" in all criminal prosecutions).

13. *See* U.S. CONST. amend. VII (providing that "the right of trial by jury shall be preserved" for all suits at common law where the value at controversy exceeds twenty dollars).

14. Peters, *supra* note 6, at 1726.

15. 28 U.S.C. §§ 1861-78 (1988).

the aid of counsel, conducts a voir dire[16] examination of the prospective jurors. Voir dire is conducted to ascertain whether the prospective jurors are acquainted with the facts of the case or the parties involved in the dispute.[17] The parties aim, through voir dire, to determine whether the prospective jurors have a predisposition regarding the merits of the case.[18] The jurors primarily are asked questions regarding their backgrounds, work, and families.[19] With this information, the litigants strive to discern whether the potential jurors harbor any biases that would hinder their ability to serve fairly and impartially as jurors.

After voir dire, the parties may strike a potential juror from the panel for cause.[20] Cause challenges must be exercised on the basis of articulated bias.[21] Such challenges allow parties to eliminate jurors for "narrowly specified, provable, and legally cognizable" reasons.[22] Pursuant to cause challenges, jurors who have exhibited actual or implied bias may be excluded. Actual bias refers to the potential juror's subjective state of mind, while implied bias is presumed by law from the existence of relationships or interests of the juror.[23] Permissible justifications for cause challenges are often codified, limiting removal of jurors to situations where, for instance, the juror has been convicted of a felony[24] or will be a witness in the litigation.[25] Trial judges are given the discretion to grant or deny cause challenges.[26]

The parties may also exercise a limited number of peremptory challenges.[27] A peremptory challenge is the right to challenge a juror without assigning, or being required to assign, a reason for the challenge. The parties may employ peremptory challenges "without a reason stated, without inquiry and without being subject to the court's control."[28]

---

16. Voir dire literally means "speak the truth." 2 WAYNE R. LAFAVE & JEROLD H. ISREAL, CRIMINAL PROCEDURE § 21.3, at 718 (1984).

17. Jo, *supra* note 8, at 1307.

18. *Id.* at 1307.

19. *Id.* at 1307.

20. Robert L. Harris, Note, *Redefining the Harm of Peremptory Challenges*, 32 WM. & MARY L. REV. 1027, 1030 (1991).

21. Karen M. Bray, Comment, *Reaching the Final Chapter in the Story of Peremptory Challenges*, 40 UCLA L. REV. 517, 519 (1992).

22. *Id.* at 519.

23. Susan L. McCoin, Note, *Sex Discrimination in the Voir Dire Process: The Rights of Prospective Female Jurors*, 58 S. CAL. L. REV. 1225, 1226 n.5 (1985).

24. ALA. CODE § 12-16-150(5) (1986); *see also* Bray, *supra* note 21, at 569 & n.4.

25. TEX. CRIM. PROC. CODE ANN. § 35.16(a)(6) (West 1989); *see also* Bray, *supra* note 21, at 569 & n.5.

26. Brent J. Gurney, Note, *The Case for Abolishing Peremptory Challenges in Criminal Trials*, 21 HARV. C.R.-C.L. L. REV. 227, 227 (1986).

27. Harris, *supra* note 20, at 1030-31.

28. Swain v. Alabama, 380 U.S. 202, 220 (1965). Generally, jurisdictions utilize one of two general challenge procedures: the "struck system" and the "sequential system." In jurisdictions that employ the struck system, the size of the jury pool is equivalent to the size of the petit jury plus the sum of peremptories allotted to both parties. If, for example, the required jury is twelve and each side is allowed five peremptory challenges, the jury pool will consist of twenty-two individuals. After the jury pool is assembled and subjected to voir dire, the parties exercise their cause challenges. Those removed for cause are replaced by other prospective jurors who also are examined in voir dire. The parties then alternate exercising their per-

Unlike the jury system as a whole, the Framers did not expressly incorporate peremptory challenges into the Sixth Amendment.[29] Despite this exclusion, the Supreme Court traditionally has given the practice an elevated position. In 1887, in *Hayes v. Missouri*,[30] the Supreme Court found that

> [e]xperience has shown that one of the most effective means to free the jurybox from men unfit to be there is the exercise of the peremptory challenge. The public prosecutor may have the strongest reasons to distrust the character of a juror offered, from his habits and associations, and yet find it difficult to formulate and sustain a legal objection to him.[31]

More recently, in the 1965 case *Swain v. Alabama*,[32] the Court implied that peremptory challenges effectuate a fair trial:

> The function of the challenge is not only to eliminate extremes of partiality on both sides, but to assure the parties that the jurors before whom they try the case will decide on the basis of the evidence placed before them, and not otherwise.[33]

This formidable history and the strong rhetoric adhered to by Supreme Court justices throughout American jurisprudential history served to fortify the lofty position the peremptory challenge has held in courtrooms in this country.[34] Recently, however, a majority of Supreme Court justices have recognized the potential and actual abuse engendered by the exercise of peremptory challenges. With the decisions in *Batson* and *J.E.B.*, the Court has begun to chip away at a legal device that no longer has a place in American society.

---

emptories against the twenty-two members of the jury pool or use all their peremptories at once, depending on the jurisdiction. *See* Gurney, *supra* note 26, at 228.

In "sequential system" jurisdictions, the number of people assembled for voir dire equals the size of the petit jury. After each individual is examined, the parties exercise both for cause and peremptory challenges. The process continues until the parties exhaust their peremptories and, after completion of the challenges for cause, enough jurors remain to form a petit jury. *Id.*

29. Harris, *supra* note 20, at 1031.

30. 120 U.S. 68 (1887).

31. *Id.* at 71; *see also* Pointer v. United States, 151 U.S. 396, 408 (1894) ("The right to challenge a given number of jurors without showing cause is one of the most important of the rights secured to the accused . . . . Any system for the empaneling of a jury that pre[v]ents or embarrasses the full, unrestricted exercise by the accused of that right, must be condemned.").

32. 380 U.S. 202 (1965).

33. *Id.* at 219.

34. While the judiciary in the United States has concentrated on eliminating racial and gender discrimination in the use of the peremptory challenge, Canada has recently equalized the number of challenges given to each side and England has eliminated the peremptory challenge entirely. Except for special circumstances, jurors under English and Canadian systems cannot be questioned about their beliefs and prejudices during voir dire nor can they be investigated prior to trial. *See* Judith Heinz, *Peremptory Challenges in Criminal Cases: A Comparison of Regulation in the United States, England, and Canada*, 16 LOY. L.A. INT'L & COMP. L.J. 201, 205-206 (1993) (comparing the history and current status of peremptory challenges in the United States, England, and Canada).

B. *The History of Racial Discrimination in Jury Selection*

In 1879, the Supreme Court in *Strauder v. West Virginia*[35] held that a West Virginia statute[36] that excluded blacks from service on grand and petit juries[37] violated the Equal Protection Clause.[38] This landmark decision established for the first time that a state could not exclude deliberately a racial group from jury service. For almost a century after *Strauder*, the Court upheld the right of racial minorities to serve on juries through proper inclusion in the jury selection process.[39] In 1965, however, the Court held in *Swain v. Alabama*[40] that a prosecutor's discriminatory intent in the use of peremptory challenges could only be demonstrated by evidence of the systematic removal of African-Americans from juries over a period of time.[41] This "crippling burden of proof"[42] effectively rendered the peremptory challenge immune from equal protection challenges.[43] The Court subsequently recognized, however, an excluded juror's interest in nondiscriminatory jury selection[44] and the harm to the judicial process caused by such discrimination.[45]

In 1986, in line with these developments, the Court overruled *Swain* in *Batson v. Kentucky*.[46] In *Batson*, the Court eased the defendant's burden of proof, ruling that evidence of a discriminatory pattern of peremptory challenges in a defendant's own trial may be sufficient to establish a prima facie case of discrimination.[47] Once a defendant establishes a prima facie case, the state must come forward with race-neutral explanations for peremptorily challenging racial minorities.[48] Since the *Batson* decision, the

---

35. 100 U.S. 303 (1879).

36. The West Virginia statute provided: "All white male persons who are twenty-one years of age and who are citizens of this State shall be liable to serve as jurors, except as herein provided." *Id.* at 305.

37. It is the duty of a grand jury to receive "complaints and accusations in criminal cases, hear the evidence adduced on the part of the state, and find bills of indictment in cases where they are satisfied a trial ought to [occur]." A petit jury is the ordinary jury called for the trial of civil or criminal actions. BLACK'S LAW DICTIONARY 855-56 (6th ed. 1990).

38. *See* U.S. CONST. amend. XIV, § 1.

39. *See* Patton v. Mississippi, 332 U.S. 463 (1947) (holding that once a prima facie case has been established, the burden is on the state to prove that the exclusion is not racially motivated); Norris v. Alabama, 294 U.S. 587 (1935) (ruling that testimony of witnesses establishing the total exclusion of blacks from jury service made out a prima facie case of the denial of equal protection). *See generally* William D. Griggs, Recent Development, 50 TENN. L. REV. 385, 389 (1983) (examining the history of racial discrimination in the jury selection process).

40. 380 U.S. 202 (1965).

41. *Id.* at 227.

42. *Batson*, 476 U.S. at 92 (overruling the burden of proof established in *Swain*).

43. Warren D. Hayes, Recent Development, State v. Knox*: The Louisiana Supreme Court Expands Equal Protection on Racially Motivated Peremptory Challenges*, 68 TUL. L. REV. 713, 715 (1994).

44. *See, e.g.*, Carter v. Jury Comm'n, 396 U.S. 320, 329 (1970).

45. *See, e.g.*, Peters v. Kiff, 407 U.S. 493, 502-03 (1972).

46. 476 U.S. 79 (1986).

47. *Id.* at 96.

48. *Id.*

Court has repeatedly affirmed and expanded the application of the Equal Protection Clause to race-based peremptory challenges.[49]

C. *The History of Gender Discrimination in Jury Selection*

During the second half of the nineteenth century, the Supreme Court held that exclusion of African-American males from jury service was unconstitutional.[50] The total exclusion of women from juries, however, endured well into the twentieth century.[51] Derived from English common law,[52] this exclusion was justified as a means by which women could be shielded from the gruesome and shocking realities of the courtroom,[53] because women were thought too delicate and naive to witness the brutal scenes depicted within the courtroom.[54]

To advance gender-motivated jury selection claims, criminal defendants have utilized both the Sixth Amendment's right to a "fair and impartial jury"[55] and the Fourteenth Amendment's Equal Protection Clause.[56] Prior to 1990, it appeared the Court would employ Sixth Amendment logic to eliminate gender-based peremptory strikes, just as criminal defendants had effectively utilized Sixth Amendment jurisprudence to rid the jury selection process of other types of gender-based discrimination.[57]

---

49. *See* Georgia v. Carr, 113 S. Ct. 30 (1992) (requiring defense to supply race-neutral explanations for peremptorily striking all white persons from the jury); Georgia v. McCollum, 112 S. Ct. 2348 (1992) (applying *Batson* to defense peremptories in criminal cases); Edmonson v. Leesville Concrete Co., 500 U.S. 614 (1991) (extending the protection of the Equal Protection Clause to private litigants in a civil case); Hernandez v. New York, 500 U.S. 352 (1991) (accepting the parties' categorization of the excluded juror as "Latino" or "Hispanic" and holding that, where members of such identifiable groups were excluded for reasons of ethnicity, the use of peremptory strikes would violate Equal Protection as interpreted by *Batson*); Powers v. Ohio, 499 U.S. 400 (1991) (ruling that a defendant could raise a *Batson* challenge even though he was not the same race as the defendant).

50. Strauder v. West Virginia, 100 U.S. 303 (1879).

51. *See, e.g.*, Hoyt v. Florida, 368 U.S. 57, 62 (1961) (holding that the exclusion of women from jury service was neither a due process nor an equal protection violation because women were "still regarded as the center of home and family life").

52. *See, e.g.*, United States v. De Gross, 960 F.2d 1433, 1438 (9th Cir. 1992) (noting that, at common law, women were excluded from juries under the doctrine of propter defectum sexus, literally the 'defect of sex') (quoting 2 WILLIAM BLACKSTONE, COMMENTARIES *362).

53. J.E.B. v. Alabama *ex rel* T.B., 114 S. Ct. 1419, 1423 (1994).

54. *Id.*

55. U.S. CONST. amend. VI. The Sixth Amendment provides:

> In 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, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor, and to have the Assistance of Counsel for his defense.

Id.

56. U.S. CONST. amend. XIV, § 1.

57. *See* Duren v. Missouri, 439 U.S. 357 (1979) (ruling that excluding women from jury venires so that they may tend to their domestic responsibilities is not a sufficiently valid interest to deprive the defendant of his or her constitutional rights guaranteed by the Sixth Amendment); Taylor v. Louisiana, 419 U.S. 522 (1975) (striking down, under the Sixth Amendment, an affirmative registration requirement that exempted women from mandatory jury service unless they volunteered to serve); Ballard v. United States, 329 U.S. 187 (1946) (holding that women may not be excluded from the venire in federal trials in states where women were eligible for jury service under local law).

In *Holland v. Illinois*,[58] however, the Court refused to extend the application of the Sixth Amendment to peremptory challenges, effectively closing the door on the possibility that the Sixth Amendment could be utilized to rid the jury selection process of gender-based peremptory strikes.[59]

Unable to eliminate gender-based peremptory strikes through application of Sixth Amendment jurisprudence, proponents of eliminating such strikes looked to the Fourteenth Amendment's Equal Protection Clause. This proved to be a difficult task. Unlike race, gender is not considered a suspect class for purposes of judicial review.[60] The Equal Protection Clause does not, therefore, subject gender-based classifications to strict scrutiny.[61] Such classifications are subject merely to intermediate review.[62] Undoubtedly influenced by the fact that gender is not a suspect classification, the Court seemingly had proscribed the application of the Equal Protection Clause to gender-motivated peremptory strikes when it refused to review the issue in three post-*Batson* cases.[63] By granting certiorari in *J.E.B.*, however, the Court agreed to examine the issue in depth.

II. J.E.B. v. ALABAMA *EX REL* T.B.

A. *Facts and Procedural History*

J.E.B. v. Alabama ex rel T.B.[64] originated in 1991 when the state of Alabama initiated a paternity and child support action against J.E.B.[65] The State alleged that J.E.B. was the father of a child born to T.B.[66] During voir dire, the state used nine of its ten peremptory challenges to strike men from the jury.[67] Subsequently, a jury of twelve women found J.E.B. to

58. 493 U.S. 474 (1990).

59. *Id.* at 480. The Court stated that although the Sixth Amendment requires a representative venire panel in order to be considered an impartial jury drawn from a fair cross section of the community, "[i]t has never included the notion that, in the process of drawing the jury, . . . initial representativeness cannot be diminished by allowing both the accused and the State to eliminate persons thought to be inclined against their interest." *Id.*

60. While sex is not a suspect class, a plurality of justices agreed that "classifications based upon sex, like classifications based upon race, alienage, and national origin, are inherently suspect and must therefore be subjected to close judicial scrutiny." Fontiero v. Richardson, 411 U.S. 677, 682 (1973) (citations omitted).

61. Strict scrutiny requires that state legislation be narrowly tailored to further a compelling governmental interest. Shaw v. Reno, 113 S. Ct. 2816, 2818 (1993).

62. Under intermediate review, classifications must serve important governmental objectives and must be substantially related to achievement of those objectives. Craig v. Boren, 429 U.S. 190, 197 (1976).

63. United States v. Nichols, 937 F.2d 1257, 1262 (7th Cir. 1991), *cert. denied*, 112 S. Ct. 989 (1992); United States v. Hamilton, 850 F.2d 1038, 1042 (4th Cir. 1988), *cert. denied*, 493 U.S. 1069 (1990); State v. Brown, 345 S.E.2d 393 (N.C.), *cert. denied*, 479 U.S. 940 (1986) (O'Conner, J., concurring) (noting that as *Batson* depends on this country's profound commitment to racial equality, it should not be applied outside of the context of race-based discrimination).

64. 114 S. Ct. 1419 (1994).

65. *Id.* at 1421.

