METHOD TO THEIR MADNESS: DISPELLING THE MYTH OF ECONOMIC RATIONALITY AS A BEHAVIORAL IDEAL

John Dobson

ABSTRACT

Although not immediately apparent, the discipline of behavioral finance is rapidly adopting an implicit prescriptive agenda. Behavioral finance does not merely describe financial market reality, it shapes it. Economic rationality is taken as the ideal toward to which individuals 'should' strive.

In this paper I show that, as a behavioral ideal, economic rationality is unjustified both from a strictly economic perspective, and from a moral perspective. In short, there is nothing inherently "wrong" with economically irrational participants in the business environment. Indeed such participants will actually enhance the efficiency, and the ethicality, of business.

1. INTRODUCTION

rationality itself, whether theoretical or practical, is a concept with a history: indeed, since there are a diversity of traditions of enquiry, with histories, there are . . . rationalities rather than rationality (Alasdair MacIntyre, *Whose Justice, Which rationality*, 1988, p. 9).

The rapidly growing discipline of behavioral finance is generally viewed as a value-free descriptive subject that makes no pronouncements on how agents "should" behave. Behavioral finance, most academics and practitioners would argue, attempts merely to describe and account for how individuals and groups actually behave in financial environments. In other words, the general consensus is that "behavioral finance" implies no particular moral or ethical agenda.

On closer inspection, however, behavioral finance does espouse a prescriptive ideal of how economic agents "should" behave. Consider the following statement in a recent behavioral finance article that appeared in the *Financial Analysts Journal*: "The way the world *should be* (the rational economic paradigm) and the way the world is (behavioral tendencies) will always be in tension, but the introduction of psychological antecedents into the analysis of financial anomalies is not a negation of the rational economic paradigm". Note the view that the way individuals "should" behave in financial markets is in a manner consistent with the rational economic paradigm. The authors compound this implication by referring to behavior that is inconsistent with economic rationality as "errant" (p. 56) behavior.

Many other financial economists who write about behavioral finance seem to think it natural to set up a dichotomy between actual observed behavior on the one hand, versus some behavioral ideal on the other; where the behavioral ideal comes from financial-economic theory's concept of rationality. Financial economists often refer to irrational behavior as behavior "off the equilibrium path", as if those individuals who behave irrationally are in some way straying or deviating from an ideal route.

Even in the formative stages of the discipline, therefore, behavioral finance has adopted a normative (i.e. prescriptive) agenda. Behavioral finance is not merely about the application of the theories and methodology of psychology in an attempt to explain behavior. Albeit implicitly, behavioral finance is also about prescribing how agents should behave: agents should behave rationally, where rationality is defined strictly in terms of financial economic theory.

In this paper I analyze behavioral finance's implicit normative agenda. Specifically, I argue that it is unjustified and injustifiable from both an economic and a moral perspective to espouse financial-economic rationality as a behavioral ideal. In short, there is no meaningful way in which we can say that investors, managers, brokers, or anyone else for that matter "should" behave in a way consistent with economic rationality.

The remainder of the paper is split into three parts. First, I describe exactly what is meant by financial-economic rationality. Second, I show why such a rationality concept is unjustified as a normative ideal from a strictly economic perspective. Third, I show why such a rationality concept is injustifiable as a normative ideal from a strictly moral perspective.

2. WHO IS THE RATIONAL AGENT IN FINANCIAL ECONOMICS?

Rationality in financial economics is founded on the five axioms of cardinal utility, as first enumerated by von Neumann and Morgenstern (1947), plus one additional axiom. These original five axioms are *comparability, consistency, independence, measurability, and ranking*:

- comparability; the individual can make comparisons between preferences.
- consistency; these comparisons are consistent over an array of alternatives.
- independence; original preference orderings are independent of new preference alternatives.
- measurability; preferences are measurable.
- ranking; preferences can be consistently and ordinally ranked.

In essence, the five axioms define rationality in terms of an individual's ability to make consistent preference orderings over a broad spectrum of choices: "We wish to find the mathematically complete principles which define "rational behavior" for the participants in a social economy, and derive from them the general characteristics of that behavior" (p. 31). Furthermore, "people are assumed to be able to make these rational choices among thousands of alternatives" (Copeland & Weston, 1988, p. 80).

