ENVY, COMPARISON COSTS, AND THE ECONOMIC THEORY OF THE FIRM

JACK A. NICKERSON* and TODD R. ZENGER
Olin Business School, Washington University, St. Louis, Missouri, U.S.A.

An economic theory of the firm must explain both when firms supplant markets and when markets supplant firms. While theories of when markets fail are well developed, the extant literature provides a less than adequate explanation of why and when hierarchies fail and of actions managers take to mitigate such failure. In this article, we seek to develop a more complete theory of the firm by theorizing about the causes and consequences of organizational failure. Our theory focuses on the concept of social comparison costs that arise through social comparison processes and envy. While transaction costs in the market provide an impetus to move activities inside the boundaries of the firm, we argue that envy and resulting social comparison costs motivate moving activities outside the boundary of the firm. More specifically, our theory provides an explanation for ‘managerial’ diseconomies of both scale and scope—arguments that are independent from traditional measurement, rent seeking, and competency arguments—that provides new insights into the theory of the firm. In our theory, hierarchies fail as they expand in scale because social comparison costs imposed on firms escalate and hinder the capacity of managers to optimally structure incentives and production. Further, hierarchy fails as a firm expands in scope for the simple reason that the costs of differentially structuring compensation within the firm to match the increasing diversity of activities also rises with increasing scope. In addition, we explore how social comparison costs influence the design of the firm through selection of production technologies and compensation structures within the firm. Copyright © 2008 John Wiley & Sons, Ltd.

INTRODUCTION

The theory of the firm has primarily focused on the question of firm boundaries, articulating when firms supplant markets and when markets supplant firms. Coase (1937) originally articulated the logic that boundary choices depend on a comparative assessment of the costs of internally and externally managing activities. When market governance is costly or markets fail, integration is likely; when internal governance is costly or organizations fail, market governance is likely. Current theory provides well-developed explanations for the failure of markets. Markets commonly fail when exchange demands highly specialized assets, involves measurement difficulty, or requires the transfer of knowledge (Williamson, 1975; Klein, Crawford, and Alchian, 1978; Barzel, 1982). However, as both Coase (1937) and Williamson (1985) note, a theory of market failure alone is not a theory of the firm, because such theory fails to resolve a basic puzzle: If firms are advantaged in managing these complexities in exchange, why is not all
production carried on by one big firm?’ (Coase, 1937: 394). What precludes the firm from selectively using the control features of hierarchy and otherwise replicating market incentives within its boundaries? If firms can replicate market incentives and utilize authority to control exchange only when needed, the firm should face no boundary limits. Indeed, absent impediments to replicating market incentives within the firm, the boundaries of the firm become rather irrelevant and managers enjoy enormous flexibility in the design and structure of organizations.

Existing theory provides little in the way of explanation for the limits of the firm or the causes of organizational failure. The prevalent view in early writings on industrial organization was that limits to firm size were due to ‘diminishing returns to management’ (Sraffa, 1926; Kaldor, 1934). Implicitly, Coase (1937) adopted a similar argument, assuming that at some point, integrating additional transactions becomes more costly than managing them externally. However, Coase provided no clear explanation as to the origin of these costs or why so-called management costs should rise with firm scale and scope. Indeed, Coase has commented that the question of why the costs of internal organization increase with firm scale and scope remains unanswered (Coase, 1988). Our goal is to develop a theory of organizational failure, more precisely a theory of organizational diseconomies of scale and scope that is rooted in processes of social comparison as discussed in the sociology and social psychology literatures. These processes of social comparison give rise to what we label ‘social comparison costs’ (Zenger, 1992; 1994). Our theory argues that while transaction costs in the market prompt access to authority through integration, social comparison costs prompt managers to limit the degree of integration and otherwise take costly organizational actions to restrict and efficiently manage these social comparison costs.

We are, of course, not the first to explore the question of organizational failure (see Hennart, 1994). Some have argued measurement difficulties escalate with firm size and limit the capacity of large firms to offer the high-powered incentives prevalent in markets (Barzel, 1982; Holmstrom, 1989; Rasmusen and Zenger, 1990; Zenger, 1992; Williamson, 1985). But, this argument alone does not explain why a large firm can’t replicate the measurement precision of markets by creating highly autonomous internal units and then structuring market-like incentives within its boundaries. Similarly, this argument does not explain why a firm cannot integrate an independent firm or activity and craft incentives that fully replicate the market incentives that previously existed.

Influence activities within organizations, as discussed by Milgrom and Roberts (1988; 1990a; 1990b), provide a partial explanation. When an activity is integrated within the boundaries of the firm, it is placed in an arena in which incentives now exist for both those managing the integrated activity and those positioned elsewhere in the firm to politically influence the distribution of rewards allocated to those managing it. They argue that such rent-seeking activity is wasteful and fully absent when the activity is autonomously controlled by the market. Absent integration, there is simply no one to politic and therefore no influence activities. While these influence activities undoubtedly play an important role in shaping the boundaries of firms, they provide at best a partial explanation. The theory does not explain why these organizational costs should increase with size or scope. Thus, while Milgrom and Roberts (1988, 1990a, 1990b, 1992) explain why selective intervention is costly within firms, they do not explain why the marginal transaction or activity becomes increasingly costly to integrate, as the firm increases in scale and scope. Since Milgrom and Roberts’ (1988, 1990a, 1990b, 1992) argument is invariant to size and scope, the theory fails to explain the limits to firm size and scope.