66. *Id.*

67. *Id.* at 1422. The trial court assembled 36 potential jurors consisting of 12 males and 24 females. After the court excused three jurors for cause, 10 of the remaining jurors were male. The state then used 9 of its 10 peremptory strikes to remove male jurors. Since J.E.B. used all but one of his strikes to remove female jurors, all the selected jurors were female. *Id.* at 1421-22.

be the father of the child.[68] J.E.B. argued the state's gender-based peremptory strikes violated the Equal Protection Clause and urged the court to afford him the procedures required under *Batson*.[69] The trial court refused, concluding that *Batson* did not extend to gender-based peremptory strikes.[70]

B. *The Majority Opinion*

In *J.E.B.*, the Supreme Court affirmed the principle that in both criminal and civil litigation, potential jurors, as well as litigants, have an equal protection right to jury selection procedures that are free from state-sponsored intentional discrimination based upon group stereotypes that endorse and fortify prejudicial views.[71] The Court held that the Equal Protection Clause prohibits discrimination in jury selection based on gender or on the presumption that a potential juror will be biased simply because the person is a man or a woman.[72] According to the Court, gender, like race, is an unconstitutional "proxy" of juror capability and objectivity.[73] Utilizing the heightened scrutiny standard traditionally afforded gender-based classifications, the Court required "an exceedingly persuasive justification" for the classification.[74] The Court held that the challenges failed heightened scrutiny because they did not substantially further the state's interest in achieving a fair and impartial jury.[75] The Court refused to accept the state's argument that gender-based peremptory challenges further this interest by eliminating a group that may be partial to a particular defendant.[76] The state's stereotypic assertion, the Court opined, would not serve as justification for gender-based peremptory challenges.[77] Just as the state's generalizations would be impermissible if made on the basis of race, they were impermissible when made on the basis of gender.[78]

---

68. *Id.* at 1422. The scientific evidence presented at trial established J.E.B.'s paternity with 99.92% accuracy. *Id.* at 1437 (Scalia, J., dissenting).

69. *Id.*

70. *Id.*

71. *Id.* at 1422.

72. *Id.* at 1430.

73. *Id.* at 1421.

74. *Id.* at 1425 (citations omitted).

75. *Id.* at 1426.

76. *Id.* The state maintained that its use of gender-based peremptory strikes was based upon the belief that men, although otherwise qualified to serve on a jury, might be more sensitive and responsive to the arguments of a man alleged to be the father of a child in a paternity action, while women might be inclined to favor the complaining woman in such suits. *Id.*

77. *Id.*

78. *Id.* at 1427. Explanations for exercising peremptory strikes against people of color that were found by lower courts to be impermissibly based on race include: United States v. Chinchilla, 874 F.2d 695, 698 (9th Cir. 1989) (holding that type of employment, age, and residence were not sufficient non-race reasons under the circumstances); Roman v. Abrams, 822 F.2d 214, 228 (2d Cir. 1987) (refusing to accept the argument that potential jurors were struck because their knowledge of electronics, bookkeeping, and computers may prevent them from accepting the reasonable doubt standard), *cert. denied*, 489 U.S. 1052 (1989); United States v. Chalan, 812 F.2d 1302, 1314 (10th Cir. 1987) (holding that general references to a juror's unsatisfactory background and unspecified dissatisfaction with answers in

The Court was quick to note, however, that its decision did not imply the elimination of all peremptory challenges.[79] The Court held that its ruling was not inimical to the state's legitimate interest in utilizing such challenges to secure a fair and impartial jury.[80] In other words, parties may still use peremptory challenges to strike individuals whom they feel are less suitable than other potential jurors or who are members of a group or class not subject to heightened scrutiny.[81] The Court's decision simply disallows parties from using gender as a proxy for bias.[82] Once a prima facie showing of intentional discrimination based on gender has been made, the party exercising the strike must offer a gender-neutral explanation for the strike.[83] The party's justification need not rise to the level of a cause challenge.[84] Instead, the explanation must simply be gender-neutral and may not serve as a pretext for discrimination.[85]

C. *Concurring Opinions*

Justice O'Connor concurred in judgment with the majority but wrote separately to discuss the costs associated with the Court's ruling against gender discrimination. According to Justice O'Connor, the Court's decision will, like *Batson*, result in the proliferation of mini-hearings[86] concerning peremptory challenges and will further erode a device that plays an essential role in securing a fair and impartial jury.[87] She also asserted

---

juror's questionnaire fail to satisfy *Batson*); Pacee v. State, 816 S.W.2d. 856, 859 (Ark. 1991) (striking juror because of "demeanor" was not a sufficiently specific race-neutral explanation); People v. Arrington, 843 P.2d 62, 64-65 (Colo. Ct. App. 1992) (removing juror peremptorily because juror had pending race discrimination suit against his employer was not race-neutral and was invalid); State v. Slappy, 522 So.2d 18, 22 (Fla.) (ruling that peremptorily striking two African-Americans because they were "liberal" was not sufficiently race-neutral explanation when party failed to question the stricken jurors about their alleged bias), *cert. denied*, 487 U.S. 1219 (1988); Tolbert v. State, 553 A.2d 228, 232 (Md. 1989) (rejecting prosecution's claim that it generally strikes young females where stricken jurors were 38 and 54 years of age respectively); Commonwealth v. Harris, 567 N.E.2d 899, 904 (Mass. 1991) (holding that prosecution failed to articulate a race-neutral explanation for peremptory challenge of sole black juror when the prosecutor claimed that juror reminded him of the defendant's mother; that juror lived in location where defense witnesses lived; and that because defendant's mother had become hysterical during the arraignment, as a black woman, the juror might not be impartial); State v. Goode, 756 P.2d 578, 582 (N.M. Ct. App.) (stating that "[b]y far the most common factor noted by courts holding a state's explanations to be pretextual is a varying treatment of white and nonwhite panel members"), *cert. denied*, 756 P.2d 1203 (N.M. 1988); State v. Walker, 453 N.W.2d 127, 135-36 (Wis.) (reversing defendant's conviction where prosecution's stated reason for striking prospective juror was that the prosecution had no information about the prospective juror), *cert. denied*, 498 U.S. 962 (1990). *See* Douglas B. Dykes, Comment, *Articulation of Non-Race Based Reasons for Peremptory Challenges After* Batson v. Kentucky, 17 AM. J. TRIAL ADVOC. 245, 264-65 & n.142 (1993) (examining unacceptable reasoning in the justification of purportedly race-based peremptory challenges).

79. *J.E.B.*, 114 S. Ct. at 1429.

80. *Id.*

81. *Id.*

82. *Id.*

83. *Id.* at 1429-30.

84. *Id.* at 1430.

85. *Id.*

86. *Id.* Justice O'Connor stated that "[i]n further constitutionalizing jury selection procedures, the Court increases the number of cases in which jury selection—once a sideshow—will become part of the main event." *Id.* at 1431 (O'Connor, J., concurring).

87. *Id.* (O'Connor, J., concurring).

that by disallowing peremptory challenges based on gender, the Court decreases the litigants' ability to base their strikes on what are often accurate gender-based assumptions about juror views.[88] Justice O'Connor concluded that the prohibition of gender-based peremptory strikes should not be applied to private civil litigants or criminal defendants because they are not, in her view, state actors when they exercise peremptory challenges.[89]

Justice Kennedy, concurring in the judgment of the Court, wrote separately to explain that, pursuant to the legal framework required by equal protection analysis, precedent leads to the conclusion that the Equal Protection Clause prohibits gender discrimination in the exercise of peremptory challenges.[90] According to Justice Kennedy, just as the Equal Protection Clause forbids sex discrimination in the selection of jurors, it prohibits peremptory challenges based on sex.[91] His concurrence emphasized the importance of individual rights in equal protection analysis, including the right of an individual to participate in the political process.[92]

D. *Dissenting Opinions*

Chief Justice Rehnquist authored a dissenting opinion in which he asserted that race and gender discrimination are different.[93] In *Batson*, the Court balanced the practice of peremptory challenges with the commands of equal protection and held that in the case of race-based peremptories equal protection was superior.[94] Chief Justice Rehnquist concluded that the differences between race and gender discrimination, however, indicate that when sex, not race, is at issue, the scales should tilt in favor of peremptory challenges.[95]

Justice Scalia, joined by Chief Justice Rehnquist and Justice Thomas, also dissented. Justice Scalia asserted that since all groups are subject to peremptory challenges, gender-based peremptories do not result in the denial of equal protection.[96] According to Justice Scalia, the Court's decision simply demonstrates the Justices' politically correct views in matters pertaining to the sexes.[97] The result of this decision, which is neither

88. Justice O'Connor noted studies indicating that in rape cases, female jurors are somewhat more likely to convict than male jurors. Furthermore, she asserted that while there have been no definitive studies, it is clear that a person's gender and resulting life experience will impact an individual's view in sexual harassment, child custody, and spousal and child abuse cases. *Id.* at 1432 (O'Connor, J., concurring).

89. *Id.* at 1432-33 (O'Connor, J., concurring).

90. *Id.* at 1433 (Kennedy, J., concurring).

91. *Id.* (Kennedy, J., concurring).

92. *Id.* at 1433-34 (Kennedy, J., concurring).

93. *Id.* at 1435 (Rehnquist, C.J., dissenting). The Chief Justice argued that while race-based classifications are subject to strict scrutiny, gender-based classifications are reviewed under a heightened, but less strict standard. Furthermore, while racial groups make up numerical minorities in American society, the population is nearly equally divided between men and women. Finally, according to the Chief Justice, racial equality has proven to be a more difficult goal to achieve than gender equality. *Id.* (Rehnquist, C.J., dissenting).

94. *Id.* (Rehnquist, C.J., dissenting) (citing *Batson*, 476 U.S. at 98-99).

95. *Id.* (Rehnquist, C.J., dissenting).

96. *Id.* at 1437 (Scalia, J., dissenting).

97. *Id.* at 1436 (Scalia, J., dissenting).

mandated nor permitted by the Constitution, is the erosion of a practice that historically has been an essential aspect of the right to a fair jury trial.[98]

III. ANALYSIS

A. *The Benefits and Harms Engendered by the Use of the Peremptory Challenge*

Proponents of the peremptory challenge point to its two distinct purposes. First, it serves as a "safety net" for challenges exercised for cause.[99] The peremptory allows the parties to reject a juror for a partiality they cannot name or explain.[100] Thus, because it enables the parties to remove presumably biased jurors, the peremptory challenge is considered one of the most effective means of securing a fair and impartial jury.[101]

Furthermore, the peremptory challenge furthers the symbolic legitimacy of a fair jury trial.[102] This view finds its roots in fourteenth century England, when Parliament eliminated the power of the king's attorneys to exercise peremptory challenges.[103] This action demonstrated the symbolic significance of a defendant's opportunity to play an active role in the determination of the composition of the jury.[104] The peremptory challenge is:

> 'a provision full of that tenderness and humanity to prisoners, for which our English laws are justly famous.' With this tool, the defendant may dismiss from the jury, for any unspoken reason, those he most hates or fears so that he is left with a 'good opinion of the jury, the want of which might totally disconcert her.'[105]

Pursuant to this principle, the defendant, who helped choose the jury, will be satisfied with its composition and, therefore, the verdict will appear fair. Further, because the verdict appears fair to the defendant, the community will also be confident that the judgment is just.[106] Although the modern American judicial system allows prosecutors, as well as defendants, to exercise peremptory challenges, the peremptory challenge still, in some respects, retains the symbolic character first recognized in medieval England.

Although the peremptory challenge may serve these important purposes, litigants' continued discriminatory exercise of this device and the resultant harms of this discrimination call for the elimination of the peremptory challenge altogether. When a litigant is allowed to exercise her

---

98. *Id.* at 1439 (Scalia, J., dissenting).
99. Hayes, *supra* note 43, at 721.
100. *See, e.g.*, MICHAEL SAKS & REID HASTIE, SOCIAL PSYCHOLOGY IN COURT 55 (1978) (asserting that jury selection was, for centuries, the product of "hunches, unsymptomatic past experience, intuition, [or] stabs in the dark"); *see also Batson*, 476 U.S. at 138 (Rehnquist, J., dissenting) (asserting that peremptory challenges are based on "seat of the pants instincts").
101. *Swain*, 380 U.S. at 217-18.
102. Tracey L. Altman, Note, *Affirmative Selection: A New Response to Peremptory Challenge Abuse*, 38 STAN. L. REV. 781, 794 (1986).
103. *Id.*
104. *Id.*
105. *Id.* (quoting 4 WILLIAM BLACKSTONE, COMMENTARIES *353) (citations omitted).
106. *Id.*

peremptory challenges in a discriminatory manner several harms result. First, without the broad range of social experiences often found in groups comprised of both sexes, juries may be ill-equipped to evaluate the facts presented.[107] For example, all male juries may not understand the fear and helplessness felt by battered wives who, in self defense, wound or murder their batterers.[108] Misunderstanding important testimony relating to gender issues, such as spousal abuse, can create the opportunity for unconscious prejudice.[109]

Secondly, when potential jurors are excluded from juries because of their gender, those excluded are deprived of their basic democratic right to participate in the community's administration of justice.[110] Like the right to vote, jury service is one of the most fundamental ways an individual citizen can participate in the democratic process.[111] Serving as a juror can be an empowering experience, especially for women who historically have been subjected to gender discrimination. When women are excluded from jury service because of their gender they are stigmatized by the implication that they are not equals and that they are unable or unwilling to be impartial.[112]

Finally, discriminatory use of peremptory challenges undermines the legitimacy of and confidence in the fairness of the justice system.[113] Women who are excluded because they are women see that the law is treating them unequally with respect to jury service, and they may come to believe that the law will treat them unfairly in other contexts as well.[114] Fairness to litigants, inclusion of citizens of both sexes, and the integrity of the justice system all demand that discrimination be recognized and eliminated in the exercise of peremptory challenges. As the following discussion demonstrates, in order to achieve this goal, peremptory challenges must be abolished.

B. *Pretextual Reasoning*

J.E.B. will not eradicate gender discrimination in the exercise of peremptory challenges. Lower courts applying *Batson* and lower courts attempting to ban gender-based peremptory challenges prior to *J.E.B.* have found it difficult to distinguish between legitimate race and gender-neutral reasons for peremptory strikes and mere pretexts for discrimination. When lower courts implement *J.E.B.*, they necessarily will have similar difficulties. While the *J.E.B.* decision is a historic step toward the elimination

---

107. Theodore McMillan & Christopher J. Petrini, Batson v. Kentucky*: A Promise Unfulfilled*, 58 UMKC L. REV. 361, 362 (1990).

108. *See, e.g.*, Deborah L. Forman, *What Difference Does it Make? Gender and Jury Selection*, 2 UCLA WOMEN'S L.J. 35 (1992).

109. Note, *Developments in the Law: Race and the Criminal Process*, 101 HARV. L. REV. 1472, 1559 (1988).

110. McMillan & Petrini, *supra* note 107, at 352.

111. *Id.*

112. *Id.*

113. *Id.*

114. *Id.*

of gender discrimination in the jury selection process, it will simply eliminate the most blatant discriminatory peremptory challenges. Thinly disguised pretexts for gender discrimination, on the other hand, may survive judicial examination. In light of this difficulty, the effort to remove discrimination from the jury selection process would be best served if peremptory challenges were eliminated altogether.

1. Pretextual Reasoning in Race-Based Peremptory Challenges

The evidentiary analysis required by *J.E.B.*[115] is identical to the analysis established by the Court in *Batson.*[116] Lower courts have had difficulty, however, implementing *Batson,* which suggests that the application of *J.E.B.* will be equally problematic. The difficulty arising from the application of *Batson* stems from litigants' use of pretextual reasoning in the justification of peremptory challenges. Litigants have become proficient at offering acceptable reasons for their strikes,[117] and the courts have readily accepted these often pretextual explanations, which enables attorneys to avoid the commands of *Batson.*[118]

Significantly, the Supreme Court has provided almost no guidance to the lower courts in assessing purportedly race-neutral explanations offered by litigants.[119] In *Batson,* the Court merely noted that a litigant "must give a clear and reasonably specific explanation of his legitimate reasons for exercising the challenges."[120] While the Court held that a litigant's reasons must be related to the particular case to be tried,[121] this requirement has failed to provide the lower courts with an adequate standard by which to examine a litigant's proffered explanations.

Pursuant to this indefinite standard, lower courts have accepted justifications based on non-racial characteristics that are, in fact, merely racial and ethnic surrogates.[122] Justifications for peremptories are clearly proxies for race and ethnicity when there is a dramatic statistical correlation between the trait and race or ethnicity.[123] Studies, for example, have indicated that residence, especially in urban areas, often serves as a surrogate

---

115. *See J.E.B.*, 114 S. Ct. at 1430.

116. *Batson*, 476 U.S. at 97.

117. *Id.* at 106 (Marshall, J., concurring).

118. Andrew G. Gordon, Note, *Beyond* Batson v. Kentucky*: A Proposed Ethical Rule Prohibiting Racial Discrimination in Jury Selection*, 62 FORDHAM L. REV. 685, 694 (1993) (discussing lower court acceptance of pretextual explanations).

119. *Id.*; *see* Hernandez v. New York, 500 U.S. 352, 371-72 (1991) (finding it unnecessary to address the issue of pretext where prosecutor struck jurors because of their Spanish-speaking ability).

120. *Batson*, 476 U.S. at 98 n.20 (internal quotation marks omitted).

121. *Id.* at 98.