The axioms are thus based on a very mathematical and instrumental notion of what it means to be rational: they are all concerned with defining instrumental rationality in terms of the consistent ranking of preferences. For example, if you are an investor choosing stocks in which to invest, and you prefer IBM to Microsoft, and you prefer Microsoft to Netscape, then to be rational you must prefer IBM to Netscape; furthermore your degree of preference for IBM over Netscape, along with your preferences for thousands of other securities, must stay constant no matter how many more stocks are added to your opportunity set.

Note that these five axioms make no normative statement concerning whether the agent has any specific goal, or what the goal of the agent should be; the axioms simply require that the agent act in a consistent manner in ordering preferences.

Financial-economic theory's *sixth* axiom, however, has just such prescriptive implications. As Thomas Copeland and Fred Weston's leading finance text puts it: "Having established the five axioms we add to them the assumption that individuals *always* prefer more wealth to less" (1988, p. 80, my emphasis). Personal wealth

maximization is thus a rational agent's *sine qua non*: no matter what the context, no other ultimate objective is allowed or considered.

In relating the five axioms to this sixth axiom, a useful distinction can be made between *instrumental* rationality and *substantive* rationality. In *The Protestant Ethic and The Spirit of Capitalism*, Max Weber made this distinction in labelling two types of rationality as "zweckrationalitat" (formal or instrumental rationality) and "wertrarationalitat" (values-based or substantive rationality). In essence, instrumental rationality concerns how the agent goes about achieving the desired objective, whereas substantive rationality concerns identifying the desired objective itself. Jennifer Moore distinguishes between the two concepts as follows:

> The primary feature of instrumental rationality is that it does not choose ends, but accepts them as given and looks for the best means to achieve them. In instrumental rationality, reason is subordinated to and placed at the service of ends outside itself. In . . . [*substantive rationality*], in contrast, reason is free ranging. It is not the servant of any end. Rather, it subjects every end to its *own* standards of evaluation and criticism (1991, p. 63).

von Neumann and Morgenstern's five axioms clearly pertain to instrumental rationality. They do not stipulate an ultimate objective but merely require that agents pursue some given objective in a consistent and logical manner. The substantive rationality premise of financial economics is provided by the sixth axiom: the opportunistic and atomistic pursuit of material gain ad infinitum. No justification is supplied in the finance literature in the form of empirical evidence to support this substantive rationality premise (indeed there is ample evidence that in many environments individuals are not motivated primarily by personal wealth maximization); nor is any normative argument supplied to defend the notion that this is how agents *should* behave.

The six axioms are merely a framing effect: a way of placing behavior in a relatively simple mathematical context. They make no claim to factual accuracy in all situations, or even in most situations, in which finance professionals find themselves. Nor are the axioms meant to make any moral claim to the effect that this is how individuals should behave in any given situation. As John Boatright observes: "Economics does not make any value judgement about the goods that people prefer or about the selfishness that is assumed" (1999, p. 48). Financial-economic rationality is a simplifying assumption, and nothing more.

3. WHY THE "RATIONAL" AGENT CANNOT BE VIEWED AS AN *ECONOMIC* IDEAL

> I carved a massive cake of beeswax into bits and rolled them in my hands until they softened . . . Going forward I carried wax along the line, and laid it thick on their ears. They tied me up, then, plumb amidships, back to the mast, lashed to the mast, and took themselves again to rowing. Soon, as we came smartly within hailing distance, the two Sirens, noting our fast ship off their point, made ready, and they sang . . . The lovely voices in ardor appealing over the water made me crave to listen, and I tried to say 'Untie me!' to the crew, jerking my brows; but they bent steady to the oars (Homer, c. –900, pp. 227–228).

A significant contributor to the canon of financial-economic theory is Stewart Myers. In "The Determinants of Corporate Borrowing" (1977), Myers employs the above excerpt from Homer's Odyssey to illustrate the paradoxical nature of rationality in financial economics. The Sirens' song in financial economics is the opportunistic pursuit of personal material gain. The term "opportunistic" implies that the agent will do whatever is necessary in the interests of this pursuit. The agent will lie, cheat, steal, etc., so long as this behavior is construed as wealth maximizing. Agents are "lashed to the mast" in the sense that they are assumed always to act opprtunistically. In Myers' model of equilibrium in debt markets, for example, borrowers pay a higher interest rate in order to compensate lenders for the assumed opportunism of borrowers: borrowers will renege on loan agreements as soon as it is in their material self-interest to do so. There is no "benefit-of-the-doubt" given here, opportunistic wealth maximization (i.e. the six axioms) is the only rationality premise considered.