In this study, we develop a theory of managerial diseconomies of scale and scope, which are organizational failures that constrain the boundaries of the firm. We begin by assuming that market failures of the type commonly described in the transaction cost economics literature provide the impetus to integrate activities within the boundaries of the firm. Cognizant of these transaction costs in the market that prompt the extension of firm boundaries, we develop a theory of countervailing social comparison costs that impel managers to restrict these boundaries or otherwise control these costs. Thus, while market failures create a type of centripetal force for moving activities out of the market and into a firm, our theory explains when and how organizational failures create a centrifugal force for moving activities out of the firm and into the market.
As a brief preview, our theory begins by examining the behaviors of both individual employees and managers in response to the infusion of market-like incentives within the firm. Consistent with a wide range of social science literature, we assert that individual employees invidiously compare their rewards with others they deem to be within their referent group (for example, see Adams, 1963; Festinger, 1954). If perceived inequity arises, the resulting negative feeling—what we refer to as envious emotion—drives individuals to expend effort to ameliorate these perceptions of inequity. Such behaviors include reduced effort, influence activities, departure, noncooperativeness, or even outright sabotage. Such behaviors impose substantial costs on the firm—costs that we reference as social comparison costs.

Managers, of course, are not passive actors in responding to these comparison costs. Managers anticipate social comparison costs or at least perceive them when they arise, and take active steps to attenuate them. Management, we argue, has three structural levers at its disposal to attenuate social comparison costs: ‘compressing’ rewards by decoupling pay and performance, shifting ‘production technology’ to reshape social comparison and social referents, and redrawing the boundaries of the firm. Our theory argues that each of these choices potentially imposes significant costs on the firm. Therefore, the probability that managers will select any one of them depends on the magnitude of social comparison costs associated with each. As social comparison costs rise, managers are more likely to respond with damped incentives, compromised production technology, or restricted boundaries to the firm. We argue that the scale and scope of the firm are primary drivers of these social comparison costs. Hence, as scale and scope increase, the need to attenuate social comparison costs through incentive dampening, production efficiency compromises, or boundary restriction increases. The result is a theory of organizational failure in which organizational costs rise with increasing scale and scope of the firm. We believe our theory provides an explanation for managerial diseconomies of scale and scope—the result is a theory of organizational failure in which organizational costs rise with increasing scale and scope of the firm. We believe our theory provides an explanation for managerial diseconomies of scale and scope—a theory that is independent from traditional measurement, rent seeking, and competency arguments—that contributes to determining firm boundaries.

In the pages that follow, we use a series of illustrations to highlight the social comparison costs that arise as managers attempt to selectively infuse market incentives within the firm. We then discuss the structure of these costs and identify and describe three levers, or organizational design choices, available to management to attenuate them. We use two levels of analysis—employee decisions about how to attenuate envious emotions and managerial policy decisions about the structure of the firm—to develop an argument for how social comparison generates diseconomies of scale and scope in firms. We then discuss the implications of our theory along with a few caveats and conclude.

Selective intervention and social comparison costs

Our theory is designed to complement, not replace, the logic of transaction cost economics. In transaction cost economics, managers choose markets over firms in order to access the superior incentives that markets provide (Williamson, 1991). In other words, managers forego integration and thus limit the size and scope of the firm in order to access the high-powered incentives of the market (Williamson, 1985). Consequently, as Williamson argues, if managers could replicate the incentives of the market within the firm, the firm would face no limits to its scale and scope, and hence no limits to its boundaries. If firms could truly complement the virtues of internal organization with the incentives of the market, markets need never arise and the boundaries of the firm face no limits. Therefore, to understand the limits to the scale and scope of firms, it is useful to first explore why firms fail as managers attempt to internally replicate market incentives.

We are not the first to observe the difficulty that firms face in selectively replicating market incentives for subunits or activities within the firm (Williamson, 1985; Zenger and Hesterly, 1997). Indeed, there is empirical literature that confirms the greater difficulty that large firms confront in structuring market incentives within the firm (Garen, 1985; Zenger and Marshall, 2000). Furthermore, the literature is replete with examples of failed attempts at such ‘selective intervention.’ By examining a few notable illustrations of these failed efforts, we hope to highlight the genesis of social comparison costs and provide a backdrop for our theory.
Illustration 1

Harvard’s efforts to manage internally their endowment portfolio provide a fascinating example of social comparison costs. Until 2005, Harvard Management Company was a wholly owned subsidiary of Harvard University that managed Harvard’s $27 billion endowment portfolio. In an effort to mimic the rewards offered to fund managers of privately managed hedge funds, the multiple fund managers within the Harvard subsidiary received compensation based purely on a formula that linked bonus compensation to their fund’s performance relative to a benchmark fund of comparable risk and classification (see Hall and Lim, 2003). These internally managed funds quite consistently outperformed the benchmark funds by a very wide margin, which resulted in extremely high compensation for these Harvard employees. In fiscal year 2004, the top two fund managers received $25 million each. In fiscal year 2003, the top fund managers received $35 million each. In 2001 and 2002 large sums were also paid out to fund managers due to exceptional relative performance, even though the funds themselves declined in value. In 2004, opposition from students, alumni, and faculty to such pay reached a feverish pitch (Harvard Magazine, 2004). Key alumni called for a university-wide forum to review the matter and threatened to withhold donations. Larry Summers, Harvard’s President, claimed that outsourcing the activity would lead to lower performance and higher fees, while Ronald Daniel, the University Treasurer, defended the practice by noting that other universities which use ‘external investment managers...do not face comparable scrutiny’ (Williamson, 1985: 73). In early 2005, Harvard Management Company significantly restricted the maximum levels of compensation within the subsidiary. By March of 2005, many of the key fund managers had departed, taking with them portions of Harvard’s endowment to manage externally.

Illustration 2

Strauss’ (1955) describes a factory that manufactured wooden toys that sought to enhance the productivity of a team of workers in the paint department through a bonus pay system based on team output. As a result of the market-like pay plan, productivity of this team increased 30 to 50 percent above what was traditionally expected. These employees consequently earned a significant premium above what other more skilled workers earned in other parts of the plant. Because of this differential, ‘[m]anagement was besieged by demands that the inequity be taken care of’ (Strauss, 1955: 94). Ultimately, management returned the painting operation to its original (and less productive) pay structure, which quickly led to the voluntary departure of workers on the team, as well as the departure of their supervisor.