122. Alschuler, *supra* note 5, at 175.

123. *See* Deborah A. Ramirez, *Excluded Voices: The Disenfranchisement of Ethnic Groups From Jury Service*, 1993 WIS. L. REV. 761, 789-91 (discussing Hernandez v. New York, 500 U.S. 352, 371-72 (1992), and the statistical impact of excluding spanish-speaking latino jurors).

for race or ethnicity.[124] As one court has stated, "[r]esidence . . . acts as an ethnic badge . . . [and] can be the most accurate predictor of race."[125]

In *United States v. Uwaezhoke*,[126] for example, an African-American defendant was on trial for participation in a drug conspiracy.[127] The defendant objected when the prosecutor peremptorily challenged an African-American woman.[128] The prosecutor explained that the woman was stricken because, as a postal worker and a single parent who rented an apartment in Newark, New Jersey, she "may be involved in a drug situation where she lives."[129] The court held that the explanation was facially race-neutral.[130] The prosecutor's explanation, however, was clearly based upon unfounded stereotypes. He simply assumed that a single, black woman living in Newark would reside in low income housing and that areas of low-income housing are drug infested.[131] Thus, the potential juror's race, her neighborhood, and crime and violence became amalgamated, one serving as a surrogate for another.[132]

Additionally, lower courts have accepted justifications for peremptory strikes against people of color where white jurors, exhibiting the same characteristics that presumably prompted the peremptory strikes, remain on the jury.[133] While some courts have held that the unequal application

---

124. *See, e.g.*, Michael F. Potter, Note, *Racial Diversity in Residential Communities: Societal Housing Patterns and a Proposal for a "Racial Inclusionary Ordinance"*, 63 S. CAL. L. REV. 1151, 1154 (1990) (finding that race determines housing patterns); Richard H. Sander, Comment, *Individual Rights and Demographic Realities: The Problem of Fair Housing*, 82 Nw. U. L. REV. 874, 875 (1988) (stating that "[e]very major metropolitan area in the United States still has a large ghetto; in many cities, over eighty percent of the black population lives in virtually all-black neighborhoods").

Litigants have argued that other characteristics also serve as racial and ethnic surrogates. *See, e.g.*, United States v. Mixon, 977 F.2d 921, 923 (5th Cir. 1992) (insufficient education); United States v. Hinojosa, 958 F.2d 624, 631-32 (5th Cir. 1992) (same); United States v. Hughes, 970 F.2d 227, 231 (7th Cir. 1992) (relatives with criminal records); United States v. Johnson, 941 F.2d 1102, 1106-07 (10th Cir. 1991) (same); United States v. Payne, 962 F.2d 1228, 1233 (6th Cir.) (group membership), *cert. denied*, 113 S. Ct. 811 (1992); United States v. Clemmons, 892 F.2d 1153, 1157 (3d Cir. 1989) (religion), *cert. denied*, 496 U.S. 927 (1990); United States v. Woods, 812 F.2d 1483, 1487 (4th Cir. 1987) (same); United States v. Cartlidge, 808 F.2d 1064, 1071 (5th Cir. 1987) (lack of substantial income); *see also* Gordon, *supra* note 118, at 699-705 (discussing courts' willingness to accept explanations for peremptories that are racial or ethnic surrogates).

125. United States v. Bishop, 959 F.2d 820, 828 (9th Cir. 1992).

126. 995 F.2d 388 (3d Cir. 1993), *cert. denied*, 114 S. Ct. 920 (1994).

127. *Id.* at 389.

128. *Id.*

129. *Id.*

130. *Id.* at 393.

131. Gordon, *supra* note 118, at 700-01.

132. "Through mental association, African-Americans, their neighborhoods, crime and violence all become amalgamated, giving rise to tenacious stereotypes—innocent and unintentional perhaps, but stereotypes nonetheless." *Bishop*, 959 F.2d at 828 (9th Cir. 1992). *Bishop* held that the peremptory challenge of an African-American woman because she lived in Compton, a neighborhood in South Central Los Angeles where seventy-five percent of the residents were African-American, was a discriminatory racial proxy. *Id.* at 822, 827; *see also* Gordon, *supra* note 118, at 702, 740 (discussing purportedly neutral explanations that are, in fact, racial and ethnic surrogates).

133. Gordon, *supra* note 118, at 706.

of peremptory challenges is unacceptable,[134] other courts have allowed such strikes even though they are clearly pretextual.[135]

In *United States v. Alvarado*,[136] for example, the prosecutor peremptorily struck an African-American juror with children the age of the defendant, claiming she might be unduly sympathetic.[137] The prosecutor, however, failed to strike white members of the venire with children of an age similar to the defendant's.[138] The trial court accepted the proffered rationale.[139] Although the Second Circuit noted that the force of a challenger's explanation for strikes against people of color is substantially weakened by evidence that white members, to whom the same explanation applies, were not challenged, the judicial officer was entitled to assess each explanation in light of all the relevant evidence.[140] Despite evidence indicating the explanation was pretextual, the Second Circuit upheld the trial court's finding.[141]

Finally, peremptory challenges based on attorneys' subjective impressions are often accepted by lower courts.[142] As such impressions are essen-

---

134. *See, e.g.*, Jones v. Ryan, 987 F.2d 960 (3d Cir. 1993) (refusing to accept the prosecution's proffered explanations because they were applied only to African-American jurors); United States v. Chinchilla, 874 F.2d 695, 698 (9th Cir. 1989) (disallowing strike against Hispanic who resided in La Mesa, California because prosecutor failed to strike a white juror who lived in La Mesa); *see also* Gordon, *supra* note 118, at 706-07, 719 n.230 (citing *Jones v. Ryan*).

135. *See, e.g.*, United States v. Clemons, 941 F.2d 321, 324 (5th Cir. 1991) (accepting proffered explanation of youth, although nineteen-year-old white juror remained unchallenged); United States v. Williams, 936 F.2d 1243, 1246 (11th Cir. 1991) (accepting rationale for strike of African-American woman because of her previous association with defense counsel, even though several white jurors also had contact with defense counsel), *cert. denied*, 112 S. Ct. 1279 (1992); United States v. Bennett, 928 F.2d 1548, 1551 (11th Cir. 1991) (allowing youth, unemployment, and relatives with drug convictions as reasons for strikes of African-American jurors, although one white juror was young and unemployed and another had been convicted of drug charges); Barfield v. Orange County, 911 F.2d 644, 648-49 (11th Cir. 1990) (allowing strike of African-American woman because she worked for the school board, while white women who worked for the school board were not stricken), *cert. denied*, 500 U.S. 954 (1991); United States v. Alston, 895 F.2d 1362, 1367 n.5 (11th Cir. 1990) (accepting strikes against African-Americans based on age, family drug problems, and misunderstanding voir dire questions, although white jurors exhibiting the same characteristics were not stricken); United States v. Lance, 853 F.2d 1177, 1180 (5th Cir. 1988) (allowing strike of African-American, who was young and single, even though white juror exhibited the same characteristics); *see also* Gordon, *supra* note 118, at 719 n.231.

136. 951 F.2d 22 (2d Cir. 1991).

137. *Id.* at 24.

138. *Id.* at 25.

139. *Id.*

140. *Id.* at 25-26.

141. *Id.* at 26.

142. Courts have accepted a variety of subjective impressions as explanations. *See, e.g.*, Brown v. Kelly, 973 F.2d 116, 119 (2d Cir. 1992) (impressions of attitude; nervousness and tone of voice), *cert. denied*, 113 S. Ct. 1060 (1993); United States v. Vaccaro, 816 F.2d 443, 457 (9th Cir.) (same), *cert. denied*, 484 U.S. 928 (1987); Dunham v. Frank's Nursery & Crafts, Inc., 967 F.2d 1121, 1124-25 (7th Cir. 1992) (impressions of eye contact); Reynolds v. Benefield, 931 F.2d 506, 512 (8th Cir.) (same), *cert. denied*, 501 U.S. 1204 (1991); United States v. Clemons, 941 F.2d 321, 323-24 (1991) (impressions of dress); United States v. Sherrills, 929 F.2d 393, 395 (8th Cir. 1991) (impressions of inattentiveness during voir dire); United States v. Rudas, 905 F.2d 38, 41 (2d Cir. 1990) (same); United States v. Hendrieth, 922 F.2d 748, 749-50 (11th Cir. 1991) (jurors' facial expressions); United States v. Ruiz, 894 F.2d 501, 506 (2d Cir.) (same), *aff'd*, 894 F.2d 501 (2d Cir. 1990); United States v. Terrazas-Carrasco, 861 F.2d

tially unverifiable, they can easily hide discriminatory intent.[143] Judges rarely notice a particular juror's mannerisms and trial courts are not in a position to assess an explanation based upon a subjective opinion.[144] Furthermore, appeals of such challenges are difficult, as written records cannot support or negate an attorney's subjective impression.[145]

In *Barfield v. Orange County*,[146] for instance, a prosecutor claimed that an African-American juror was peremptorily struck because she "was looking at me, and looking at my client, and looking at the Defendant's table with an expression that conveyed to me some hostility, and it was my gut feeling, based on her facial expression that she was likely to not be fair and impartial to the [defendant]."[147] The trial court accepted the prosecutor's explanation, and the Eleventh Circuit affirmed, ruling that "[h]ostile facial expressions and body language are legitimate" rationales for the exercise of a peremptory challenge.[148] Subjective impressions, however, such as that offered by the prosecutor in *Barfield*, are often affected by an attorney's realized or unrealized racism.[149] Similarly, a judge's own bigotry may lead her to accept an attorney's rationale as plausible.[150]

Because of the similarities between race and gender discrimination present in both the character of the prejudice and the analysis required by the court, the problematic nature of the *Batson* decision necessarily foreshadows the difficulties lower courts will encounter in the application of *J.E.B.* Under *Batson*, litigants satisfy their rebuttal burdens by reciting unreviewable explanations that are merely pretexts for racial discrimination. Because it is simply too difficult to review intelligently and accurately litigants' motives in the peremptory exclusion of jurors in the racial context and, by implication, in the gender context, the peremptory challenge system should be eliminated.

2. Pretextual Reasoning in Gender-Based Peremptory Challenges

Prior to the Court's decision in *Batson*, state courts, unable to utilize Sixth Amendment jurisprudence in the examination of the validity of gender-based peremptories,[151] "turned to their own state constitutions and an analysis of the Equal Protection Clause of the United States Constitution."[152] Therefore, years before the Supreme Court decided *Batson* and *J.E.B.*, numerous state courts disallowed the use of both race-based and

---

93, 95 n.1 (5th Cir. 1988) (impressions of body language); United States v. Lance, 853 F.2d 1177, 1181 (5th Cir. 1988) (impressions of demeanor); United States v. Forbes, 816 F.2d 1006, 1010-11 (5th Cir. 1987) (same); United State v. Cartlidge, 808 F.2d 1064, 1071 (5th Cir. 1987) (same); *see also* Gordon, *supra* note 118, at 719 nn.243-252 (citing court decisions where peremptory challenges based on litigants' subjective impressions were accepted).

143. Gordon, *supra* note 118, at 709.
144. *Id.*
145. *See id.*
146. 911 F.2d 644 (11th Cir. 1990), *cert. denied*, 500 U.S. 954 (1991).
147. *Id.* at 646.
148. *Id.* at 648.
149. *See Batson*, 476 U.S. at 106 (Marshall, J., concurring).
150. *Id.*
151. *See* discussion *supra* Part I.C.
152. *See, e.g.*, Jo, *supra* note 8, at 1315.

gender-based peremptories.[153] Furthermore, after the Court's decision in *Batson*, several state and federal courts held that the *Batson* rationale extended to peremptory strikes motivated by gender.[154] Thus, while lower courts have not yet specifically applied *J.E.B.*, an analysis of lower court cases that proscribed the use of gender-based peremptory challenges demonstrates that, while *J.E.B.* may eliminate the most flagrant instances of gender discrimination, the implementation of the decision will prove problematic because courts routinely accept the pretextual explanations offered by attorneys in defense of discriminatory peremptory challenges.

Like many purportedly race-neutral explanations, facially neutral rationales offered for the peremptory removal of women are often simply surrogates for gender. Justifications for peremptories are surrogates for gender when there is a high statistical correlation between the trait and gender. Just as residence may serve as a proxy for race, employment in a particular field often serves as a surrogate for gender. For example, women enter the field of nursing in dramatically higher numbers than men, demonstrating a correlation between gender and nursing.[155]

In *State v. Burch*,[156] a pre-*J.E.B.* decision, the Washington Court of Appeals held that peremptory challenges on the basis of gender violated both federal equal protection and the state's equal rights amendment.[157] The *Burch* court held that, under the circumstances, the defendant had established a prima facie case of purposeful gender discrimination.[158] During rebuttal, the prosecutor claimed that one of the excluded women jurors had been stricken because, as a nurse who worked at a women's clinic, she would find the defense witness, also a woman, "extremely credible."[159] The court found this explanation sufficiently gender-neutral.[160]

The prosecutor's rationale, however, was the product of uncorroborated stereotypes about women. She assumed that a nurse in a women's clinic who acted as a caregiver to women would be sympathetic towards a female witness and unable to be impartial. This explanation is not gender neutral because it is "founded on gender stereotypes which view women as generally governed by emotion, instinct, and feeling rather than reason, judgment or common sense."[161] In this sense, the court allowed the fu-

153. *See, e.g.*, People v. Wheeler, 583 P.2d 748 (Cal. 1978); Commonwealth v. Soares, 387 N.E.2d 499, 516 (Mass. 1979).

154. *See, e.g.*, United States v. De Gross, 913 F.2d 1417, 1426 (9th Cir. 1990); New York v. Irizarry, 195 N.Y.S.2d 279, 280 (N.Y. App. Div. 1990). *But see* United States v. Hamilton, 850 F.2d 1038, 1041 (4th Cir. 1988), *cert. denied*, 493 U.S. 1069 (1990); State v. Oliviera, 534 A.2d 867, 870 (R.I. 1987).

155. Telephone interview with Lori Brotzman, Program Assistant, Office of Student and Academic Support, University of Colorado School of Nursing, (June 2, 1994) (stating that 647 women and 43 men are enrolled in the University of Colorado School of Nursing).

156. 830 P.2d 357 (Wash. Ct. App. 1992).

157. *Id.* at 361-63.

158. *Id.* at 365.

159. *Id.* at 366.

160. *Id.*

161. *Id.*

sion of the potential juror's sex, her profession, and a particular set of characteristics, enabling one to serve as a surrogate for another.[162]

Furthermore, as in the context of race-based peremptory challenges,[163] courts have accepted justifications for peremptories against women where men jurors with the same characteristics remain on the jury. Although the mere presence of the unequal application of peremptories evidences pretextual reasoning, courts continue to rule that such application does not warrant a finding of improper use of peremptories.

The court in *People v. Irizarry*[164] found, pursuant to state prohibitions and the federal constitution, gender-based peremptory challenges were unlawful.[165] The court held that the prosecution's challenges against female jurors established a prima facie case of gender-based discrimination.[166] Against a panel of an unusually high number of men, the prosecutor challenged nine women and only one man.[167] Furthermore, the manner in which the peremptories were exercised showed an apparent attempt to rid the jury of women.[168]

The challenged women had diverse employment and family backgrounds, and, for the most part, the voir dire provided no anti-prosecution inclination.[169] In fact, several of the women had relatives who were police officers, and others had been crime victims.[170] The men who were not challenged disclosed information about themselves similar to the facts revealed by the challenged women.[171] After the prosecution gave specific explanations for the removal of the female jurors, the court ruled that seven of the nine peremptory challenges were exercised for reasons independent of gender.[172] Despite the prosecutor's failure to strike male members of the venire with characteristics similar to those of the challenged females, the court ruled that the prosecutor's use of peremptory challenges was not pretextual and, thus, not unlawful.

Finally, courts have accepted explanations for the peremptory removal of prospective jurors based on litigants' subjective impressions. As in the context of race,[173] subjective impressions can camouflage unlawful motives. Litigants wishing to keep jurors of a particular sex off the jury can merely provide the court with a subjective rationale based on an unverifiable impression, and, as noted above, courts are not equipped to judge the validity of such impressions.

---

162. *See supra* note 132 and accompanying text.
163. *See supra* notes 133-41 and accompanying text.
164. 536 N.Y.S.2d 630 (N.Y. App. Div. 1988).
165. *Id.* at 638.
166. *Id.* at 643.
167. *Id.* at 644.
168. *Id.*
169. *Id.*
170. *Id.*
171. *Id.*
172. *Id.* at 645.
173. *See supra* notes 142-50 and accompanying text.