What is important to note here is that no one actually gains from this opportunistic wealth maximization; it is self-defeating. Returning to Myers' model, the borrowers have to pay higher interest costs and so borrow less, which in turn hurts the lenders. So economic rationality ensures that its own self-declared objective is never attained: wealth is not maximized for borrowers, for lenders, or indeed for the economy in aggregate. But why *must* this be so?

To answer this question we must look more closely at the behavior, and specifically the interaction, of these "rational" agents. To model behavior, Myers employs game theory, and it is through applying game theory to financial markets – what is generally called agency theory – that we can see the self-defeating nature of economic rationality. Indeed, as a theoretical discipline, behavioral finance can be seen as an application of agency theory. From its humble beginnings in the corporate finance theory of Myers and others, the 1970s agency theory has now become the predominant methodoloy of theoretical finance.

Agency theory analyses the situation, ubiquitous in financial markets, in which "one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision-making authority to the agent" (Jensen & Meckling, 1976, p. 308). These agency-theory models can be loosely categorized into two types, namely *adverse-selection* and *moral-hazard*. The difference between these two categories is essentially a function of the nature and degree of uncertainty inherent in the contractual situation.

Consider first moral-hazard-type agency problems, here the contractual situation is ostensibly a simpler one in which there is only one *type* of agent. There may be no informational asymmetry and the agency problem may simply stem from the principal's inability to *control* the actions of the agent. Or there may be informational asymmetry in which case the agency problem stems from the principal's inability to observe directly some information that affects the actions or the performance of the agent (e.g. it might be hard for the principal to discern whether a stockbroker was genuinely committed to executing a client's security transaction at the best available price).

The classic agency problem of this type is managerial perquisite consumption. As a firm moves from private to public ownership, there is a separation of ownership and control. The owners bear the cost of managers' perquisite consumption (e.g. business lunches, corporate jets, generous stock-option packages, etc.), but the managers make the decisions on how many "perqs" to consume. Barring effective accountability – in other words barring a resolution to the agency problem – a "rational" wealth-maximizing management, who no longer bears the full cost of its perqs, may be predisposed to consume perqs to an excessive degree: specifically, to a degree that compromises the value of the firm as a whole. Potential shareholders and bondholders, cognizant of management's "rational" predisposition, will lower the price at which they are willing to buy the firm's equity or debt.

Once again, therefore, the *cost* of agency is invariably borne by the agent. Hence the "finance paradox" that assumes agents are unable to resist the Sirens' Song of opportunism, even though they must rationally realize that such a predisposition is self-defeating. In short, a predisposition to opportunistic wealth maximization does not maximize wealth for either the individual opportunist, or for the economy in aggregate: it is second best.

This point is illustrated clearly by the following simple game.

Figure 1 illustrates a simple game between two players: "A" and "B". Each player represents a stakeholder or group of stakeholders. So, for example, player "A" might represent a group of shareholders considering an investment in a company whose management is represented by player "B". David Kreps summarizes the game's play as follows:

> First A must choose whether or not to trust his opponent. If he (A) elects not to trust B, then both A and B get nothing. If he elects trust, B is made aware of this fact and is given the option either to honor that trust or to abuse it. If A trusts B and she (B) chooses to honor that trust, both get $10. But if A trusts B and she chooses to abuse it, B gets $15 and A loses $5 (1984, p. 12).

Assume that each player's payoff from the game is common knowledge. In other words there is no informational asymmetry and, to the extent that there is an agency problem, it would be characterized as one of simple moral hazard. As Kreps