Illustration 3

In 1980 Tenneco Inc., acquired a relatively small company, Houston Oil and Minerals Corporation (HOMC) (see Williamson, 1985: 158; and Milgrom and Roberts, 1992: 194). To encourage the retention of HOMC’s exploration talent, Tenneco offered special salaries, bonuses, and benefits to HOMC employees—payments that were not offered to others at Tenneco. ‘Tenneco also agreed to keep [HOMC] intact and operate it as an independent subsidiary rather than consolidate the acquisition.... Despite initial enthusiasm, [HOMC’s] managers and its geologists, geophysicists, engineers, and landmen left in droves during the ensuing year’ (Williamson, 1985: 158). The implementation of the customized compensation package had been delayed, because, as Tenneco’s vice president for administration maintained, ‘We have to ensure internal equity and apply the same standard of compensation to everyone...’ (Williamson, 1985: 158).

Illustration 4

Recently, MBA students at a North American business school, which will remain anonymous, complained about the lack of access to the more ‘visible and famous’ professors at the school. Seeking to satisfy student demand while recognizing that these same professors were in high demand for consulting and speaking engagements, the dean proposed individually negotiating overload teaching payments to faculty. This approach would enable the payment of ‘market rates’ for internal teaching services. However, discussion with faculty quickly led to concerns over fairness, which could be resolved only if there was a set rate for all faculty to do overload teaching. It is interesting to note that no such fairness concerns were voiced.
about high speaking and consulting fees generated from outside the university. Negotiation with individual faculty on overload remuneration did not proceed.

Illustration 5

At International Harvester in the 1960s, the union requested that management increase the speed of particular assembly lines where employees were rewarded in part based on output. Despite clear benefits to the firm, management refused, concerned that higher pay for this group would promote dissension elsewhere among those not benefiting from the line speed increase (reported in McKersie, 1967: 222).

Illustration 6

In the early 1970s, Syntex and Varian Corporation established a joint venture to commercialize jointly developed technology. The intent was to create an entrepreneurial unit free from the bureaucracy of the larger parent companies. Consistent with the desire to create the feel of a truly new start-up, the head of this new venture was given a compensation package characteristic of a small, entrepreneurial firm. He was granted an equity stake and assigned a relatively small salary. As soon as the details of this package were revealed to senior executives at Syntex, a major internal uproar ensued. Long before any value from the joint venture could be realized, the compensation package of the head of the joint venture was realigned so as to be consistent with others at a similar level within Syntex (former vice president of Human Resources to second author, personal communication).

Overview

There are several common themes revealed in these illustrations. In each, managers within the organization sought to develop a high-powered, market-like incentive mechanism, custom-tailored to a specific group or activity. In each case where the incentive plan was in place long enough to measure performance, the performance results for those engaged in the activity were quite positive. However, due to this success or anticipated success, the plans produced wide variance in compensation and, in particular, some very high compensation levels that others in the organization perceived as inequitable. The costs imposed by the reactions (or anticipated reactions) of these other employees prompted the plans’ cessations. Note that in each case management was forced to make a trade-off between the incentive benefits of those affected by the plan, in terms of higher effort or the attraction of superior talent, and comparison costs imposed by those not attached to the plan. In each case the social comparison costs were determined to outweigh the incentive benefits.

These illustrations suggest that organizations may fail relative to markets because social comparison costs limit the flexibility of managers in offering the ‘optimal’ market-like incentives for specific activities within the firm. Below, we illuminate more completely the mechanisms that underlie this argument. Moreover, we argue that as a firm’s scale and activity scope increase, comparison costs also increase and with them the manager’s inflexibility in tailoring incentives for different internal activities. These rising costs and inflexibility amount to organizational failures that constrain the boundaries of the firm.

Social comparison and human behavior

Our theory begins with the assumption that individuals engage in a process of socially comparing their rewards to those received by salient referents. Thus, decisions about how rewards are allocated define in part the nature of these social comparisons. Research on social comparison processes and their effect on individuals are legion across the social sciences (Homans, 1961; Festinger, 1954; Adams, 1963; Martin, 1981). Many theories indicate that people engage in comparison activities, focusing especially on relative income and the basis by which relative income is determined. Moreover, these comparisons tend to be invidious in the sense that individuals identify salient referents as those who are economically better off. These theories assert that people care about inequity; they react negatively to outcomes they deem unfair; and they exert efforts to reduce negative feelings proportionate with the inequity they perceive (Adams, 1963). This propensity to compare income to salient referents and the willingness to expend effort to reduce perceived inequity in relative income, we assert, constitutes a fundamental behavioral assumption about human nature upon which our theory rests.

Modern research on social comparison processes begins with Festinger (1954), who assumed that
humans have an innate propensity to self-evaluate ‘based on comparison with other persons’ (1954: 138). Homans’s (1961) study constructed similar arguments, emphasizing that individuals particularly compare income. 1 Equity theory as developed by Adams (1963) posits that ‘[i]nequity exists for [an individual] whenever his perceived job inputs and/or outcomes [typically measured by actual or expected income] stand psychologically in an obverse relation to what he perceives are the inputs and/or outcomes of [others]’ (Adams, 1963: 424). Relative deprivation theory similarly concludes that feelings of deprivation arise from comparisons of rewards to some referent (Martin, 1981). It is important to note that perceptions of inequity are not symmetric. That is, individuals commonly have inflated perceptions of personal contributions or performance (Meyer, 1975; Zenger, 1994). The problem of establishing equity perceptions is further exacerbated by individuals’ propensity to compare pay only to those perceived as comparable in performance, but earning more (Martin, 1981). This combination of upwardly biased self-assessment and a propensity to compare income to those earning more nearly always ensures at least some level of inequity perception. Thus, varying perspectives on social comparison and inequity perceptions are consistent with the notion that individuals are envious of salient referents perceived to receive superior income relative to their contributions. 2

Social comparisons that lead to inequity perceptions in individuals create a willingness to expend effort to attenuate these feelings of inequity. Such efforts increase the greater is the perceived inequity. For instance, Adams (1963: 427) argued that ‘[t]he presence of inequity in [an individual] creates tension in him ... proportional to the magnitude of inequity present’. The presence of inequity ‘will drive him to reduce it.’ As with equity theory, feelings of relative deprivation prompt behavioral responses to alleviate or avoid such feelings (Martin, 1981). Generally speaking, the literature argues that those who perceive inequity relative to their social referents expend effort to eliminate this tension by restoring equity (see Walster, Berscheid, and Walster, 1973; Adams and Freedman, 1976; Ma and Nickerson, 2007; for empirical reviews see Greenberg, 1982; Suls and Williams, 1991).