In *Mouzon v. Philadelphia Housing Authority*,[174] a woman employee brought suit against her employer alleging gender bias and sexual harassment.[175] The plaintiff claimed the defendant's three peremptory challenges were exercised against female jurors on the basis of gender in violation of the Equal Protection Clause.[176] During voir dire, one of the excluded female jurors said she would find sexually explicit testimony embarrassing.[177] The defendant claimed that, due to this statement and her demeanor at the time she made the statement, the juror was hypersensitive and an unacceptable juror.[178] The court found this explanation to be gender neutral.[179] It is possible, however, that the juror appeared hypersensitive, especially with respect to sexual matters, simply because traditionally women are seen as virginal and innocent. Thus, the attorney's subjective impression of the juror, and the court's subsequent acceptance of his explanation, may have been affected by sexist attitudes.

In light of lower court experience concerning the evaluation of litigants' purportedly gender-based explanations, it is apparent that *J.E.B.* is an imperfect remedy for gender discrimination in the exercise of peremptory challenges. Lower courts are not equipped to assess explanations that are surrogates for gender-based classifications. The eradication of gender discrimination in the context of the peremptory challenge can only be achieved through the elimination of this device.

IV. AFFIRMATIVE SELECTION

As discussed above, the peremptory challenge serves as a "safety net" for challenges exercised for cause and represents the symbolic legitimacy of the jury system.[180] As the preceding discussion illustrates, however, despite judicial regulation of peremptory challenges, such challenges allow litigants to discriminate in the jury selection process.[181] A new system should be introduced into American courtrooms that will serve the valid goals of peremptory challenges and preclude discrimination in the jury selection process. An affirmative selection[182] system will accomplish these tasks.

---

174. No. CIV.A.93-3686, 1994 WL 197165 (E.D. Pa. May 19, 1994).
175. *Id.* at *1.
176. *Id.*
177. *Id.*
178. *Id.*
179. *Id.* at *2.
180. *See* discussion *supra* Part III.A.
181. *See* discussion *supra* Parts III.B.1. and III.B.2.
182. This method was originally proposed in Tracey L. Altman, Note, *Affirmative Selection: A New Response to Peremptory Challenge Abuse*, 38 STAN. L. REV. 781 (1986). Since the publication of Altman's Note, numerous scholars have endorsed this method, including Alschuler, *supra* note 5; Harris, *supra* note 20; Heinz, *supra* note 34; and Hans Zeisel, *Affirmative Peremptory Jury Selection*, 39 STAN. L. REV. 1165 (1987).

A number of commentators have suggested alternatives other than affirmative selection, including the replacement of peremptories with a system of cause challenges, Gurney, *supra* note 26, at 257-62; the expansion of the jury pool, *id.* at 262-66; the improvement of the effectiveness of cause challenges, *id.* at 266; disallowing reliance on jurors' self-assessment of bias, *id.* 266-68; the improvement of voir dire, *id.* at 268-73; the reduction of the number of peremptories given to each side, *id.* at 274; the enactment of statutes that would restrict the

A. *The Procedural Requirements*

Like the current jury selection process, in an affirmative selection system, twenty-four venirepersons would be drawn randomly from the jury pool.[183] These twenty-four individuals would then be subject to voir dire and challenges for cause.[184] After challenges for cause, each party would choose twelve jurors and list them in order of preference. The lists would be submitted to the judge. Regardless of their rank, the judge would seat on the petit jury those venirepersons who appear on both lists.[185] Next, alternating between the lists, the judge would seat each party's selections in descending order until the proper number of jurors is reached.[186]

B. *Affirmative Selection Satisfies the Traditional Objectives of the Peremptory Challenge*

While the peremptory challenge system acts as a "safety net" for challenges exercised for cause and presumably allows the parties to secure an impartial jury,[187] the affirmative selection system also allows for the selection of an impartial jury. First, affirmative selection guarantees a fair contest between the litigants. Since the parties have equal power to select jurors, neither party has the opportunity to stack the jury in her favor.[188] Furthermore, because jurors who appear on the lists of both parties are seated first, jurors who are suitable to both parties become part of the petit jury. Finally, because the parties have equal ability to choose acceptable

---

application of peremptory challenges, Barbara A. Babcock, *A Place in the Palladium: Women's Rights and Jury Service*, 61 U. CIN. L. REV. 1139, 1176 (1993); the incorporation of an ethical rule into the American Bar Association's Model Rules of Professional Conduct that would prohibit discrimination against members of the venire during jury selection, Gordon, *supra* note 118, at 713; and the adoption of a due process standard, Note, *Due Process Limits on Prosecutorial Peremptory Challenges*, 102 HARV. L. REV. 1013 (1989).

183. Altman, *supra* note 182, at 806.

184. *Id.*; *see* discussion *supra* Part I.A.

185. Altman, *supra* note 182, at 806.

186. *Id.* One commentator, in an effort to expand Altman's affirmative selection system, recommends that after the judge seats those jurors who appear on both parties' lists the remaining jurors should be subject to removal if, to the court's satisfaction, the moving party provides a neutral explanation justifying their dismissal. The trial judge should permit these challenges alternately, allowing each side to utilize as many strikes as the peremptory challenge system allows. Under this system, the trial judge should assess the similarities and differences between the challenged and unchallenged venirepersons to determine whether the explanations for removal are based on specific biases not shared by other panel members. It is in this manner, argues the commentator, that the trial judge should determine whether the explanations are merely pretexts for discrimination. If both parties expend their challenges and the petit jury has not been empaneled, the judge should randomly select the remaining jurors from the jury panel. Harris, *supra* note 20, at 1063.

This variation on Altman's affirmative selection system fails to remedy the problem of pretextual reasoning engendered by peremptory challenges. As in the peremptory challenge context, courts are not equipped to evaluate the validity of a litigant's purportedly neutral explanation. A superficial examination of the similarities and differences between challenged and unchallenged venirepersons will not enable the judge to ferret out explanations that are, in fact, surrogates for race or gender-based classifications, unequally applied to individuals on the venire panel, or based on a litigant's subjective impression of a particular panel member. *See* discussion *supra* Parts III.B.1. and III.B.2.

187. *See* discussion *supra* Part III.A.

188. Altman, *supra* note 182, at 807.

jurors if there is an insufficient number of mutually acceptable jurors, each has the opportunity to seat those individuals whom she thinks may be partial toward her case.[189] Therefore, assuming the litigants' choice of jurors is influenced by their respective interests, affirmative selection produces a balanced jury. It "lets the parties determine which biases or community members are important to their cases, and their competing interests should affect a balance of biases on the jury."[190]

While the symbolic legitimacy engendered by the defendant's ability to shape the composition of the jury through the exercise of peremptory challenges is an important purpose of the challenge, this objective can be accomplished in an equally satisfactory manner through the implementation of an affirmative selection system. Through this system, the defendant has the opportunity to choose one half of the jury and the jury will not be seen as merely an extension of the prosecution.[191] Furthermore, to the extent that affirmative selection allows for the inclusion of more minorities and for a more equal representation of both genders, "it promotes public confidence in the jury since popular participation instills respect for the system."[192] The replacement of peremptory challenges with affirmative selection will not destroy the confidence litigants and the public have in the jury system, but will, in fact, promote it. Like the peremptory challenge system, affirmative selection produces an impartial jury and promotes the symbolic legitimacy of the jury trial. Unlike the peremptory challenge system, however, affirmative selection does not create a forum in which pretexts for discrimination are accepted as reasonable explanations for the exclusion of individuals from petit juries.[193]

CONCLUSION

Eliminating gender-motivated peremptory challenges will take more than the pronunciation by the Court in *J.E.B.*, because, despite the law, attorneys will continue to employ peremptories in a discriminatory manner. The case of race-based peremptory challenges provides an unfortunate example of the persistence of racial stereotypes in courtrooms through the use of pretextual reasoning. Similarly, in state and federal courts where gender-motivated peremptory challenges have been disallowed, explanations given by attorneys and subsequently accepted by judges to justify the removal of women jurors echo attempts to exclude African-American jurors while presumably adhering to the mandates of *Batson*. Experience, thus, demonstrates the necessity of abolishing the peremptory challenge system in order to rid our jury selection process of discrimination, not just in theory, but in fact.

Pamela R. Garfield

189. *Id.*
190. *Id.*
191. *Id.* at 808.
192. *Id.*
193. *See* discussion *supra* Parts III.B.1. and III.B.2.

CENTRAL BANK OF DENVER V. FIRST INTERSTATE BANK: PRUNING THE JUDICIAL OAK BY SEVERING THE AIDING AND ABETTING BRANCH

Congress and those charged with enforcement of the securities laws stand forewarned that unresolved questions concerning the scope of [private 10(b)] causes of action are likely to be answered by the Court in favor of defendants[1]

INTRODUCTION

Section 10(b) of the Securities Exchange Act of 1934 makes it unlawful to use deceptive devices or make misleading statements "in connection with the purchase or sale of any security." Under section 10(b), private actions have been said to represent a "judicial oak which has grown from little more than a legislative acorn,"[2] as courts increasingly have accorded leniency to such actions. For nearly three decades, lower federal courts have allowed private parties to bring suit based on aiding and abetting claims pursuant to section 10(b).[3] Over the same time period, the Supreme Court has declined to rule on the validity of such claims.[4] The Supreme Court's silence recently ended with its decision in *Central Bank v. First Interstate Bank*,[5] which held that the conduct proscribed by section 10(b) does not include aiding and abetting a primary violator. Consequently, private parties no longer can maintain suits based on such claims.[6] The Court's holding overrules substantial lower court precedent allowing aiding and abetting liability. This decision illustrates the continuation of the Contraction Era of securities law[7] and for the time being halts causes of action based on aiding and abetting section 10(b) violations.[8]

---

1. Virginia Bankshares, Inc. v. Sandberg, 501 U.S. 1083, 1115 (1991) (Kennedy, J., concurring in part and dissenting in part).

2. Blue Chip Stamps v. Manor Drug Stores, 421 U.S. 723, 737 (1975).

3. *See, e.g.*, Monsen v. Consolidated Dressed Beef Co., 579 F.2d 793 (3d Cir.), *cert. denied*, 439 U.S. 930 (1978); Brennan v. Midwestern United Life Ins. Co., 259 F. Supp. 673 (N.D. Ind. 1966).

4. *See* Herman & MacLean v. Huddleston, 459 U.S. 375, 379 n.5 (1983); Ernst & Ernst v. Hochfelder, 425 U.S. 185, 193 n.7 (1976).

5. 114 S. Ct. 1439 (1994).

6. *Id.*

7. The term "Contraction Era" refers to the post-1975 decisions in which courts denied, restricted, and criticized implied private causes of action under the securities laws. *See* Alan R. Bromberg & Lewis D. Lowenfels, *Aiding and Abetting Securities Fraud: A Critical Examination*, 52 ALB. L. REV. 637, 648 n.64 (1988).

8. Senator Christopher Dodd (D-Conn.) and Senator Howard Metzenbaum (D-Ohio) both announced their disapproval of the Court's decision. Senator Dodd declared the Court "has laid down the gauntlet for Congress." *SEC Advocates Legislation to Preserve Section 10(b) Aiding and Abetting Liability*, 26 SEC. REG. & L. REP. (BNA) 691, 691 (May 13, 1994). Senator Metzenbaum, deeming the Court's reasoning "bizarre," introduced the Securities Exchange Act of 1934 Amendment Act of 1994 in response. 140 CONG. REC. S9460 (daily ed. July 21, 1994).

This Comment analyzes the Court's decision in *Central Bank*. Part I examines the historical development of aiding and abetting liability and the judicially conservative trend of recent Supreme Court decisions in the securities law arena. Part II provides the factual background, procedural history, and majority and dissenting rationales of *Central Bank*. Part III analyzes the Court's decision. The analysis praises the majority's judicial restraint, criticizes the dissent's reliance on suspect methods of statutory construction, and discusses possible implications of the decision. The larger policy issue of whether aiding and abetting liability might serve as a valuable addition to section 10(b) is beyond the scope of this Comment. Though aiding and abetting liability may be desirable to some commentators, this Comment takes the position that section 10(b), as it currently reads, cannot and should not be construed to allow such liability.

I. BACKGROUND

In the aftermath of the stock market crash of 1929 and the subsequent economic depression, Congress enacted the Securities Act of 1933 ("1933 Act")[9] and the Securities Exchange Act of 1934 ("1934 Act").[10] Congress passed the Acts to protect investors from fraudulent conduct in connection with securities transactions,[11] thereby substituting a philosophy of full disclosure for the prevailing philosophy of caveat emptor.[12] Together, the Acts create an extensive scheme of liability, with the 1933 Act regulating initial distribution of securities and the 1934 Act regulating post-distribution trading.[13]

The use of aiding and abetting liability by private parties has proliferated in recent years, becoming the most important secondary liability doctrine used in section 10(b) actions.[14] Prior to *Central Bank*, many commentators felt the doctrine had become so established that the Supreme Court would never reject it.[15] Other commentators and courts questioned the validity of private aiding and abetting claims given the Supreme Court's restrictive textual approach to securities law in recent years.[16]

---

9. Securities Act of 1933, 15 U.S.C. §§ 77a-77aa (1988 & Supp. IV 1992).

10. Securities Exchange Act of 1934, 15 U.S.C. §§ 78a-78ll (1988 & Supp. IV 1992).

11. Cheryl L. Pollak, Comment, *Rule 10b-5 Liability After* Hochfelder*: Abandoning the Concept of Aiding and Abetting*, 45 U. CHI. L. REV. 218, 218 (1977).

12. Affiliated Ute Citizens v. United States, 406 U.S. 128, 151 (1972) (citing SEC v. Capital Gains Research Bureau, 375 U.S. 180, 186 (1963)).

13. Central Bank v. First Interstate Bank, 114 S. Ct. 1439, 1445 (1994).

14. Thomas L. Riesenberg, *Supreme Court to Examine Aiding and Abetting Liability Under Rule 10b-5*, 7 INSIGHTS, no. 8, August 1993, at 34. In the past, aiding and abetting liability was used almost exclusively by the SEC. More recently, almost every private § 10(b) action contains an aiding and abetting claim. *Id.*

15. *See, e.g.*, William H. Kuehnle, *Secondary Liability Under the Federal Securities Laws—Aiding and Abetting, Conspiracy, Controlling Person, and Agency: Common-Law Principles and the Statutory Scheme*, 14 J. CORP. L. 313, 315-18 (1988).

16. *See* Bromberg & Lowenfels, *supra* note 7, *passim*.

A. *Primary Violations and Aiding and Abetting under Section 10(b)*

Section 10(b) prohibits the use of deception or manipulation, as specified by the Securities and Exchange Commission ("SEC"), in connection with the purchase or sale of any security.[17] Pursuant to the authority granted in section 10(b), the SEC promulgated Rule 10b-5, which requires traders to either disclose inside information or abstain from trading based on such information.[18]

While the text of section 10(b) does not provide for a private right of action, courts have inferred such a right based on the maxim *ubi jus ibi remedium*, which is to say, where there is a right there is a remedy.[19] In a suit based on a section 10(b) violation, a purchaser or seller[20] of any security can bring a claim against any person engaged in manipulation[21] or deception in connection with the purchase or sale of the security.[22] For the claim to succeed, the plaintiff must satisfy three judicially created thresholds. First, the defendant must have engaged intentionally or knowingly in manipulative or deceptive conduct[23] on which the plaintiff re-

---

17. Securities Exchange Act § 10, 15 U.S.C. § 78j (1988). Section 10 provides in part:

> It shall be unlawful for any person, directly or indirectly, by the use of any means or instrumentality of interstate commerce or of the mails, or of any facility of any national securities exchange-
>
> . . . .
>
> (b) To use or employ, in connection the with purchase or sale of any security registered on a national securities exchange or any security not so registered, any manipulative or deceptive device or contrivance in contravention of such rules and regulations as the [SEC] may prescribe as necessary or appropriate in the public interest or for the protection of investors.

Id. Section 10 is a catchall provision, "but what it catches must be fraud." Chiarella v. United States, 445 U.S. 222, 234-35 (1980).

18. 17 C.F.R. § 240.10b-5 (1994). The Rule provides:

> It shall be unlawful for any person, directly or indirectly, by the use of any means or instrumentality of interstate commerce, or of the mails or of any facility of any national securities exchange,
>
> (a) To employ any device, scheme, or artifice to defraud,
>
> (b) To make any untrue statement of a material fact or to omit to state a material fact necessary in order to make the statements made, in the light of the circumstances under which they were made, not misleading, or
>
> (c) To engage in any act, practice, or course of business which operates or would operate as a fraud or deceit upon any person, in connection with the purchase or sale of any security.

Throughout this Comment, the author refers to § 10(b); however, any construction limiting conduct actionable under § 10(b) also limits Rule 10b-5 because the Rule gains its authority from § 10(b). Rulemaking authority granted to an agency gives it the power to adopt regulations to carry into effect the will of Congress, not the power to make law. Ernst & Ernst v. Hochfelder, 425 U.S. 185, 213-14 (1976).