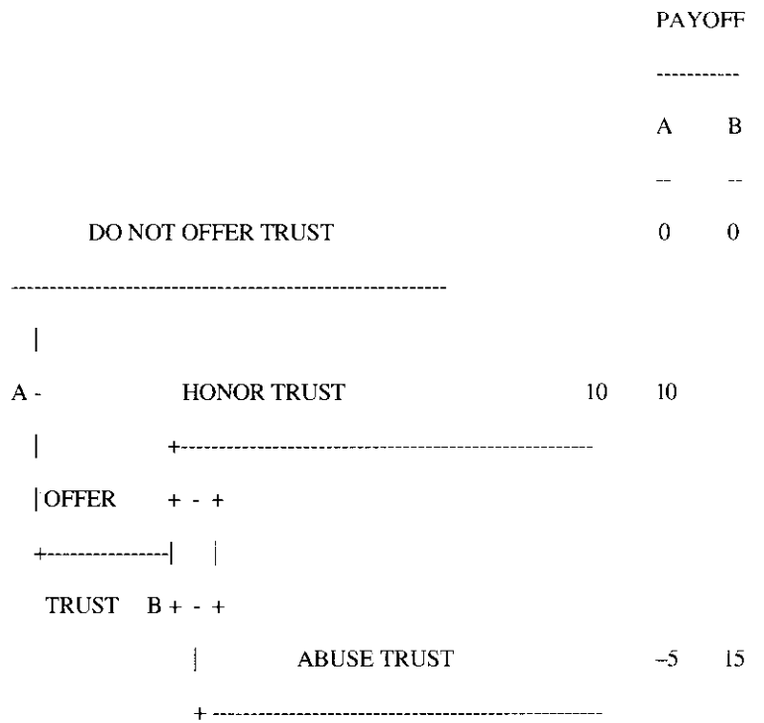

Fig. 1. A Game Tree Illustrating a Contractual Relation between Two Parties in which an Inherent Conflict of Interest Exists.

explains, the game begins with player "A" deciding whether or not to trust player "B". If he ("A") does decide to trust "B", then she ("B") must decide whether to honor or abuse that trust. Readers familiar with Game Theory will recognize Fig. 1 as a one-sided version of the infamous *Prisoners' Dilemma* game. If we assume that both players are rational and thus are primarily motivated to maximize their payoff, then presumably, if called upon to move, "B" will abuse the trust vested in her by "A". Realizing this, "A" will never offer trust and a contract between these two players will not be entered into. The most "reasonable" outcome for this game, therefore, is for each player to receive a payoff of $0. Formally, this is the unique *Nash* equilibrium, in which each player's move is "rational" given the move of the other player.

Such an outcome is clearly not the most desirable, however, either from the point of view of the two players as individuals or from the point of view of the economy as a whole, in that the maximum total payoff of $20 is not attained (this

would be the *first-best* outcome). The unwillingness of player "A" to trust player "B" has cost both players $10. But then why should "B" honor trust when her immediate payoff is maximized by abusing it? And whatever "B" might actually plan on doing, why should "A" assume "B" is going to honor trust when he can see that abusing it yields her the higher payoff? In short, economically "rational" agents are unable to reach the desirable outcome; once again simple self-interest has proved self-defeating.

But, returning to Fig. 1, what if the game were repeated. In reality individuals generally deal with one another more than once, and in such a repeated-game environment would not a rational agent now honor trust? Thus, could not enlightened self-interest thus overcome the contractual impasse?

This is the much vaunted "reputation solution" in game theory: the agent honors trust and cooperates in order to build a reputation for cooperation. Game theory, however, indicates that equilibria based on reputation are quite fragile. In the context of Fig. 1, for example, Dave Kreps shows that reputation will enforce a contract between the two players only when there is some uncertainty about the length of the game (i.e. the number of iterations), or when there is uncertainty about the rationality of one of the players (i.e. whether they are both self-interested wealth maximizers). Indeed, even when one of these conditions is met, the agent's desire to maintain her reputation may induce her to honor the contract only some of the time. Similarly, in one of the most extensive game-theory models yet developed – namely Diamond's (1989) model of reputation acquisition in debt markets – agents never actively strive to build reputations. Some merely acquire reputations for timely debt repayment through luck. Once acquired these reputations *may* be actively maintained until the endgame is reached: at which point agents revert to opportunistic behavior.

The economic undesirability of "rational" self-interest becomes even more apparent when we consider adverse-selection, where the uncertainty stems from an *asymmetry of information* that precludes the principal from costlessly identifying the *type* of agent. For example, imagine two stockbrokers: "Churn-and-burn" and "Buy-and-hold". Buy-and-hold tends to offer relatively superior long-run returns to clients than does Churn-and-burn, however prospective clients are not able readily to observe this difference. In other words, the different earnings prospects of the two brokers is not readily apparent from their financial statements or other generally available information. Thus, the contractual environment is one characterized by informational asymmetry.

The ability of the principals, in this context the clients, to make optimal (i.e. wealth maximizing) investment decisions is a function of their ability to distinguish between the two agents, in this context the two brokers. "Buy-and-hold" must in some way signal unambiguously to investors that he/she is superior to "Churn-and Burn".

With this signalling solution to adverse-selection-type agency problems, the challenge is for the "good" agent to devise a signal that cannot be mimicked by the "bad" agent. In addition, this signal must not be so costly that it is uneconomic for even the good agent to emit. If the good agent is able to devise and emit such a signal then it engenders what agency theorists term a *separating equilibrium* in which the two agents become distinguishable to principals and thus the informational asymmetry is overcome. If such a signal does not exist, then the informational asymmetry endures and a *pooling equilibrium* ensues (Spence, 1976).