An important issue underlying this research is identifying which referents are salient. In other words, with whom do individuals socially compare themselves? The general conclusion in the literature is that spatial proximity, degree of interaction, and availability of information are primary determinants of the choice of salient referents (e.g., Festinger, 1954; Kulik and Ambrose, 1992; Williams, 1975). Spatial proximity means not only propinquity, but also a variety of other demographic measures of social distance such as age, tenure, education, gender, and the like. As with spatial proximity, the degree of interaction and the availability of information may be endogenous to firm boundaries and other organizational and production technology choices. These conclusions are reminiscent of Aristotle’s observation (Rhetoric, 1388, quoted in Goel and Thakor, 2003): ‘We envy those who are near us in time, place, age, or reputation.’ We return to these issues below.

The literature suggests that individuals pursue several different behaviors or strategies to reduce envious emotions that accompany perceived inequity. First, individuals alter their own behavior. In particular, they reduce effort, bringing contributions more in line with rewards (Adams,

---

1 Homans’s classic on social behavior (1961), which builds on Blau’s (1955) social exchange theory, identifies an equivalent propensity. Homans claims that individuals engage in exchanges, whether they are economic or social, because they receive some benefit in excess of their costs (Homans, 1961: 72). Further, individuals ‘perceive and appraise their rewards, costs, and investments in relations to the rewards, costs, and investments of other men’ (Homans, 1961: 76). Homans’s key proposition in this regard is that ‘the more to a man’s disadvantage the rule of distributive justice fails of realization, the more likely he is to display the emotional behavior we call anger’ (Homans, 1961: 75). This proposition equates to individuals caring not only about the absolute level of their income but also about relative income. Moreover, Homans acknowledges that such anger translates into action because individuals ‘learn to do something about it’ (Homans, 1961: 77).

2 At least since Aristotle, philosophers have recognized envy and jealousy as fundamental propensities of human nature. Envy is the emotion that arises when one desires something currently possessed by another (Salovey, 1991). Envy is one of the seven deadly sins (Silver and Sabini, 1978) and perhaps the most pervasive but under-acknowledged one (Epstein, 2003). Arguably, it is envy and jealousy that generate feelings from social comparison that give rise to actions to reduce such feelings because people systematically care more about others ahead of them compared to those behind them. For instance, while Festinger (1954) did not delve into the emotional underpinnings of social comparison theory, Salovey (1991) argues that envy and jealousy are the specific emotions that accompany such appraisals. Indeed, invidious comparison, envy, jealousy, or some combination is mentioned as the implicit or explicit motivation for behavior in all three of the psychological perspectives mentioned above.
Thus, the organizing challenge of the manager is to economize on social comparison costs while simultaneously providing effective incentives and efficient production technology.

Economizing on social comparison costs

While transactions cost economics contends that the manager’s job is to economize on transaction costs, our theory suggests that the manager’s job is to (also) economize on social comparison costs. We maintain that managers have at their disposal three structural or organizational levers with which to attenuate social comparison costs: wage compression, the choice of production technology, and the choice of firm boundaries. All contribute to explaining organizational diseconomies of scale. Below, we not only highlight how these levers economize on social comparison costs, but also examine how each choice imposes its own costs on the organization. After discussing these levers, we consider their relationship to managerial diseconomies of scale and scope.

Wage compression

Scholars from a range of disciplines suggest that weakening the link between pay and performance is a common response to social comparison processes (Akerlof and Yellen, 1990; Frank, 1985; Konrad and Pfeffer, 1990; Zenger, 1992). By taking steps to reduce differences in income among workers, managers reduce the impetus for workers to envy each other’s income. Indeed, by adopting uniform compensation, managers essentially avoid issues of distributive justice. When all pay

3 In addition to increasing or reducing effort to bring contributions more in line with rewards, workers might also (over) invest in capabilities through training or education based on the expectation of future rewards. To the extent that workers choose increasing effort and investing in capabilities to reduce envious emotions, these behavioral responses can benefit the firm. We set aside these behaviors that might benefit the firm while developing our theory and return to them in the discussion.

4 Even when compensation is uniform and distributive concerns are rendered rather moot, procedural justice issues may still arise. Distributive justice focuses on perceived fairness of the distribution of rewards (Homans, 1961) whereas procedural justice focuses on perceived fairness of the process by which decisions are reached (Greenberg, 1990). We do not incorporate procedural justice concerns into our argument for three reasons. First, empirical research indicates that distributive justice contributes to variance of satisfaction with pay more than twice that of procedural justice (Tyler, Rasinski, and McGraw, 1985; Folger and Konovsky, 1989), which suggests that distributive justice may be far more important than procedural justice. Second, Long, Bendersky, and Morrill (2007) find that when tasks are separate and pay contingent, workers principally care about distribution of rewards—a perspective consistent with our view. They also find that when tasks are interdependent and wages are not contingent (i.e., wages are compressed), workers may then care more about procedural justice. Their finding suggests that distributive justice is a more important driver of structural responses by managers.
is equal, perceptions of inequity must stem from differing perceptions of performance rather than pay. A small but growing economics literature has incorporated envy and social comparison into formal models. These economists acknowledge that workers care about relative income, which creates social forces that affect the wage profiles within and between firms. In an early comment, Hicks (1955: 390) states that 'economic forces do affect wages, but only when they are strong enough to overcome these social forces.' Economists acknowledge that disparities in compensation can ‘induce discontent among employees. . . . followed by uncooperative and unaccommodating behavior’ (Pencavel, 1977: 239). Thus, economists theoretically and empirically note that managers can and do attenuate social comparison costs by flattening the wage structure compared to the marginal product schedule. In doing so, the firm elevates pay beyond the marginal product for some and constrains pay below the marginal product for others, which may precipitate departure—a phenomenon called wage compression. These economic arguments clearly resonate with those of equity theory, relative deprivation, and social exchange theory discussed above.