19. Kardon v. National Gypsum Co., 69 F. Supp. 512, 513 (E.D. Pa. 1946).

20. *See, e.g.*, Blue Chip Stamps v. Manor Drug Stores, 421 U.S. 723, 731-32 (1975) (illustrating the history of lower federal court holdings allowing only actual purchasers and sellers to sue under § 10(b)).

21. In the securities law context, "manipulation" is a term of art referring to practices such as wash sales, matched orders, or rigged prices, that artificially affect market activity. Santa Fe Indus. v. Green, 430 U.S. 462, 476 (1977).

22. Herman & MacLean v. Huddleston, 459 U.S. 375, 382 (1983).

23. *See* Ernst & Ernst v. Hochfelder, 425 U.S. 185, 197 (1976). The Court also rejected negligence as a basis for liability. *Id.* at 199; *see also* Chiarella v. United States, 455 U.S. 222, 228 (1980) (holding omissions to be actionable only if defendant had a duty to disclose); Santa Fe Indus. v. Green, 430 U.S. 462 (1977) (limiting the conduct actionable to misstatements or omissions of material facts and manipulation of securities pricing).

lied.[24] Second, the plaintiff must establish that the defendant's conduct proximately caused an injury.[25] Finally, the defendant must have used the United States mail, a national security exchange, or interstate commerce to further the manipulation or deception.[26]

Additional judicial construction of the section eventually expanded liability and allowed for private actions based on a defendant aiding and abetting the primary violator.[27] This construction of the section resulted in defrauded investors suing banks,[28] accountants,[29] lawyers,[30] underwriters,[31] and stock exchanges[32] due to the typical insolvency of the primary violators in the wake of failed securities schemes.[33]

In contrast to the judicially implied private right of action for section 10(b) violations, several sections of the 1933 Act and 1934 Act expressly provide for private remedial measures.[34] However, none of these sections provide for aiding and abetting liability. While none of the sections provide an express private remedy for aiding and abetting, other sections of both the 1933 Act and the 1934 Act provide for secondary liability in the form of "controlling person" liability.[35] The 1934 Act allows the SEC to censure or restrict the activities of persons associated with a broker-

---

24. Basic Inc. v. Levinson, 485 U.S. 224, 243 (1988). Along with finding reliance a requirement, the Court also clarified the definition of materiality in regard to reliance as the "substantial likelihood that the disclosure or the omitted fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information made available." *Id.* at 231 (citing TSC Indus. v. Northway, 426 U.S. 438, 449 (1976)); *see also* Affiliated Ute Citizens v. United States, 406 U.S. 128, 153 (1972) (holding plaintiff need not show reliance when basing action on an omission).

25. *See* Cooke v. Manufactured Homes, Inc., 998 F.2d 1256, 1261 (4th Cir. 1993); *In re* Control Data Corp. Securities Litigation, 983 F.2d 616, 618 (8th Cir. 1991); Harris v. Union Elec. Co., 787 F.2d 355, 362 (8th Cir.), *cert. denied*, 479 U.S. 823 (1986).

26. Perez-Rubio v. Wyckoff, 718 F. Supp. 217, 232 (S.D.N.Y. 1989).

27. *See* Brennan v. Midwestern United Life Ins. Co., 259 F. Supp. 673, 676 (N.D. Ind. 1966).

28. *See, e.g.*, Monsen v. Consolidated Dressed Beef Co., 579 F.2d 793 (3d Cir.), *cert. denied*, 439 U.S. 930 (1978).

29. *See, e.g.*, H.L. Green Co. v. Childree, 185 F. Supp. 95 (S.D.N.Y. 1960).

30. *See, e.g.*, SEC v. Coven, 581 F.2d 1020 (2d Cir. 1978), *cert. denied*, 440 U.S. 950 (1979).

31. *See, e.g.*, IIT v. Cornfeld, 462 F. Supp. 209 (S.D.N.Y. 1978), *aff'd in part and rev'd in part*, 619 F.2d 909 (2d Cir. 1980).

32. *See, e.g.*, Pettit v. American Stock Exch., 217 F. Supp. 21 (S.D.N.Y. 1963).

33. *See* John T. Vangel, Note, *A Complicity-Doctrine Approach to Section 10(b) Aiding and Abetting Civil Damages Actions*, 89 COLUM. L. REV. 180, 180 (1989).

34. *See* 15 U.S.C. § 77k(a) (1988) (creating private cause of action for false statements or omissions in registration statements); 15 U.S.C. § 77*l* (1988) (same for the sale of securities by way of a material misstatement or omission); 15 U.S.C. § 78i(e) (1988) (same for manipulative practices, e.g., wash sales and matched orders); 15 U.S.C. § 78p(b) (1988) (same against owners, officers, and directors who engage in short-swing trading); 15 U.S.C. § 78r(a) (1988) (same for misleading statements in forms filed with the SEC). In contrast to the express rights of action granted in the 1933 and 1934 Acts, the Commodity Exchange Act expressly provides for a private cause of action based on aiding and abetting a violator. 7 U.S.C. § 25(b)(3) (1988).

35. *See, e.g.*, Securities Act of 1933 § 15, 15 U.S.C. § 77o (1988) (providing for "controlling person" liability, unless such person "had no knowledge of or reasonable ground to believe in the existence of the facts by reason of which the liability of the controlled person is alleged to exist"); Securities Exchange Act of 1934 § 20, 15 U.S.C. § 78t (1988) (providing for controlling person liability with the same 'good faith' exception).

dealer[36] who willfully aid and abet or otherwise induce a securities violation by another.[37] In addition to the foregoing express provisions for secondary liability, the 1934 Act imposes supervisory duties on exchanges and clearing agencies.[38]

The first case to address aiding and abetting liability in the securities arena, *SEC v. Timetrust*,[39] illustrates its criminal law antecedents. The SEC alleged that Timetrust had engaged in a scheme to defraud purchasers of securities[40] and that Timetrust had done so with the "aid and abetment" of others.[41] In deciding whether the SEC could invoke properly an injunction based on aiding and abetting a securities violation of section 17(a), the court noted that the Criminal Code of the United States provided for aiding and abetting liability in a criminal proceeding.[42] The court then concluded that "no good reason appears why this same rule should not apply in an injunctive proceeding to restrain a violation of the same statute."[43]

The landmark decision allowing civil liability for aiding and abetting a violation of section 10(b) came in *Brennan v. Midwestern United Life Insurance*.[44] The district court relied on tort law principles to allow a private action based on an aiding and abetting theory. In *Brennan*, Dobich Securities Corporation, from whom the plaintiffs had purchased defendant corporation's stock, used a short selling scheme[45] to create an artificially high market price for the stock.[46] The plaintiffs alleged that Midwestern knew of Dobich's conduct but permitted it to continue by failing to report Dobich to the SEC.[47] The plaintiffs argued that by failing to report the improper activity, the defendants "knowingly and purposely encouraged an artificial build-up in the market for its stock."[48] As a result of such aiding and abetting, Midwestern "was allegedly in a more favorable position for potential mergers" being negotiated.[49]

---

36. "Broker-dealer" is a term of art referring to a securities brokerage firm. BLACK'S LAW DICTIONARY 193 (6th ed. 1990).

37. 15 U.S.C. § 78o(b)(4)(E) (1988). Pursuant to authority granted by the Investment Advisers Act and the Investment Company Act, the SEC also has authority to discipline investment advisers who aid or abet a securities law violation. 15 U.S.C. §§ 80a-9(b)(3), 80b-3(e)(5) (Supp. IV 1992).

38. 15 U.S.C. §§ 78f, 78q (1988 & Supp. IV 1992).

39. 28 F. Supp. 34 (N.D. Cal. 1939).

40. *Id.* at 37.

41. *Id.* at 43.

42. *Id.*

43. *Id.*

44. 259 F. Supp. 673 (N.D. Ind. 1966).

45. Short selling occurs when a seller contracts for the sale of stock the seller does not own, "so as to be available for delivery at the time when, under rules of the exchange, delivery must be made." BLACK'S LAW DICTIONARY 1379 (6th Ed. 1990). Short selling is not illegal in and of itself, but in *Brennan* Dobich did not cover (buy and deliver) as the rules required. *Brennan*, 259 F. Supp. at 675.

46. *Brennan*, 259 F. Supp. at 675.

47. *Id.*

48. *Id.*

49. *Id.*

Midwestern attacked the complaint by challenging the aiding and abetting theory, contending section 10(b) did not provide authority for the claim and that the legislative history indicated that Congress chose not to proscribe aiding and abetting.[50] The district court rejected Midwestern's contentions, reasoning courts had allowed private section 10(b) actions based on general principles of tort law and that these same principles should guide the construction of section 10(b) with respect to the issue of aiding and abetting liability.[51] The court relied on the Restatement of Torts section 876, stating that principles formulated therein are the "logical and natural complement" to a section 10(b) implied right of action.[52]

Other courts ultimately followed *Brennan*'s rationale and similarly relied on tort law to hold peripheral defendants liable as aiders and abettors.[53]

All federal courts of appeals considering whether aiding and abetting liability exists under section 10(b) have concluded that it does.[54] The majority rule, as represented by the Second, Third, Eighth, and Tenth Circuits, requires: (1) an independent violation of section 10(b); (2) the aider and abettor's knowledge of the violation; and (3) substantial assistance by the aider and abettor in the achievement of the primary violation.[55] The First, Fifth, Sixth, and Eleventh Circuits articulate the elements in a slightly different manner, requiring that the aider and abettor knowingly and substantially assist in the primary violation.[56]

---

50. *Id.* at 675-77.

51. *Id.* at 680-82 (citing Crist v. United Underwriters, Ltd., 343 F.2d 902 (10th Cir. 1965); Kardon v. National Gypsum Co., 69 F. Supp. 512 (E.D. Pa. 1946)).

52. *Brennan*, 259 F. Supp. at 680. The Restatement (Second) of Torts, reading substantially the same today as the Restatement of Torts did in 1939, provides:

> For harm resulting to a third person from the tortious conduct of another, one is subject to liability if he
>
> (a) does a tortious act in concert with the other or pursuant to a common design with him, or
>
> (b) knows that the other's conduct constitutes a breach of duty and gives substantial assistance or encouragement to the other so to conduct himself, or
>
> (c) gives substantial assistance to the other in accomplishing a tortious result and his own conduct, separately considered, constitutes a breach of duty to the third person.

RESTATEMENT (SECOND) OF TORTS § 876 (1977).

53. *See, e.g.*, SEC v. Coffey, 493 F.2d 1304, 1316 (6th Cir. 1974), *cert. denied*, 420 U.S. 908 (1975); Landy v. FDIC, 486 F.2d 139, 162 (3d Cir. 1973), *cert. denied*, 416 U.S. 960 (1974); *see, e.g.*, Daniel R. Fischel, *Secondary Liability Under Section 10(b) of the Securities Act of 1934*, 69 CAL. L. REV. 80, 84 n.29 (1981).

54. Central Bank v. First Interstate Bank, 114 S. Ct. 1439, 1456 & n.1 (Stevens, J., dissenting).

55. *See Landy*, 486 F.2d at 162-63; *see also* National Union Fire Ins. Co. v. Turtur, 892 F.2d 199, 206-07 (2d Cir. 1989); Walck v. American Stock Exch., Inc., 687 F.2d 778, 791 (3rd Cir. 1982), *cert. denied*, 461 U.S. 942 (1983); Stokes v. Lokken, 644 F.2d 779, 782-83 (8th Cir. 1981); Edward J. Mawod & Co. v. SEC, 591 F.2d 588, 595-96 (10th Cir. 1979). The Ninth Circuit also uses this formula but additionally requires actual knowledge of the violation by the aider and abettor. *See* Harmsen v. Smith, 693 F.2d 932, 943 (9th Cir. 1982), *cert. denied*, 464 U.S. 822 (1983).

56. *See, e.g.*, Bane v. Sigmundr Exploration Corp., 848 F.2d 579, 581 (5th Cir. 1988); Moore v. Fenex, Inc., 809 F.2d 297, 303 (6th Cir.), *cert. denied*, 483 U.S. 1006 (1987); Woods v. Barnett Bank, 765 F.2d 1004, 1009 (11th Cir. 1985). Some commentators view the differ-

The Seventh Circuit employs a stricter analysis than the other circuits. Aside from requiring satisfaction of the three-pronged majority rule, the Seventh Circuit requires as a threshold matter that the aider and abettor possess the same level of scienter as the primary violator and that the aider and abettor commit one of the manipulative or deceptive acts.[57] In *Barker v. Henderson, Franklin, Starnes & Holt*,[58] the plaintiffs brought a claim against an accounting firm and a law firm for aiding and abetting. The Seventh Circuit granted summary judgment, stating that an aider and abettor must meet the same standards as a primary violator but need not actually sell the security.[59]

B. *Aiding and Abetting Application Problems*

No clear distinctions exist between primary and secondary liability in the securities law context. Generally, plaintiffs sue the same person as both a primary violator and as an aider and abettor,[60] and courts spend minimal time analyzing the distinction.[61] At times, courts impose liability based on both capacities.[62]

Some courts distinguish between primary liability and aiding and abetting liability by using a direct/indirect participant duality.[63] Under this type of analysis, a primary violator directly participates in the fraud, while a secondary violator only indirectly participates in the fraud. *SEC v. Coffey*[64] illustrates this distinction. In *Coffey*, the State of Ohio purchased notes from an issuing company.[65] The SEC alleged the issuing company made misrepresentations to the state and to a rating agency to achieve a "prime" rating for the notes.[66] The SEC named the company's financial vice-president as one of the defendants, and, in deciding whether to impose primary liability on the officer, the court examined whether the financial vice-president had direct contacts with the misled parties.[67]

Courts occasionally base the distinction between a primary violator and an aider and abettor on the role the person played in the transaction. In *DMI Furniture, Inc. v. Brown, Kraft & Co.*,[68] the court limited primary

---

ences between the Second, Third, Eighth, and Tenth Circuits' test and the test applied by the First, Fifth, Sixth, and Eleventh Circuits as inconsequential. *See, e.g.*, Joel S. Feldman, *The Breakdown of Securities Fraud Aiding and Abetting Liability: Can a Uniform Standard be Resurrected?*, 19 SEC. REG. L.J. 45, 73 (1991).

57. Schlifke v. Seafirst Corp., 866 F.2d 935, 947 (7th Cir. 1989); *see* LHLC Corp. v. Cluett, Peabody & Co., 842 F.2d 928, 932 (7th Cir.), *cert. denied*, 488 U.S. 926 (1988); Barker v. Henderson, Franklin, Starnes & Holt, 797 F.2d 490, 495 (7th Cir. 1986).

58. 797 F.2d 490 (7th Cir. 1986).

59. *Id.* at 495.

60. Bromberg & Lowenfels, *supra* note 7, at 640.

61. Kuehnle, *supra* note 15, at 318.

62. *See, e.g.*, Herman & Maclean v. Huddleston, 459 U.S. 375, 379 & n.5 (1983) (recognizing, but not discussing, the dual liability found by the trial court).

63. *See* Smith v. Ayres, 845 F.2d 1360, 1365 (5th Cir. 1988).

64. 493 F.2d 1304 (6th Cir. 1974), *cert. denied*, 420 U.S. 908 (1975).

65. *Id.* at 1308.

66. *Id.*

67. *Id.* at 1315.

68. 644 F. Supp. 1517 (C.D. Cal. 1986).

liability to actual buyers and sellers of securities, persons acting in roles that statutes expressly specify as liable, or persons acting in roles that constitute an integral part of the statutory scheme.[69]

The lines drawn in *DMI* and *Coffey* do not offer valid distinctions because the text of section 10(b) covers "any person."[70] If "any person" can "directly or indirectly" commit a primary violation, the distinctions fail to separate sufficiently aiders and abettors from primary violators.

The Seventh Circuit limits primary liability to those persons otherwise covered by the 1933 and 1934 Acts.[71] Under this view, the court limits primary liability to buyers, sellers, issuers of securities, the board of directors of the issuer, persons signing or preparing a prospectus, and any controlling persons.[72] As noted previously, the Seventh Circuit also requires that the aider and abettor engage in a manipulative or deceptive act. Arguably, under the Seventh Circuit's approach, aiding and abetting liability constitutes a mere label attached to accountants, lawyers, trustees, and such whose actions constitute primary violations of section 10(b).[73]

While some commentators and courts have suggested that no distinction exists between primary and aiding and abetting liability, others have defined secondary liability in terms of duty.[74] Still others beg the question, describing a primary violator as one who commits acts directly proscribed by law and a secondary violator as one upon whom the court imposes liability because of a relationship with the primary violator.[75] In any event, the jurisprudence prior to *Central Bank* provided no clear or meaningful distinction between primary and aiding and abetting liability.[76]

C. *Supreme Court Decisions in the Contraction Era*

Courts developed private remedies and aiding and abetting liability in the Expansion Era of securities law.[77] During this era, courts increasingly

---

69. *Id.* at 1519.