Even if a separating equilibrium is achieved, however, there are generally costs involved. Whatever form the signal takes, the "good" agent bears the cost of emitting the signal. In other words, if there had been no initial informational asymmetry or if the "bad" agent chose openly to reveal its type, then the bad agent would be no worse off and the good agent would be better off through not having to fund the signal. Formally, the equilibrium is said to be *second-best* because there is a deadweight or "dissipative" cost levied on the economy in aggregate (a cost to the good agent not recouped by the bad agent). Jensen and Meckling (1976) label this deadweight cost the "residual loss"; it results directly from the contractual enforcement problem between principal and agent. Thus, agency problems, even if they are overcome, are not zero-sum games. We are not dealing merely with a redistribution of wealth from principal to agent, but rather with an absolute wealth loss to the economy in aggregate. This absolute wealth loss is a direct result of the bad firm's pursuit of self-interest "with if necessary guile and deceit" (Noreen, 1988, p. 359). Note that in equilibrium the bad firm does not benefit from this opportunistic behavior. Indeed, as part of the economy, it too suffers the consequences of the dead-weight loss.

Reputation or signalling, therefore, may work as contractual enforcement mechanisms in some stages of some contractual situations. But the above studies indicate that there are many environments in which they will not work, or at least not work costlessly. Furthermore, in equilibrium, the cost is invariably borne by the agent. Thus, once again, agents do not benefit from opportunistic behavior; wealth maximization is not achieved, either for the individual or the aggregate. In a broader context, Bowie makes a similar observation; "It only pays to lie or cheat when you can free ride off the honesty of others . . . The conscious pursuit of self-interest by all members of society has the collective result of undermining the interests of all" (1991, pp. 11–12).

Economic game theory, therefore, provides no normative justification for viewing financial-economic rationality as a behavioral ideal. An economy populated with agents adhering strictly to the above six axioms of rationality would, at best, be an inefficient economy; and would, at worst, degenerate into what the Eighteenth Century moral philosopher Thomas Hobbes called a "war of all against all" in which life is "nasty, brutish, and short"!

4. WHY THE "RATIONAL" AGENT CANNOT BE VIEWED AS A *MORAL* IDEAL

In the case of economic efficiency, this paper has made clear that opportunistic agents inevitably engender equilibria that are inefficient in that wealth is not maximized either for the agents involved or for the economy in aggregate. The equilibria are "second-best". But, regardless of their economic efficiency, are economically "rational" agents morally desirable?

Even the most cursory review of ethical theory reveals that this notion of rationality as essentially economic opportunism is not a favorable behavioral attribute. About the best moral defense that one could make is that it is harmless: it is merely a behavioral assumption that expedites finance theory and has no impact on actual behavior. But, as we've already seen, even within their theoretical constructs these agents are not harmless in that they levy a cost on the financial system. More importantly from a moral perspective, several recent empirical studies find evidence that these behavioral assumptions have influence beyond the boundaries of financial modelling. These studies find that the assumptions of financial theory influence financial practice: "is" does indeed imply "ought". For example, from his experience as a business professor, Norman Bowie supplies anecdotal evidence that exposure to this narrow rationality paradigm and related rubrics modifies behavior: "They [MBA students] believe that they will have to be unethical to keep their jobs. They believe that everyone else will put their [own] interests first" (p. 9). But he goes on to note that "the evidence here is not merely anecdotal . . . economics graduate students are more inclined to behave in a self-interested fashion" (p. 9).

For example, in one recent study involving 267 prisoners'-dilemma-type scenarios, economics students defected (i.e. failed to cooperate) 60% of the time, while non-economists defected only 30% of the time. Also, when compared to students in different disciplines, economics students were found to be less honest in hypothetical situations, and both economics students *and professors* were found to be less likely to donate to charity (Frank, Gilovich & Regan, 1993). Bowie observes that "people change their behavior when confronted with assumptions about how other people behave" (1991, p. 9). In *Challenging The Egoistic Paradigm*, he concludes that "[l]ooking out for oneself is a natural, powerful motive that needs little, if any, social reinforcement . . . Altruistic motives, even if they too are natural, are not as powerful: they need to be socially reinforced and nurtured" (p. 19). Such nurturing is not to be found in the behavioral assumptions of finance.