It is useful to distinguish between two types of wage compression. Horizontal wage compression refers to the adoption of rather flattened or more uniform compensation for individuals in equivalent or identical jobs. Thus, as the costs imposed on firms by social comparison processes increase, firms adopt more horizontally compressed or more uniform income for each position (e.g., Hicks, 1955; Addison and Burton, 1981; Hamermesh, 1975; Dunlop, 1947; Pencavel, 1977). Indeed, it is well documented in labor economics that workers undertaking the same job or occupying positions within the same grade typically have minor differences in relative income even though their marginal products may vary greatly. This uniformity of income—income that is invariant with performance—is a decision made by management.

A second type of wage compression—vertical wage compression—arises among workers holding different positions, or performing completely different jobs. Vertical wage compression involves reducing the variation in income across differing positions or jobs, despite potentially widely varying marginal products associated with these jobs and positions. Addressing the vertical wage compression phenomenon, Frank (1984a; 1984b; 1985) offered a formal theory how invidious comparisons and variation in utility for status lead to vertical wage compression in firms. Akerlof and Yellen (1990) in their fair wage hypothesis argue that vertical wage compression is a best response by management to counter reduced effort, which occurs as a response to envy among individuals with different skills. These and other scholars argue that preferences for perceived inequity can cause managers to vertically compress wages by narrowing income dispersion across levels and jobs (e.g., Frank, 1984a; 1984b; Akerlof and Yellen, 1990; Hicks, 1955; Addison and Burton, 1981; Hamermesh, 1975; Dunlop, 1947; Pencavel, 1977; Zenger, 1992).

Our prediction is that as social comparison costs increase, managers are more likely to engage in wage compression both vertically and horizontally. Such wage compression offers an economic lever to ameliorate social comparison costs by reducing the stimulus of envy. Its application, however, necessarily implies opportunity losses as weak incentives discourage high effort and trigger a pattern of adverse selection among workers in which the highly productive but underpaid depart, while the less productive but overpaid remain.
Technology choice

A second economic lever at management’s disposal in response to social comparison costs relates to shaping a firm’s production technology, broadly defined. From a classical microeconomic perspective, for any given desired output there is an optimal production technology that defines the configuration of assets and the structure of individual and collective tasks and activities within the firm. In addition to this ‘technical’ efficiency, the choice of production technology also defines the degree to which output is team produced, the spatial proximity of workers, and the degree to which workers interact. Thus, managers shape both production costs and social architecture when choosing a production technology.

The social architecture defined by a production technology is important because it shapes social comparison processes, which can impose costs on the organization in two ways. First, social architecture affects the cost of accessing information about peers’ performance and productivity, which molds processes of social comparison. For instance, workers located in close proximity and performing similar tasks are likely to identify each other as salient referents and compare income, which, if different, can lead to social comparison costs. By increasing the physical distance among workers, management can restrict the scope of interaction and information sharing, thereby reducing the salience of these workers as referents. Thus, management can reduce social comparison costs by increasing the physical and informational distance between jobs, which, however, departs from the technically efficient production technology.

Second, the choice of production technology, specifically the extent of team production, shapes social comparison costs by defining the precision with which individual performance is measured. In the absence of team production, workers and managers can clearly observe and verify relative individual contributions (Alchian and Demsetz, 1972). While variance in worker income may stimulate envy, the absence of team production limits the efficacy of behavioral strategies that impose comparison costs on the firm. For instance, in the absence of team production, shirking imposes large costs on the worker in terms of lost productivity and income and imposes rather modest costs in lost productivity for the firm. The verifiability of individual performance causes influence activities that appeal to management to raise pay or lower others’ pay to be credibly met with the response that the worker, not management, is responsible for earning less. Noncooperative behavior and sabotage are less feasible strategies because they lead to a reduction in the worker’s own productivity and income. Furthermore, independent jobs reduce the opportunities for such behavior. By contrast, the envy promoted by pay variance with independent jobs may trigger behavioral strategies that positively affect worker effort and benefit the firm, since additional effort translates directly into higher income. Thus, social comparison processes when jobs are independent may lead to social comparison benefits instead of costs.

By contrast, social comparison costs are likely quite high when managers in a team production environment attempt to link income to individual performance. As the scope of team production increases, accurately observing and agreeing on relative contribution becomes more difficult for workers and managers alike. In a team production setting, individual contributions are of necessity subjectively determined. Consequently, if managers assign pay based on these subjective evaluations, this engenders influence activities because managerial subjectivity can be blamed for differences in pay. Additionally, individuals without access to verifiable information rather easily develop inflated perceptions of their marginal contributions (Meyer, 1975; Zenger 1994), which may amplify willingness to engage in influence activities. Moreover, team production makes feasible other behavioral strategies like noncooperative behavior and sabotage as well as shirking that are not effective in the absence of team production. Coincidentally, because of measurement difficulties, team production also reduces the effectiveness of behavioral strategies like increasing effort as a means of raising income.

The underlying logic here is that in team production settings the social comparison costs that accompany efforts to link pay and individual performance typically overwhelm the benefits.
Consequently, managers increasingly compress wages to reduce envy as the degree of team production rises. Thus, Leventhal (1976) and Deutsch (1985) find that individuals prefer equal rewards rather than performance-based rewards when they anticipate collaborating on tasks. Similarly, Frank (1985) and Konrad and Pfeffer (1990) argue that production technologies that encourage collaboration hinder pay based on individual contributions.