70. 15 U.S.C. § 78j (1988).

71. *See* Barker v. Henderson, Franklin, Starnes & Holt, 797 F.2d 490, 494-95 (7th Cir. 1986).

72. *See* Mary T. Doherty, Note, *Aiding and Abetting Securities Fraud*, 25 IND. L. REV. 829, 839-40 (1992) (arguing the Seventh Circuit's distinction is different from other circuits in that it focuses on the defendant's role in the statutory scheme rather than his role in the fraudulent scheme).

73. *See id.* at 852 (suggesting also that aiding and abetting has been all but eliminated in the Seventh Circuit).

74. *See, e.g.*, David S. Ruder, *Multiple Defendants in Securities Law Fraud Cases: Aiding and Abetting, Conspiracy, In Pari Delicto, Indemnification, and Contribution*, 120 U. PA. L. REV. 597, 600 (1972) (classifying persons owing a duty to the public as primary violators).

75. *See* Kuehnle, *supra* note 15, at 318-20.

76. *See, e.g.*, Feldman, *supra* note 56, at 46 (contending that the distinction defies precise definition).

77. Bromberg & Lowenfels, *supra* note 7, at 648; *see also* Lewis D. Lowenfels, *Recent Supreme Court Decisions Under the Federal Securities Laws: The Pendulum Swings*, 65 GEO. L.J. 891, 923 (1977) (arguing the Court has shifted its views on federal securities law from an expansion of implied rights during the Expansion Era to a more recent defendant-oriented construction).

formulated new private actions and remedies.[78] Beginning in 1975, the Supreme Court entered a new phase in which its decisions began to limit the doctrines developed during the Expansion Era.[79] The Court expressed concern over the implications of the expansive liability created in the Expansion Era and began to restrict such liability.[80] The ensuing "contraction" period consequently resulted in courts rejecting new private actions and criticizing the existing private actions.[81]

The Supreme Court's decisions during the Contraction Era demonstrate a recurring trend to look continually to the language of the statute and the statutory scheme to ascertain congressional intent.[82] This textual approach to statutory construction differs dramatically from the original approach used by the courts in recognizing private actions and aiding and abetting claims. The difference in the two approaches caused courts and commentators to question the validity of private actions based on aiding and abetting.[83]

The Court's securities law decisions in the Contraction Era can be grouped roughly into three interrelated categories. The first category consists of cases in which the Court considered whether a particular section implied a private cause of action.[84] Cases in the second category dis-

78. *See, e.g.*, SEC v. National Sec., Inc., 393 U.S. 453, 468-69 (1969) (holding that a merger involves both a purchase and sale and recognizing Rule 10b-5 application in a proxy solicitation situation); J.I. Case Co. v. Borak, 377 U.S. 426, 431-34 (1964) (finding an implied cause of action for violation of Rule 14a-9 and instructing courts to provide effective remedial measures); SEC v. Capital Gains Research Bureau, Inc., 375 U.S. 180, 192 (1963) (holding scienter unnecessary in an SEC injunction proceeding under the Investment Advisors Act of 1940); Marx v. Computer Sciences Corp., 507 F.2d 485, 492 (9th Cir. 1974) (ruling that an earnings forecast constituted a fact and allowing a 10b-5 claim based on such a forecast); White v. Abrams, 495 F.2d 724, 733 (9th Cir. 1974) (finding 10b-5 broad enough to include negligence); A.T. Brod & Co. v. Perlow, 375 F.2d 393, 397 (2d Cir. 1967) (allowing 10b-5 claims for all forms of fraud, not just traditional ones); Vine v. Beneficial Fin. Co., 374 F.2d 627 (2d Cir. 1967) (easing the buyer/seller requirement of 10b-5); Miller v. Bargain City, USA, Inc., 229 F. Supp. 33 (E.D. Pa. 1964) (allowing an open market buyer to maintain a 10b-5 suit based on false statements filed with the SEC); New Park Mining Co. v. Cranmer, 225 F. Supp. 261, 266 (S.D.N.Y. 1963) (allowing a corporation to sue its officers and directors for 10b-5 violations and relaxing privity); Cochran v. Channing Corp., 211 F. Supp. 239, 243-45 (S.D.N.Y. 1962) (recognizing a 10b-5 violation by silence and dispensing with privity requirement).

79. Bromberg & Lowenfels, *supra* note 7, at 648 n.64.

80. *Id.*

81. *Id.*

82. *See* Musick, Peeler & Garrett v. Employers Ins. of Wausau, 113 S. Ct. 2085, 2090 (1993); Virginia Bankshares, Inc. v. Sandberg, 501 U.S. 1083, 1102-04 (1991); Chiarella v. United States, 445 U.S. 222, 234-35 (1980); Blue Chip Stamps v. Manor Drug Store, 421 U.S. 723, 756 (1975).

83. *See* Akin v. Q-L Inv., Inc., 959 F.2d 521, 525 (5th Cir. 1992); Benoay v. Decker, 517 F. Supp. 490, 495 (E.D. Mich. 1981), *aff'd*, 735 F.2d 1363 (6th Cir. 1984); *see also* Fischel, *supra* note 53, at 91-96.

84. *See generally* Tamar Frankel, *Implied Rights of Action*, 78 VA. L. REV. 553 (1981) (providing a detailed analysis of the Court's decisions involving implied private remedies for securities violations).

cuss issues concerning the extent of the prohibited conduct.[85] Cases in the third category address aspects of the private liability scheme.[86]

During the Expansion Era, the Court followed an enforcement rationale approach in deciding whether a particular section implied a private remedy. Under this rationale, the Court allowed private actions based on the necessity of supplementing SEC enforcement of the statutory scheme.[87] In *Cort v. Ash*,[88] the Court restricted this enforcement rationale and articulated a four part test for deciding whether an implied private remedy existed.[89] Subsequently, the Court further narrowed the test and articulated the dispositive inquiry as whether Congress intended to create a private remedy.[90] This line of cases demonstrates a withdrawal from the earlier tort law rationale of implied private remedies and highlights the Court's focus on statutory construction.[91]

Ernst & Ernst v. Hochfelder[92] illustrates the second category of cases, in which the Court considered the scope of the conduct on which a plaintiff could base a cause of action. In *Hochfelder*, the plaintiffs premised their cause of action on a theory of "negligent nonfeasance."[93] While the plaintiff's complaint alleged the defendant had aided and abetted a violator of section 10(b), the Court declined to address this issue.[94] The Court, relying on the language of the statute, ruled that a section 10(b) claim requires scienter.[95] Finding the language clear in the overall statutory

85. *See, e.g.*, Dirks v. SEC, 463 U.S. 646, 657-58 (1983) (holding that a duty to disclose does not arise from the mere possession of non-public information); Aaron v. SEC, 446 U.S. 680, 691 (1980) (holding that the scienter requirement is the same for SEC injunctive proceedings and private party claims); *Chiarella*, 445 U.S. at 228 (holding that nondisclosure violates § 10(b) only when the defendant owes a disclosure duty to the plaintiff); Santa Fe Indus. v. Green, 430 U.S. 462, 478-80 (1977) (holding that regardless of the fairness of the terms, Delaware short-form mergers do not violate § 10(b) where there is full disclosure); Ernst & Ernst v. Hochfelder, 425 U.S. 185, 201 (1976) (requiring intent to deceive).

86. *See, e.g.*, Musick, Peeler & Garrett v. Employers Ins. of Wausau, 113 S. Ct. 2085, 2091-92 (1993) (recognizing a right of contribution for § 10(b) violations); Basic Inc. v. Levinson, 485 U.S. 224, 243 (1988) (imposing a reliance requirement on the plaintiff); Bateman Eichler, Hill Richards, Inc. v. Berner, 472 U.S. 299, 315 (1985) (denying the use of the *in pari delicto* defense to liability); *Blue Chip Stamps*, 421 U.S. at 733 (limiting standing to buyers and sellers).

87. *See, e.g.*, J.I. Case Co. v. Borak, 377 U.S. 426, 430-31 (1964) (implying a private cause of action under § 14(a) in order to supplement enforcement of the statutory scheme).

88. 422 U.S. 66 (1975).

89. The four-part test established in *Cort* for implying a private remedy asks: (1) Is the plaintiff a member of the beneficiary class of the statute?; (2) Is there any explicit or implicit indication that Congress intend a private remedy?; (3) Is a private remedy consistent with the legislative scheme?; and (4) Is the cause of action one of basically state concern typically relegated to state law? *Id.* at 78.

90. "The question whether a statute creates a cause of action, either expressly or by implication, is basically a matter of statutory construction." Transamerica Mortgage Advisors v. Lewis, 444 U.S. 11, 15 (1979).

91. *See* Bromberg & Lowenfels, *supra* note 7, at 650-661 (analyzing the validity of aiding and abetting liability under the *Cort* four-part test); Fischel, *supra* note 53, at 90-94 (discussing the validity of § 10(b) secondary liability under the Court's Contraction Era view); Frankel, *supra* note 84, *passim* (discussing the evolution of the Court's implied remedy analysis).

92. 425 U.S. 185 (1976).

93. *Id.* at 190.

94. *Id.* at 193 n.7.

95. *Id.* at 201.

scheme, the Court held that the text of the section controlled the decision.[96]

Musick, Peeler & Garrett v. Employers Insurance of Wausau[97] illustrates the third category of cases, in which the Court faced issues concerning the procedural aspects of the private liability scheme. In *Musick*, the Court considered whether defendants in a private section 10(b) action could seek contribution from other defendants.[98] The Court held in the affirmative, basing its decision on the overall statutory scheme.[99] The Court stated that when faced with questions as to the individual components of an implied private action, it must infer how the 1934 Congress would have addressed the issue had the action been an express provision.[100] In making its inference, the Court announced its continuing goal to avoid "conflict with Congress' own express rights of action, to promote clarity, consistency and coherence for those who rely upon or are subject to 10b-5 liability, and to effect Congress' objectives in enacting the securities laws."[101] *Musick* demonstrates the Court's consistent use of the textual approach, even when it expanded the components of a private right of action.

II. INSTANT CASE

A. *Facts and Procedural History*

Colorado Springs-Stetson Hills Public Building Authority (the "Authority") issued bonds in 1986 and 1988 with a total face value of $26 million.[102] Central Bank, the petitioner, acted as an indenture trustee for the bond issue.[103] Landowner assessment liens secured the bonds.[104] The bonds contained covenants requiring the value of the land subject to

---

96. *Id.*

97. 113 S. Ct. 2085 (1993).

98. *Id.* at 2086.

99. *Id.* at 2090-91. The Court looked first to the language of § 10(b), but noted it "provides little guidance." *Id.* at 2090. The Court next turned to sections 9 and 18 of the 1934 Act, stating that the sections were of "particular significance in determining how Congress would have resolved the question" as the sections target the same dangers as § 10(b). *Id.* at 2090. The Court found sections 9 and 18 both to contain nearly identical rights of contribution, and thus inferred a right of contribution for § 10(b) defendants. *Id.* at 2091. In dissent, Justice Thomas asserted that the majority's decision "unfortunately nourishes 'a judicial oak which has grown from little more than a legislative acorn.' " *Id.* at 2092 (Thomas, J., dissenting) (quoting Blue Chip Stamps v. Manor Drug Stores, 421 U.S. 723, 737 (1975)).

100. "Our task is not to assess the relative merits of the competing rules, but rather to attempt to infer how the 1934 Congress would have addressed the issue. . . ." *Id.* at 2089-90.

101. *Id.* at 2090 (citation omitted); *see also Blue Chip Stamps*, 421 U.S. at 737-44 (promoting clarity); *Santa Fe Indus. v. Green*, 430 U.S. 462, 477-78 (effecting Congress's objectives).

102. Central Bank v. First Interstate Bank, 114 S. Ct. 1439, 1443 (1994).

103. *Id.* An indenture, in the context of business financing, refers to a contract that establishes the terms of the bond issuance. BLACK'S LAW DICTIONARY 770 (6th ed. 1990). An indenture trustee performs duties specifically set out in the indenture and is only liable for the performance of those duties. The indenture trustee thus has the duty to examine all evidence to ensure it conforms to the requirements of the indenture. 15 U.S.C. 77ooo(a) (1988 & Supp. IV 1992).

104. *Central Bank*, 114 S. Ct. at 1443.

the liens to be worth at least 160% of the total outstanding principal and interest.[105]

The covenants additionally required the developer of Stetson Hills, AmWest Development, to provide evidence that the 160% threshold was met each year.[106] The 1988 appraisal, provided in January, contained land values virtually unchanged from the 1986 appraisals.[107] In contrast, property values in the Colorado Springs area generally had declined over the same period of time.[108] Because of the conflict between the appraisal and the general Colorado Springs real-estate market, Central Bank had its in-house appraiser review the 1988 evaluation.[109] The in-house appraiser concluded the 1988 evaluation appeared overly optimistic and recommended hiring an outside appraiser for an independent review of the appraisal.[110]

After a series of letters, Central Bank agreed to postpone the independent evaluation until December 1988.[111] By the time the 1988 bond issue closed in June, First Interstate Bank had bought $2.1 million worth of the 1988 bonds.[112] The independent review began in December, but before completion of the review, the Authority defaulted on the 1988 bonds.[113]

First Interstate sued the Authority, an underwriter, a junior underwriter, and a director of AmWest, alleging that the parties had violated section 10(b).[114] First Interstate also sued Central Bank, alleging Central Bank aided and abetted the other defendants in violating section 10(b).[115]

The district court granted Central Bank's motion for summary judgment on the ground no genuine issue of material fact existed, holding that allegations of recklessness did not satisfy the scienter requirement for aiding and abetting.[116] On appeal, the Tenth Circuit held that recklessness did satisfy the scienter element for an aiding and abetting claim.[117] The court of appeals did not consider the existence or validity of aiding and abetting liability under section 10(b).[118]

Central Bank filed a petition for writ of certiorari, which the Supreme Court granted.[119] The original petition raised two issues: (1) whether an indenture trustee can be liable for aiding and abetting if no duties have

105. *Id.*
106. *Id.*
107. *Id.*
108. *Id.*
109. *Id.*
110. *Id.*
111. *Id.*
112. *Id.*
113. *Id.*
114. *Id.*
115. *Id.*
116. *See* First Interstate Bank v. Pring, 969 F.2d 891, 900 (10th Cir. 1992), *rev'd sub nom.* Central Bank v. First Interstate Bank, 114 S. Ct. 1439 (1994).
117. *Pring*, 969 F.2d at 903.
118. *Id.* at 898-904.
119. Central Bank v. First Interstate Bank, 113 S. Ct. 2927 (1993).

been breached; and (2) whether recklessness satisfies the scienter requirement.[120] The Court, sua sponte, directed the parties to brief the issue of whether a private party could bring suit based on a claim of aiding and abetting.[121]

B. *Majority Opinion*

In the majority opinion written by Justice Kennedy, the Court held that section 10(b) proscribes conduct involving manipulation or deception but does not proscribe aiding and abetting a primary violator.[122] The Court considered whether aiding and abetting fell under "the scope of conduct prohibited," or whether it concerned the elements of a private right of action.[123] The Court stated that determining the elements of the section 10(b) private liability scheme poses difficult problems because the statute fails to create expressly a private cause of action and provides no guidance in regard to the elements of such a scheme.[124] With respect to the scope of conduct, however, the Court held that the text of the statute controls any decision.[125]

Concluding that aiding and abetting constitutes conduct, the Court turned to the text of section 10(b) and noted the language did not mention aiding and abetting.[126] The Court explained that it had refused to allow section 10(b) claims based on conduct not expressly proscribed by the text of the statute and emphasized its continued adherence to the language of the statute.[127]

Respondent and the SEC argued the use of the phrase "directly and indirectly" in section 10(b) demonstrated congressional intent to prohibit aiding and abetting.[128] The Court considered the statutory scheme of securities laws and rejected the argument, noting that Congress used such

---

120. Central Bank v. First Interstate Bank, 61 U.S.L.W. 3463, 3464 (Jan. 5, 1993) (No. 92-854).

121. *Central Bank*, 113 S. Ct. at 2927.

122. *Central Bank*, 114 S. Ct. at 1455.

123. *Id.* at 1445. *See generally supra* notes 91-100 and accompanying text (discussing cases that fall under the first and second prong of the analysis). Under the first prong, the text of the statute controls the decision while under the second prong, the Court is required "to infer how the 1934 Congress would have addressed the [private liability scheme] issue" had an express private right of action for aiding and abetting been included in the 1934 Act. *Central Bank*, 114 S. Ct. at 1446.

124. *Central Bank*, 114 S. Ct. at 1446.