In a broader context, the susceptibility and suggestibility of human behavior was made very clear in the famous laboratory experiments conducted by Stanley Milgram. In these experiments volunteers were asked to administer progressively stronger electric shocks to some individual. Even though these volunteers could see the "victim" in considerable distress from the shocks, the volunteers were generally willing to administer ever higher voltages given the encouragement of an "authority figure". Milgram concluded that:

> Ordinary people, simply doing their jobs, and without any particular hostility on their part, can become agents in a terrible destructive process . . . even when the destructive effects of their work become patently clear, and they are asked to carry out actions incompatible with fundamental standards of morality, relatively few people have the resources needed to resist authority (1974, p. 6).

Similarly, Gregory Dees argues that the value systems of business theory influence those of business practice. He observes that "how concepts are introduced in an academic setting can have a significant influence on their use later on" (1992, p. 38). While commenting on the value system underlying business theory, Ronald Duska notes that "as it gets accepted as a legitimating reason for certain behavior in our form of life, it becomes subtly self-fulfilling" (1992, p. 149). Thus, the behavioral finance paradigm presents itself as morally neutral, but fails to recognize that its narrow and rigid invocation of self-interest has *moral* implications. Alasdair MacIntyre makes a similar point:

> Managers themselves and most writers about management conceive of themselves as morally neutral characters whose skills enable them to devise the most efficient means of achieving whatever end is proposed. Whether a given manager is effective or not is in the dominant view a quite different question from that of the morality of the ends which his effectiveness serves or fails to serve. Nonetheless there are strong grounds for rejecting the claim that effectiveness is a morally neutral value (1984, p. 74).

Models developed within the behavioral finance paradigm, therefore, do not merely endeavor to explain observed phenomena. They do not present merely a morally neutral perspective. To a significant degree they promote a rationality concept that lacks a defensible moral foundation.

5. CONCLUSION

Behavioral finance does not merely describe financial market reality, it shapes it. Economic rationality is taken as the ideal toward to which individuals "should" strive. In this paper I show that, as a behavioral ideal, economic rationality is unjustified both from a strictly economic perspective, and from a moral perspective.

By assuming away other motivations and thus elevating wealth-maximization to the status of a necessary law of nature, behavioral finance may be sanctioning behavior that society at large regards as, at best, morally questionable, and,

at worst, strictly immoral. In the corporate milieu, by assuming unbridled self-interest, behavioral finance promotes unbridled self-interest. Furthermore, even if empirical evidence were to overwhelmingly support wealth maximization as the dominant motivation among contemporary economic agents (which, as we've just seen, it does not), behavioral finance's normative dimension would still obligate it to consider alternatives.

REFERENCES

Copeland, T., & Weston, F. (1988). *Financial theory and corporate policy* (3rd ed.). Addison-Wesley.

Dees, G. J. (1982). Principals, agents, and ethics. In: N. E. Bowie & R. E. Freeman (Eds), *Ethics and Agency Theory*. New York: Oxford University Press.

Diamond, D. W. (1989). Reputation acquisition in debt markets. *Journal of Political Economy*, *97*, 828–861.

Duska, R. F. (1992). Why be a loyal agent? A systematic ethical analysis. In: N. E. Bowie & R. E. Freeman (Eds), *Ethics and Agency Theory*. New York: Oxford University Press.

Frank, R., Gilovich, T., & Regan, D. (1993). Does studying economics inhibit co-operation? *Journal of Economic Perspectives* (Spring).

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305–360.

Kreps, D. (1984). Corporate culture and economic theory. Stanford University, Working Paper.

MacIntyre, A. (1984). *After virtue* (2nd ed.). Notre Dame: University of Notre Dame Press.

MacIntyre, A. (1988). *Whose justice? Which rationality?* Notre Dame: University of Notre Dame Press.

Milgram, S. (1974). *Obedience to authority*.

Moore, J. (1991). Autonomy and the legitimacy of the liberal arts. In: R. E. Freeman (Ed.), *Business Ethics: The State of The Art* (p. 63). New York: Oxford University Press.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, *5*, 147–175.

Noreen, E. (1988). The economics of ethics: A new perspective on agency theory. *Accounting Organizations and Society*, *13*, 359–369.

Spence, A. M. (1976). Job market signalling. *Quarterly Journal of Economics*, *87*(3), 355–374.

Weber, M. V. (1948). *The protestant work ethic and the spirit of capitalism*. New York: Scribner's, (quote in text p. 182).