Choosing an inefficient production technology for the purpose of reducing social comparison costs obviously imposes its own set of costs. For instance, in our illustration of International Harvester and the toy factory, management purposefully chose a technically inefficient assembly line speed in order to avoid income dispersion that would generate social comparison costs. Attenuating envy, while attempting to reward performance, often calls for a departure from the technologically efficient choice of production technology; it often demands spatially separating workers and limiting information flows to reduce social comparison. Thus, to reduce social comparison costs to a level sufficient to allow aggressive pay for performance, managers may need to alter production technology to reduce collaboration, increase physical distance among employees, or actively contain information about productivity and income. Thus, our prediction is that as social comparison costs rise, managers are more likely to compromise the efficiency of production technology to mitigate these costs.

Constraining firm boundaries

The final and most dramatic structural means to constrain social comparison costs is to limit the boundary of the firm. While social comparison occurs both within and across firms, the scope of such comparison and the corresponding costs imposed play out very differently within firms than across firms. These differences allow management to use the boundary of the firm as a means of influencing social comparison costs. The shift in social comparison costs at the firm’s boundary occurs for two reasons. Either reason is theoretically sufficient to generate our boundary prediction. First, within the boundaries of the firm, the presence of the central manager who assigns compensation changes the cost-benefit analysis that individuals face in determining their various individual behavioral strategies for reducing feelings of envy. We argue that the behavioral responses to social comparisons that occur within a firm are more costly than the behavioral responses to social comparisons that occur across firm boundaries. Second, the firm’s boundary may define an individual’s salient reference group, which implies that altering firm boundaries alters the scope of salient referents, and thus the scope of social comparison and its accompanying costs. We describe both mechanisms below.

Our first contention is that behavior stemming from invidious comparison manifests differently within the firm than across a market interface. Within the firm, the presence of a central manager shifts the costs and benefits of various behaviors for reducing envious feelings. Consequently, the comparison costs imposed by these behaviors accrue differently within firms than across markets. When comparison occurs within the firm, the costs of both reduced effort and intense lobbying to restrict others’ pay are felt very directly by the firm. By contrast, when a market interface separates focal individuals from those with whom they compare themselves, the comparison costs that these focal individuals can impose are minimal. Across this market interface, there is no central authority who commonly assigns rewards. Consequently, there is very little that those who perceive inequity across this market interface can do in response. Certainly, the external manager is unlikely to entertain pleas for cross boundary pay equity from those employed with an outside firm.

To further illuminate this argument, consider two actors, Gen and Will. Assume they are engaged in similar tasks, that Will views Gen as a salient referent, and that Gen receives a higher income, which is known to Will. Because relative income is transparent to both parties, Will envies Gen and is motivated to take action to reduce this perceived inequity. The critical question is how Will’s actions to reduce feelings of envy differ when Will and Gen are in the same versus different firms.

When Gen and Will are employed by the same firm, Will can engage in influence activities that affect Gen’s income or behavior. Will can lobby the central manager to reduce Gen’s pay or to change the nature of her job and thereby indirectly reduce her pay. Moreover, he can affect her behavior by engaging in subtle and perhaps not-so-subtle forms of retribution and sabotage. Within the firm, Will is likely to have close access to
Gen’s work location and possess information useful to Gen, which could allow him to interfere with Gen’s tasks. Such efforts in the anthropology literature (e.g., Knauf 1991) are sometimes referred to as leveling and are exemplified in all of our illustrations. While policies against retribution or sabotage may limit such behavior, the monitoring required to reduce these behaviors or to monitor effort are clearly costly themselves.

In the case that Gen and Will are employed by different firms, Will can take actions to affect his own behavior such as working harder or investing in his ability if it would increase income, engaging in influence activities to increase his own income, or reducing his effort. Alternatively, Will can depart from the referent group. Will has little opportunity to affect Gen’s behavior or more importantly Gen’s income because, as Milgrom and Roberts (1988, 1990a) argue, influence activities are not effective across firms. Will might consider noncooperation and sabotage outside of the firm; but, engaging in these activities may lead to intervention by civil authorities or social approbation.7

While the costs and benefits of these alternative responses depend on particular circumstances or context, it is nonetheless clear that the costs and benefits of the alternative actions Will can take to reduce feelings of envy shift when salient referents are within the firm’s boundaries rather than across a market interface. The primary effect of the boundary of the firm in this context is to define the ways in which Will can affect Gen’s behavior and rewards. This effect may also come in a more indirect manner. Gen, knowing that Will can successfully alter her output and pay even if she works hard, may simply shirk in anticipation of such behavioral responses by Will. Thus, envied workers may find it in their best interest to share output or shirk to attenuate envy, which is a well-known and common response (Foster, 1972; Elster, 1991; Mui, 1995). Such behavior introduces efficiency loss because it affects the envied worker’s incentive for high effort.

The boundaries of the firm not only influence the costs imposed when rewards are compared, but these boundaries also define patterns of comparison. By defining these comparison patterns, firm boundaries further shape the magnitude of social comparison costs. Individuals within a firm seldom view those outside the firm as salient referents and vice versa. Individuals outside the firm are unlikely to impose comparison costs on the focal firm, in part because those inside the firm are less likely to compare compensation with those outside. A variety of arguments may explain the greater intensity of internal comparison. Closer physical proximity or more intense social interaction within the firm provides a partial explanation, though many who are not employees of the firm may still have close physical proximity and extensive social interaction with those within the firm. Organization scholars argue that firm boundaries shape individuals’ sense of identity (e.g., Kogut and Zander, 1992). Because individuals ‘identify’ with their employer, they compare pay to those within the organization. The hierarchical structure in which pay levels are essentially ratified or endorsed supports this identification with the firm. The hierarchical structure used to allocate rewards also creates a de facto comparison process across gaps in the organization, either geographic or structural, which would otherwise prevent direct comparison among employees. While distance or the absence of social contact may preclude direct comparisons among employees, managers compare the compensation of their subordinates with the compensation of other managers’ subordinates. Equity concerns and the capacity of managers to socially compare promote implicit comparisons among all those within the borders of the organization. As a consequence, there are strong pressures to standardize pay policies even across geographically distinct organizational units (Beer et al., 1984). Thus, our prediction is that as the social comparison costs associated with maintaining high powered incentives within an activity rise, managers are increasingly likely to outsource that activity.