125. *Id.*

126. *Id.* at 1446-48. "Our consideration of statutory duties, especially in cases interpreting § 10(b), establishes that the statutory text controls the definition of conduct covered by § 10(b). That bodes ill for the respondents, for 'the language of Section 10(b) does not in terms mention aiding and abetting.' " *Id.* at 1447 (quoting Brief for SEC as Amicus Curiae 8).

127. *Id.* at 1446.

128. *Id.* at 1447.

"directly and indirectly" language numerous times[129] and at no time did such language consequently impose aiding and abetting liability.[130]

After finding that the text of section 10(b) dictated the outcome, the Court concluded that an analysis of the express rights of action granted by the 1934 Act would reach the same result.[131] The Court found that while some of the express causes of action in the securities acts specify categories of possible defendants, none of the express causes of action in the 1934 Act proscribe aiding and abetting.[132] The majority concluded that interpreting section 10(b) to proscribe aiding and abetting would create an anomaly because such liability would not attach to any of the express private rights of action in the Act.[133]

In addition to considering textual analysis, the Court considered post-legislative history and policy arguments.[134] The respondents cited two post-legislative committee reports[135] the Court dismissed as containing only "oblique references to aiding and abetting liability."[136] Moreover, the Court stated the reports were merely an interpretation by a Congress that was not responsible for passing section 10(b).[137] The respondents, invoking the acquiescence doctrine, argued that congressional silence demonstrated approval of aiding and abetting liability.[138] The petitioner pointed out that Congress had rejected three prior proposed amendments that expressly would have incorporated aiding and abetting liability.[139] The majority reasoned that the acquiescence doctrine[140] de-

---

129. *See, e.g.*, 15 U.S.C. § 78g(f)(2)(C) (1988) (addressing direct and indirect ownership of stock); 15 U.S.C. § 78i(b)(2)-(3) (1988) (addressing direct or indirect interest in puts, calls, straddles, or options); 15 U.S.C. §§ 78m(d)(1), 78p(a) (1988) (addressing direct or indirect ownership of securities); 15 U.S.C. § 78t (1988) (addressing direct or indirect control of person violating the Act).

130. *Central Bank*, 114 S. Ct. at 1447; *see also* Fischel, *supra* note 53, at 95 n.83 (positing that "[o]ne plausible interpretation of the 'direct or indirectly' language is that it allows liability to be imposed upon a defendant even though such defendant does not himself use the jurisdictional means (i.e., mail a letter in interstate commerce)"). *But see In re* Atlantic Fin. Management, 784 F.2d 29, 32 (1st Cir. 1986) (interpreting the "direct or indirect" language as encompassing principal/agent liability).

131. *Central Bank*, 114 S. Ct. at 1448-49. In undertaking this analysis, the Court essentially concluded that the same outcome would result under either prong of its test. Thus, the Court's holding did not ultimately depend on the framing of the issue as "conduct" or as "an element of the private liability scheme."

132. *Id.* at 1449.

133. *Id.*

134. *Id.* at 1452-54.

135. H.R. REP. No. 910, 100th Cong., 2d Sess. 27-28 (1988), *reprinted in* 1988 U.S.C.C.A.N. 6043, 6044-45; H.R. REP. No. 355, 98th Cong., 2d Sess. 10 (1983).

136. *Central Bank*, 114 S. Ct. at 1452.

137. *Id.*

138. *Id.*

139. *Id.* at 1453.

140. In addition to the "acquiescence doctrine," the Court addressed the "rejected proposal doctrine" and the "reenactment doctrine." *See id.* at 1450-53. "Acquiescence doctrine" refers to a theory according to which judicial gloss obtains the force of legislation by congressional inaction. The "rejected proposal doctrine" posits that rejected proposals indicate that courts cannot construe statutes to resemble the rejected proposals. The "reenactment doctrine" incorporates settled statutory construction when the Congress reenacts the statute without disturbing the settled construction. For a summary of these doctrines as well as an in-depth discussion on the dangers of using congressional inaction as an indication of legisla-

served little weight and that the post-legislative history did not point to any definitive answer.[141]

The Court addressed policy arguments by considering the uncertain and expansive litigation created by aiding and abetting liability and conceded the availability of policy arguments favoring aiding and abetting liability.[142] The Court, however, refrained from basing its decision on policy rationales, stating that "policy considerations cannot override our interpretation of the text and structure of the [1934] Act, except to the extent that they may help to show that adherence to the text and structure would lead to a result 'so bizarre' that Congress could not have intended it."[143]

C. *Dissenting Opinion*

Justice Stevens wrote the dissenting opinion in which Justices Blackmun, Souter, and Ginsburg joined. The dissenters argued that the majority had given "short shrift to a long history of aider and abettor liability under [section] 10(b),"[144] adding that every court of appeals to have considered the issue has upheld such liability.[145] The dissenters argued that if any confusion existed, it concerned the elements of aiding and abetting, not its existence or validity.[146] The dissent further stated that the majority's rationale imperiled other firmly rooted theories of secondary liability not expressly addressed in securities statutes.[147]

In contrast to the majority's approach, the dissenters framed the issue as whether a plaintiff has a right to sue a person who aids and abets a primary violator.[148] The dissent reasoned that because the 1934 Act had been adopted against a backdrop of liberal construction in which "courts regularly assumed . . . that a statute enacted for the benefit of a particular class conferred" the right to sue violators, section 10(b) should confer a right of action against aiders and abettors.[149]

Putting aside the liberal backdrop underlying the passage of the 1934 Act, the dissent argued that the doctrine of aiding and abetting liability

---

tive intent, see William N. Eskridge, Jr., *Interpreting Legislative Inaction*, 87 MICH. L. REV. 67 (1988).

141. *Central Bank*, 114 S. Ct. at 1453. Here the acquiescence doctrine favored allowing aiding and abetting liability, but the rejected proposal doctrine indicated the invalidity of aiding and abetting liability. The Court noted that Congress had not reenacted § 10(b); therefore, the reenactment doctrine did not bolster the respondents arguments. *Id.* at 1452.

142. *Id.* at 1453-54.

143. *Id.* at 1453-54 (quoting Demarest v. Manspeaker, 498 U.S. 184, 191 (1991)).

144. *Id.* at 1456 (Stevens, J., dissenting).

145. *Id.*

146. *Id.* at 1457.

147. *Id.* at 1456, 1460 n.12 (mentioning conspiracy, respondeat superior, and common law agency principles).

148. The dissent argued the Court should uphold "the *private right of action* against aider and abettors . . . [and follow] the traditional common law presumption, that a statute enacted for the benefit of a particular class conferred on the members of that class the *right to sue*." *Id.* at 1456-57 (emphasis added). Thus, the dissent did not concern itself with whether § 10(b) proscribed aiding and abetting, but rather focused on a private plaintiff's right to sue—a distinct and separate issue.

149. *Id.*

had become so well established that the Court should not disturb it.[150] The dissent noted that federal criminal law imposes liability on aiders and abettors of section 10(b) violators[151] and argued that imposing civil liability on aiders and abettors would not place an unfair duty on those who Congress has opted to leave unregulated.[152] While conceding that Congress, not courts, should create rights, the dissent argued that in this case, the long history of aiding and abetting liability provided a secure foundation for the imposition of liability.[153]

III. ANALYSIS

A. *Flaws in the Dissent's Rationale*

The dissent's argument misses the point and ignores the statutory construction goals of the modern Court. The liberal backdrop of the 1934 Act allowed private rights of action against persons engaged in conduct proscribed by statute. Aiding and abetting liability, properly seen, does not hinge on whether a plaintiff has a right to sue; rather, it focuses on whether the defendant has engaged in conduct prohibited by the statute.[154] Implying a private remedy redresses conduct prohibited by the statute, while implying liability for aiding and abetting changes the proscription of the statute by expanding its scope.[155] In the first instance, Congress *has* proscribed the conduct and the Court merely allows a remedy. In the second instance, the Court proscribes conduct *and* implies a remedy. Thus, expanding the scope represents a greater usurpation of Congressional power.

The dissent compounds its misconception of the issue by failing to validate aiding and abetting liability under the rationale of *Musick, Peeler & Garrett v. Employers Insurance of Wausau*.[156] During the Contraction Era,

---

150. *Id.* at 1457-58.

151. *Id.* at 1459 (citing 18 U.S.C. § 2 (1988)). In response, the majority noted that the Court has "been reluctant to infer a private right of action from a criminal prohibition alone." *Id.* at 1455. The majority further pointed out the illogic of the dissent's argument by stating "[i]f we were to rely on this reasoning now, we would be obliged to hold that a private right of action exists for every provision of the 1934 Act, for it is a criminal violation to violate any of its provisions." *Id.* (citing 15 U.S.C. § 78ff (1988)).

152. *Id.* at 1459.

153. *Id.* at 1460.

154. *See* Texas & Pacific Ry. v. Rigsby, 241 U.S. 33, 39 (1916). On their own accord, the dissenters contend that the liberal backdrop confers "the right to sue *violators of that statute*." *Central Bank*, 114 S. Ct. at 1457 (Stevens, J., dissenting) (emphasis added).

155. *See* Fischel, *supra* note 53, at 93. The dissent consistently argues in favor of a plaintiff's right, but does not recognize that a private remedy redresses the wrong proscribed by Congress.

156. 113 S. Ct. 2085 (1993). At first blush, one might think the Court's decision in *Central Bank* directly conflicts with *Musick* as the former eliminates a form of secondary liability while the latter provides for the right of contribution among violators of § 10(b). This contention fails to understand that the right of contribution established in *Musick* is the right of contribution between persons jointly liable for the violation. *Id.* at 2086. Thus, the right of contribution established in *Musick* is not dependent on the existence of secondary liability. *See supra* notes 97-101 and accompanying text; WILLIAM L. PROSSER, HANDBOOK OF THE LAW OF TORTS § 50, at 309 (4th ed. 1971) (addressing contribution and the joint tortfeasor); Robert A. Leflar, *Contribution and Indemnity Between Tortfeasors*, 81 U. PA. L. REV. 130, 136 (1932)

courts have used a textual approach in seeking out congressional intent when implying a private remedy or defining the contours of such a cause of action.[157] The statutory scheme of the 1934 Act weighs heavily against imposing aiding and abetting liability.[158] None of the sections of the 1933 Act or the 1934 Act expressly providing for a private cause of action imposes aiding and abetting liability. Additionally, where Congress did impose controlling person liability, a form of secondary liability, it provided a good faith defense.[159] The dissent undermines its rationale by failing to address congressional intent and the negative implication of the statutory scheme.

Moreover, the dissent erred by relying on the extensive history of aiding and abetting liability in the context of section 10(b) claims. This error consists of two components: relying on congressional silence and relying on settled lower court precedent.

Congressional silence, the passage of time, and agreement among the courts of appeals does not add validity to an erroneous decision.[160] When Congress acts, it does so collectively and in an affirmative manner.[161] Congress creates law by enacting statutes and cannot legislate without such action. Thus, courts cannot reliably use the silence of Congress to infer legislative intent.[162]

(addressing contribution and equality); William L. Prosser, *Joint Torts and Several Liability*, 25 CAL. L. REV. 413, 429-43 (1937) (addressing contribution and several liability).

157. *See supra* notes 77-101 and accompanying text (discussing the Supreme Court's Contraction Era decisions).

158. *See* Bromberg & Lowenfels, *supra* note 7, at 653 (asserting that both the negative implication and post legislative history lead to the conclusion that Congress did not intend to impose aiding and abetting liability); Fischel, *supra* note 53, at 94-99 (arguing that both statutory scheme and legislative history indicate that § 10(b) does not impose aiding and abetting liability).

159. *See supra* note 35 and accompanying text.

160. *See, e.g.*, Zuber v. Allen, 396 U.S. 168, 185 (1969) (stating that "[l]egislative silence is a poor beacon to follow"; *see also* Aaron v. SEC, 446 U.S. 680 (1980). *Aaron* presented the issue of whether the SEC had to establish scienter for an injunction to enforce Rule 10b-5. Four years earlier, the Court had held that private parties must establish scienter, but lower courts had interpreted the decision as not compelling a scienter requirement in injunctive proceedings. The SEC built a strong acquiescence argument, but the Court responded by stating, "it is our view that the failure of Congress to overturn the Commission's interpretation falls far short of providing a basis to support a construction of § 10(b) so clearly at odds with its plain meaning." *Id.* at 694 n.11; *see also* Virginia Bankshares, Inc. v. Sandberg, 501 U.S. 1083, 1104 (1991) (rejecting an acquiescence argument as to § 14(a) of the 1934 Act). *See generally* Eskridge, *supra* note 140 (criticizing the use of congressional inaction as a form of precedent).

161. *See, e.g.*, U.S. CONST. art. I, § 7, cl. 2 (stating that "[e]very Bill which shall have passed the House of Representatives and the Senate, shall, before it becomes a Law, be presented to the President of the United States"); Immigration & Naturalization Serv. v. Chadha, 462 U.S. 919, 946 (1983) (stating that bicameral enactment and presentation are "integral parts of the constitutional design for the separation of powers").

162. "It does not follow . . . that Congress's failure to overturn a statutory precedent is reason for this Court to adhere to it." Patterson v. McLean Credit Union, 491 U.S. 164, 175 (1989); *see* REED DICKERSON, THE INTERPRETATION AND APPLICATION OF STATUTES 181 (1975); Reed Dickerson, *Statutory Interpretation: Dipping into Legislative History*, 11 HOFSTRA L. REV. 1125, 1133 (1983); Ernst Freund, *Interpretation of Statutes*, 65 U. PA. L. REV. 207, 214-15 (1917); John C. Grabow, *Congressional Silence and the Search for Legislative Intent: A Venture into "Speculative Unrealities"*, 64 B.U. L. REV. 737 (1984); *see also* Laurence H. Tribe, *Toward a Syntax of the Unsaid: Construing the Sounds of Congressional and Constitutional Silence*, 57 IND. L.J. 515

The acquiescence of a subsequent Congress does not indicate the intent of the enacting Congress.[163] The Court has often stated that "views of a subsequent Congress form a hazardous basis for inferring the intent of an earlier one."[164] "[E]ven when it would otherwise be useful, subsequent legislative history will rarely override a reasonable interpretation of a statute that can be gleaned from its language and legislative history prior to its enactment."[165] If the affirmative statements of a subsequent Congress do not provide a meaningful basis for inferring legislative intent, the silence of a subsequent Congress provides even less.[166]

The parties in *Central Bank* advanced competing arguments based upon post-legislative history. While the respondent argued that acquiescence and reference to secondary liability in committee reports weigh in favor of allowing aiding and abetting liability, the petitioner argued the rejected amendments that would have incorporated aiding and abetting language into section 10(b) weigh against such liability.[167] In favoring the acquiescence doctrine over the rejected proposal doctrine, the dissent fails to support its decision with any reasonable rationale. The post-legislative history of section 10(b), at best, provides an inconclusive answer.

While all previous courts of appeals that considered the question may have agreed that aiding and abetting liability existed, considerable confusion existed among the same courts regarding the proper application of such liability.[168] The confusion among circuit courts may stem from a

---

(1982) (approving only limited use of congressional silence as an interpretive tool). The dissent's acquiescence argument would carry more weight if the Supreme Court previously had upheld the validity of aiding and abetting liability, yet the Court had twice reserved for decision the validity of such liability. *See supra* note 4 and accompanying text.

163. *See, e.g.*, Lawrence C. Marshall, *"Let Congress Do It": The Case for an Absolute Rule of Statutory Stare Decisis*, 88 MICH. L. REV. 177, 186-200 (1989) (contending that ignorance, inertia, interpretational ambiguity, and irrelevance make it difficult to infer congressional approval from congressional inaction); Richard A. Posner, *Statutory Interpretation—in the Classroom and in the Courtroom*, 50 U. CHI. L. REV. 800, 809-10 (1983) (suggesting that reliance on post-enactment legislative materials usurps prior Congressional power without "going through the constitutionally prescribed processes for repeal"); Patricia M. Wald, *Some Observations on the Use of Legislative History in the 1981 Supreme Court Term*, 68 IOWA L. REV. 195, 205 (1983) (stating that "it is particularly risky to draw inferences from subsequent congressional refusals to act").

164. *See* Andrus v. Shell Oil Co., 446 U.S. 657, 666 n.8 (1980) (quoting United States v. Price, 361 U.S. 304, 313 (1960)).

165. Consumer Prod. Safety Comm'n v. GTE Sylvania, Inc., 447 U.S. 102, 118 n.13 (1980).

166. Grabow, *supra* note 162, at 750.

167. *See supra* note 139 and accompanying text.