In sum, our arguments suggest that managers face a fundamental trade-off between enduring
the comparison costs that accompany high powered rewards, and enduring the productivity losses that accompany compressed compensation structures, suboptimal production or task structures, and compromised boundary decisions. Managers must balance the social comparison costs that arise from envy with the costs imposed by horizontally and vertically compressing wages, adopting inefficient production technology, and reconfiguring organizational boundaries. We contend that management’s best response to expected social comparison costs is often to use the structural levers at its disposal to reduce the stimuli of envy, by narrowing the set of salient referents, or shifting the behavioral strategies individuals’ employ in their effort to reduce envious emotions. Figure 1 describes the basic relationships within our theory. Management selects the firm’s production technology and social architecture, its compensation and incentive structure, and its boundaries. These choices not only create an incentive structure and production technology that employees operate within for producing output, but also create a particular pattern of social comparison within and without the firm. Individuals respond to this pattern of social comparison and impose social comparison costs. Management expecting these social comparison costs or updating its selection chooses production technology, firm boundaries, and incentives structures in a way that balances the costs imposed by compromising choices of organizational design and boundaries with social comparison costs.

Comparison costs and organizational diseconomies of scale and scope

Our arguments, thus far, highlight the role that the choice of incentives, production technology, and boundary decisions play in shaping comparison costs. Our theory suggests that for any given activity a manager examines the comparative governance costs, both transaction and comparison costs, of several alternatives: (1) governance through the market with high powered incentives, (2) governance within the firm with high powered incentives, though perhaps with a production technology that reduces social comparison, and (3) governance within the firm, but with rather weak, compressed incentives. Nothing to this point, however, highlights how the selection among these choices is influenced by the scale or scope of the existing firm. Yet, empirically there is evidence that scale and scope do matter in selecting among these alternatives. For instance, empirically, large firms are more inclined to compress wages than small firms (Garen, 1985; Zenger, 1994; Rasmusen and Zenger, 1990). We argue below that as scale or scope increase, social comparison costs rise and consequently firms are more likely to respond with wage compression, production technology compromise, or a limit to the scope of the firm. Thus, decisions about how to govern the marginal activity depend on the scale and scope of existing activities. To advance this argument, we examine a simple organization that can vary along two dimensions: the scope of activities that

---

Figure 1. Balancing social comparison costs and suboptimal organizational design

it performs and the scale of (i.e., the number of) participants within these activities.

Activity scale

We begin by discussing how the scale of an activity within the firm shapes the comparison costs that accompany efforts to link individual pay and individual performance within that activity. Of course, the sensitivity of comparison costs to activity scale depends in part on the nature of the production technology. When jobs are completely separate and independent, comparison costs in linking pay to individual performance are quite limited and thus increasing the scale of the workforce introduces no real scale diseconomy. Individual performance is completely observable to all and employees through their behavioral strategies can impose rather few costs on the firm. At the same time, when managers simply forego efforts to link individual performance and pay, due to high comparison costs, and instead adopt horizontally and fully compressed (flat) wages in response to high levels of team production, then again there is no real scale diseconomy present. While there is a clear degradation in employee effort, horizontally flat wages are fully scalable within an activity, that is, comparison costs per worker do not rise when wages show no dispersion. However, our interest is in showing that when attempting to link pay to individual performance within an intermediate level of team production, activity scale increases pressures to compress wages, adopt a less efficient production technology, or even restrict the scale of an activity.

To illustrate, consider a firm that performs a single activity staffed by multiple employees engaged in collaborative production. Assume that the firm draws workers from a labor market pool in which the workers are heterogeneous in their ability to contribute. As a starting point for our logic, assume that the workers are paid what the firm estimates to be each worker’s marginal product of labor. However, because output is collaboratively produced and inputs are imprecisely measured, the assignment of individual contributions is both somewhat arbitrary and imprecise, which provides no verifiable information to counter inflated perceptions of their marginal contribution. Also assume that each worker considers all other workers within the activity (or in this case, the firm) as social referents.

Envy arises when differences in income emerge, especially among workers engaged in an activity involving team production. In our simple, single activity firm, differences in compensation arise because of differences in the manager’s assessment of each worker’s marginal product—assessments influenced both by ability differences and by measurement error. A difference in pay between two workers gives rise to perceived inequity and envy within the lower paid worker. If we assume that feelings of envy, and hence the effort the worker is willing to expend to reduce these feelings increase, the greater is the pay differential between the two workers, then, at an aggregate level, perceptions of inequity increase as the dispersion of pay increases among workers engaged in identical activities.

We contend that dispersion (the difference between the maximum and all others) in the marginal productivity of labor, and therefore the dispersion of resulting pay among those engaged in an activity, increases with the scale (or number) of individuals engaged in that activity. Under most, if not all reasonable distributional assumptions, dispersion of ability and hence dispersion in pay increases with the number of workers, because the likelihood of drawing extremely high or low ability workers from the distribution increases with the number of workers drawn (i.e., firm scale). Based on this model, the more employees that a firm has, the greater will be the dispersion of worker abilities. Because greater dispersion in ability generates focal individuals of very high ability and therefore very high pay, as scale increases so does the magnitude of envy and resulting social comparison costs. Thus, unless management takes steps to alleviate these envious feelings, the firm can expect social comparison costs from reduced effort, non-cooperativeness, influence activities, departures, sabotage, or departure to increase with scale.8

8 While there are clearly limits to the number of comparisons in which an individual can engage, the hierarchical structure of organizations generates de facto comparisons; workers compare to focal workers and managers compare pay of their subordinates with other managers’ subordinates. Our assumption is that even in the absence of direct personal contact, word of inequity travels rapidly through the hierarchy. In particular, word of the highest paid worker, what we refer to as the focal worker, travels the fastest.