168. *See, e.g.*, Feldman, *supra* note 56, at 72-73 (claiming that such confusion must eventually be addressed). Additionally, the dissent may have given "short shrift" to lower court questioning of the validity of private actions based on aiding and abetting. *See, e.g., id.* at 72-73 (claiming that such confusion must eventually be addressed). The dissent also may have given "short shrift" to lower court questioning of aiding and abetting liability. *See* Akin v. Q-L Inv., Inc., 959 F.2d 521, 525 (5th Cir. 1992) (stating that there is a "powerful argument that . . . aider and abettor liability should not be enforceable by private parties pursuing an implied right of action"); Congregation of the Passion, Holy Cross Province v. Kidder Peabody & Co., 800 F.2d 177, 183 (7th Cir. 1986) (providing that courts have "frankly acknowledged that, in light of recent Supreme Court cases, there is some ambiguity about the existence of a civil cause of action for aiding and abetting"); Little v. Valley Nat'l Bank, 650 F.2d 218, 220 n.3 (9th Cir. 1981) (providing that "[t]he status of aiding and abetting as a basis for liability

conceptual problem with the use of aiding and abetting in the section 10(b) context. The theory of aiding and abetting developed mainly from criminal law and, to a lesser extent, tort law.[169] These areas of law mainly revolve around physical acts.[170] The aiding and abetting concept in criminal and tort law rests on degrees of physical presence and action, which in the physical context indicate culpability.[171]

In the context of physical torts and crimes, one can rely on physical presence and action to distinguish between primary violators and aiders and abettors. In the securities context this analysis results in confusion.[172] Reliance on these antiquated doctrines results in irreconcilable conceptual problems and uncertainty in application.[173] The confusion and conflicting results in applying the aiding and abetting doctrine evidences this lack of usefulness.

In the spirit of "if it ain't broke, don't fix it," the dissent argued that the majority would stand on firmer footing if aiding and abetting liability interfered with the effective operation of securities laws.[174] Such an approach ignores at least two adverse consequences of aiding and abetting liability. The first problem becomes evident when comparing aiding and abetting liability with "controlling person" liability. Sections 15 of the 1933 Act and 20 of the 1934 Act impose "controlling person" liability.[175] In addition, these sections provide for a good faith defense. To circumvent the use of the good faith defense, plaintiffs sought to impose aiding and abetting liability instead of, or in addition to, controlling person liability.[176] This use of aiding and abetting frustrates congressional purpose by

---

under the securities laws is in some doubt"); Benoay v. Decker, 517 F. Supp. 490, 495 (E.D. Mich. 1981) (providing that "[i]t is also doubtful that a claim for 'aiding and abetting' or 'conspiracy' will continue to exist under 10(b) [because *Hochfelder*] implicitly holds that aiding and abetting liability will not exist apart from liability for a direct violation"), *aff'd*, 735 F.2d 1363 (6th Cir. 1984); Seattle-First Nat'l Bank v. Carlstedt, 101 F.R.D. 715, 722-23 (W.D. Okla. 1984) (providing that "[t]he notion that aiding and abetting securities fraud constitutes a justiciable violation of law is itself a questionable assertion"), *rev'd*, 800 F.2d 1008 (10th Cir. 1986). Again, the dissent makes no attempt to validate its reasoning in light of the confusion in the application of aiding and abetting liability.

169. Tort law does not fully embrace the concept of primary and secondary liability; however, the concept of joint tortfeasance is widely accepted. This concept recognizes independent contributions to an indivisible tort, with each tortfeasor equally liable for the entire harm. *See* Pollak, *supra* note 11, at 246-47.

170. *See* RESTATEMENT (SECOND) OF TORTS § 876 (1989) (illustrating only physical harms to the plaintiff or the plaintiff's property). The petitioner directed the Court's attention to this point but the Court did not discuss the difficulties of importing a physical tort concept into the economic arena of securities law. *See* Petitioner's Opening Brief 28 n. 20. Notably, courts do not usually apply § 876 outside the context of physical torts except in § 10(b) cases. 4 BROMBERG & LOWENFELS, *supra* note 7, § 8.5(614)(4); *see* Herman & MacLean v. Huddleston, 495 U.S. 375, 388 (1983) (claiming that securities law is not coextensive with common law).

171. *See* RESTATEMENT (SECOND) OF TORTS § 876 (1989); 2 WAYNE LAFAVE & AUSTIN W. SCOTT, CRIMINAL LAW § 6.2, at 495-98 (1979).

172. Ruder, *supra* note 74, at 621-22 (suggesting that tort law offers little or no help in the securities context).

173. Feldman, *supra* note 56, *passim*.

174. *Central Bank*, 114 S. Ct. at 1459.

175. *See supra* notes 34-35 and accompanying text.

176. For a discussion of the distinction between aiding and abetting liability and other forms of secondary liability, see Sally T. Gilmore & William H. McBride, *Liability of Financial*

allowing plaintiffs to prevent defendants' use of a good faith defense. By eliminating the use of aiding and abetting as a basis for a section 10(b) cause of action, the Court rightfully has prevented plaintiffs from circumventing the good faith defense provided by Congress.

A second problem with aiding and abetting liability arises when a private party bases a cause of action on silence. To recover from a primary violator for silence, the plaintiff must show a duty existed and that he relied on that duty.[177] A plaintiff could, however, bring an aiding and abetting claim based on silence and recover without showing either duty or reliance.[178] This functionally allows plaintiffs to side-step the Court's restrictions on section 10(b) liability.

Efficient financial markets, while inherently chaotic, require a stable and predictable legal foundation upon which they can rely. The inconsistent, ad hoc manner in which courts apply aiding and abetting liability frustrates the market system. In the past, the resulting uncertainty has left facilitators of financial markets unsure about the boundaries of permissible conduct and the scope of possible liability.[179] Additionally, the uncertainty has resulted in defendants having to settle even the most frivolous claims.[180] These increased costs eventually flow to the investor.[181] In the

---

Institutions for Aiding and Abetting Violations of Securities Laws, 42 WASH. & LEE L. REV. 811, 813-814 (1985).

177. *See, e.g.*, Chiarella v. United States, 445 U.S. 222, 228 & n.9 (suggesting silence is actionable only when a duty exists between the parties).

178. *See, e.g.*, Doherty, *supra* note 72, at 851-52 (concluding that an aider and abettor can be found liable for failure to disclose despite absence of duty). Courts focus on a defendant's intent in a claim based on silence, rather than on duty. Tort law, however, does not require intent as an element of an aiding and abetting claim. Patrick J. McNulty & Daniel J. Hanson, *Liability for Aiding and Abetting by Silence or Inaction: An Unfounded Doctrine*, 29 TORT & INS. L.J. 14, 39-43 (1993).

179. Feldman, *supra* note 56, at 73.

180. Herrick K. Lidstone, Jr. & Michael J. Norton, *Professional Advisors: "Am I My Brother's Keeper?"*, 23 COLO. LAW. 1795, 1795 (1994). Because of the uncertainty surrounding § 10(b) claims, the defendant often chooses to pay a settlement rather than run the risk of an adverse judgment. *See, e.g.*, 2 LOUIS LOSS, SECURITIES REGULATIONS 1792 (2d ed. 1961) (relatively few 10b-5 cases go to trial on the merits); Michael P. Dooley, *Enforcement of Insider Trading Restrictions*, 66 VA. L. REV. 1, 27 n.129 (1980) (plaintiffs have an incentive to bring groundless and contrived claims because of the likelihood of settlement as opposed to adjudication); Thomas M. Jones, *An Empirical Examination of the Resolution of Shareholder Derivative and Class Action Lawsuits*, 60 B.U. L. REV. 542, 545 (1980) (presenting the results of a study showing that only 2.3% of the litigated shareholder claims resulted in judgment for the plaintiffs).

181. *See, e.g.*, Ralph K. Winter, *Paying Lawyers, Empowering Prosecutors, and Protecting Managers: Raising the Cost of Capital in America*, 42 DUKE L.J. 945, 962-65 (1993) (discussing the effects of unnecessarily raising the cost of capital). "When a corporation pays a judgment or a settlement, the value of the corporation's stock may fall and its cost of capital rise. In that event, not only do shareholders sustain a loss, but the productivity of the firm may decline—at some cost to the society." Frankel, *supra* note 84, at 577-78.

end, the beneficiary[182] of the statute pays the cost of the uncertainty, a result which frustrates congressional purpose.[183]

B. *Majority's Rationale*

The majority's approach follows the textual analysis laid out in its previous Contraction Era decisions. Instead of treading upon the "quicksand" of the acquiescence doctrine,[184] the majority directs courts to look at the text of the statute and the statutory scheme to ascertain congressional intent. While this approach does not result in absolute precision, it allows participants in the financial market to predict adequately the scope of the statute.[185] This predictability will minimize unnecessary costs in financial markets and thereby benefit investors.

The Court's decision establishes a two-prong analysis for deciding issues presented by the implied private remedy litigation under section 10(b). If the Court faces a question as to the components of a private cause of action, it will decide the issue based on congressional intent as ascertained from the overall statutory scheme. If the Court faces a question as to what conduct a party can base a private cause of action on, it will turn to the text of the section at issue. While the second prong involves a fairly simple analysis, the first prong may cause hesitation because the Court is forced to infer how the enacting Congress would have decided the issue.[186]

This approach will help prevent securities law from becoming loose-jointed because it focuses on consistency and conformity with the overall statutory scheme. The dissent's rationale would allow erroneous lower court precedents to flourish. The majority's statutory construction analysis curbs the growth of unintended secondary liability and leaves policy decisions to Congress.

C. *Possible Ramifications*

Plaintiffs and courts often used aiding and abetting as a "misnomer," labeling the defendant as an "aider and abettor" when "primary violator" would more accurately describe the defendant.[187] Since the inception of

---

182. Some economists have concluded that empirical evidence does not support the belief that investors have benefitted from securities regulation. SUSAN M. PHILLIPS & J. RICHARD ZECHER, THE SEC AND THE PUBLIC INTEREST (1981); George J. Benston, *Required Disclosure and the Stock Market: An Evaluation of the Securities Exchange Act of 1934*, 63 AM. ECON. REV. 132 (1973); Greg A. Jarrell, *The Economic Effects of Federal Regulation of the Market for New Security Issues*, 24 J.L. & ECON. 613 (1981); George J. Stigler, *Public Regulation of the Securities Markets*, 37 BUS. LAW. 721 (1964). For a general discussion of the economics of regulation, see STEPHEN BREYER, REGULATION AND ITS REFORM (1982).

183. *See* Akin v. Q-L Inves., 959 F.2d 521, 525 (5th Cir. 1992).

184. *See* Helvering v. Hallock, 309 U.S. 106, 121 (1940).

185. While the textual analysis will not allow participants to predict flawlessly the application of a statute, it does assure participants that courts cannot impose liability for a violation not expressly stated in the statute.

186. In considering the scope of conduct prohibited, "the text of the statute controls our decision." *Central Bank*, 114 S. Ct. at 1446.

187. *Akin*, 959 F.2d at 526.

aiding and abetting liability, courts have seldom sought to make a meaningful distinction between aiders and abettors and primary violators.[188] The majority in *Central Bank* concedes that courts need not apply the scope of primary liability narrowly.[189] As a result, courts and litigants will now begin to flesh out the limits of primary liability.

Opponents of the Court's decision in *Central Bank* claim the decision will permit fraudulent conduct to run rampant.[190] These opponents believe that *Central Bank* will immunize lawyers, accountants, bankers, and other similarly situated parties from liability for fraudulent conduct.[191] This fear ignores the main focus of the Court's decision: to violate section 10(b), the defendant must have committed some fraudulent act. If an accountant, lawyer, banker, or any other party fraudulently provides an investor with false information, the investor can bring suit against that person as a primary violator.[192] Most likely, the Court's decision will act to exonerate only those persons who did no vouching or played no role in the falsifying itself.[193] Thus, the decision does not give the "green light" to fraudulent action but instead forces plaintiffs to allege and subsequently prove fraud on the part of the defendant.

While the Court's decision probably will not disturb the implied private cause of action for violations of section 10(b),[194] it casts doubt on the validity of other forms of judicially imposed secondary liability under section 10(b) such as respondeat superior, agency, and conspiracy.[195] At the outset, it would seem that the Court's decision jeopardizes all of these

---

188. *See supra* notes 60-76 and accompanying text.

189. "Any person or entity, including a lawyer, accountant, or bank, who employs a manipulative device or makes a material misstatement (or omission) on which a purchaser or seller of securities relies may be liable as a primary violator under 10b-5, assuming all of the requirements for primary liability under Rule 10b-5 are met." *Central Bank*, 114 S. Ct. at 1455.

190. Senator Metzenbaum stated that *Central Bank* "gives clearly fraudulent behavior the green light." 140 CONG. REC. S9460 (daily ed. July 21, 1994).

191. *Id.*

192. *See, e.g.*, Breard v. Sachnoff & Weaver, Ltd., 941 F.2d 142, 144-45 (2d Cir. 1991) (finding allegations against preparer of an offering circular sufficient to survive motion for summary judgment without relying on an aiding and abetting analysis); SEC v. Washington County Util. Dist., 676 F.2d 218 (6th Cir. 1982) (holding that primary liability does not require face-to-face contact).

193. However, prior to the *Central Bank* decision courts hesitated to impose aiding and abetting liability when the defendant did "not engage in conduct that intentionally misleads or lulls a victim." IX LOUIS LOSS & JOEL SELIGMAN, SECURITIES REGULATION 4486 (3d ed. 1992). Also, aiding and abetting liability remains a potential theory in private actions based on state securities law violations. *See generally* Douglas M. Branson, *Collateral Participant Liability Under State Securities Laws*, 19 PEPP. L. REV. 1027 (1992). Plaintiffs' counsel often may seek remedies in state courts due to the Supreme Court's decisions restricting private actions under the federal securities laws. MARC I. STEINBERG, UNDERSTANDING SECURITIES LAW § 7.03, at 133 (1989).

194. The existence of a private cause of action is "simply beyond peradventure." Herman & MacLean v. Huddleston, 459 U.S. 375, 380 (1983). *But see* Michael J. Kaufman, *A Little "Right" Musick: The Unconstitutional Judicial Creation of Private Rights of Action Under Section 10(b) of the Securities Exchange Act*, 72 WASH. U. L.Q. 287, 297-335 (1994) (arguing that the implied private cause of action under section 10(b) is unconstitutional).

195. *See, e.g.*, A.J. White & Co. v. SEC, 556 F.2d 619, 624 (1st Cir. 1977) (holding a brokerage firm can act only through its agents and is accountable for the actions of its officers); Harrision v. Dean Witter Reynolds, Inc., 715 F. Supp. 1425, 1431 (N.D. Ill. 1989) (discussing respondeat superior liability in the § 10(b) context); Eastwood v. National Bank of Com-

forms of secondary liability, yet a distinction may exist between conspiracy and the two other forms of secondary liability mentioned above. Conspiracy, similar to aiding and abetting, imposes liability based upon conduct (an agreement to engage in proscribed conduct). Respondeat superior and agency represent a more vicarious form of liability based on a relationship between the primary violator and those with the burden of secondary liability.[196] Both respondeat superior and agency are more akin to "controlling person" liability and therefore courts might view them as more consistent with the statutory scheme. If so, conspiracy liability would not survive while respondeat superior and agency would be fertile claims. It appears unlikely, however, that courts would chose to utilize this distinction.[197]

CONCLUSION

The continuing strict textual approach applied by the majority in *Central Bank* injects a degree of certainty and clarity into the complex area of securities law. The Court's rationale provides securities attorneys and their clients with a meaningful basis from which they can determine the scope of permissible conduct and potential liability. Contrary to the opinion of some alarmists, the Court's decision does not give a "green light" to fraudulent conduct but rather forces courts to more carefully consider the culpability of defendants.[198]

Paul Dmitri Zier

---

merce, 673 F. Supp. 1068, 1079-81 (W.D. Okla. 1987) (discussing conspiracy liability for a § 10(b) violation).

196. While the terms secondary liability and vicarious liability are often used interchangeably, the notion of vicarious liability would be limited more properly to liability based on relationships rather than liability based on conduct. *See* Kuehnle, *supra* note 15, at 318 n.28.

197. *Central Bank* echoes the substance of Justice Kennedy's concurring-in-part and dissenting-in-part opinion in *Virginia Bankshares*. *See supra* note 1 and accompanying text. Moreover, it would seem more onerous to allow judicially created forms of liability such as respondeat superior and agency to be applied when Congress's intent was to impose controlling person liability. *See* Christoffel v. E.F. Hutton & Co., 588 F.2d 665, 667 (9th Cir. 1978); Rochez Bros., Inc. v. Rhoades, 527 F.2d 880, 884-86 (3d Cir. 1975).

198. No longer will investors be able to "have their cake and eat it, too" as they will not be able to recover from advisors on an aiding and abetting theory if the investment loses money for any reason. *See* Lidstone & Norton, *supra* note 180, at 1796.