9 While not a focus of our study, note that firms will often seek to actively reduce the variance in ability by screening out those at either the low or high tail of the organization. Some
Consequently, as the scale of the activity increases, the pressure to attenuate the stimulus to these rising comparison costs also increases. As the social comparison costs of rewarding the marginal productivity of labor increase, the manager becomes more likely to either horizontally compress pay, for instance, paying everyone the average marginal productivity of labor, or to compromise the technologically efficient production technology in a way that reduces interaction among workers and limits their ability to view each other as salient referents. Either choice imposes its own set of costs on the firm. A decision to compress pay imposes predictable costs of reduced worker effort and adverse selection. For instance, high ability workers are likely to exit the firm in search of firms that will pay them their marginal product of labor, while low ability workers will be particularly attracted to this compensation policy from other firms paying the marginal product of labor (see Zenger, 1992). While the manager may elevate pay to retain these high ability workers (Akerlof and Yellen, 1990), this choice has predictable costs. The manager must either overpay for the low ability employees also attracted to the firm or heavily invest in screening procedures that reduce their number.

A decision to subdivide an activity into geographically or socially isolated groups may decrease social comparison and social comparison costs, but at the expense of any scale economies that arise from team production that is compromised by this technology choice. For instance, subdividing workers into colocated groups may decrease social comparisons across the entire organization by shaping social referents and inhibiting information flow and affect individual responses to envy by decreasing opportunities to engage in non-cooperative behavior. Yet, it is not clear that such internal subdivisions are particularly effective in constraining these social comparisons. Managers also are susceptible to invidious comparisons and, according to our view, invest in actions to reduce their own feelings of envy that arise from comparisons with other managers regarding the compensation of subordinates. For instance, managers experience envy or anticipate feelings of envy from their subordinates, when they observe others managers’ subordinates earning more than their subordinates. Consequently, subdividing an activity within the firm may do little to increase the manager’s capacity to differentially reward these groups. To truly reduce comparison costs may require changing the boundaries of the firm and reducing activity scale.

Thus, as the scale of an activity increases, the social comparison costs of rewarding the marginal product of labor rise. As a consequence, the manager must choose to either weaken incentives, assigning a common level of pay to all workers, or simply compromise production technology. This compromise can take two forms: breaking the firm into smaller internal units that are somewhat socially or geographically isolated, or more dramatically curtailing the boundary of the firm itself. Thus, our prediction is that in attempting to link individual pay to performance, as the scale of an activity increases, social comparison costs rise and as a consequence compressing wages, compromising production technology, or constraining firm boundaries all become more likely.

Activity scope

We now turn our attention to how a firm’s activity scope influences social comparison costs. Our argument is that social comparison costs grow as the marginal productivity across activities becomes increasingly dispersed. To explore this argument, consider a firm that manages an array of heterogeneous activities that differ in their marginal productivity. For convenience, assume that all employees engaged in each activity receive the average marginal product of labor for the activity. In essence, we offer employees a horizontally compressed wage within activities in order to focus on the effects of variation across activities. We say a firm displays scope when it integrates into two or more activities with varying average marginal products. The scope of the firm increases the more dispersed the average marginal product of activities. Also assume that the production technology of the firm requires at least some interaction among the diverse activities, which implies that there exist efficiency benefits from colocation of activities and that the combination of this colocation and interaction of workers across the different activities ensures that the firm represents each employee’s salient reference group. Note again organizations conclude that extremely high ability individuals are simply ‘too costly’ to retain, where these costs are not merely the direct costs paid to the retained high ability individuals, but rather the additional comparison costs triggered throughout the remainder of the organization.
that some of this social comparison may occur indirectly through the managerial hierarchy.

Under these assumptions, a firm with only one activity incurs low social comparison costs because pay is uniform within the activity and there is no other activity with a different pay level with which to compare. When the firm integrates a second activity that differs in its average marginal product from the first, workers within these two activities become the salient referent group and the potential for envy arises. We assume that the greater the scope of activities, where scope is here defined as the dispersion (the difference between the maximum and all others) of average marginal products, the greater will be feelings of inequity among those workers receiving lower income. According to our logic, in response workers pursue behaviors to reduce these envious emotions thereby imposing social comparison costs on the firm. Thus, feelings of inequity and the corresponding comparison costs increase directly with the scope of activities within the firm.

If disparate activities—activities with dispersed average marginal products of labor—are bundled within one firm, the manager faces a range of options to address the potentially high comparison costs that would result from efforts to compensate based on average marginal productivity. Managers can compress wages vertically and no longer pay the average marginal product for each activity. However, much like horizontal wage compression, such vertical wage compression also imposes significant costs on the firm. If those employed in activities with high average marginal productivity of labor are paid well below that average, then those engaged in these high productivity activities are likely to exit the firm in search of firms that do not vertically compress wages or that compress wages less so. Our Harvard, Tenneco, and toy factory illustrations all provide evidence consistent with this scenario. If the firm instead chooses to simply elevate the wages of those engaged in activities with lower average marginal product of labor to be more similar to those with higher average marginal product of labor, this imposes its own obvious costs of overpayment.

Alternatively, the manager can constrain social comparison costs by departing from the technologically efficient production technology, for instance, isolating disparate activities and limiting what may be highly valuable interaction among them. Of course, such efforts may only partially constrain comparison costs and impose their own additional costs in terms of foregone productivity. In determining their best response, managers must consider the fundamental trade-off between social comparison costs arising from envy and productivity losses coming from the adoption of inefficient compensation and technology. Moreover, as activities become more dispersed in terms of the average marginal product of labor, both the comparison costs of rewarding average marginal product and the costs of vertically compressing wages rise. As a consequence, as activities become more dispersed in their average marginal products, constraining the vertical boundaries of the firm may become the effective response.

Note that the increase in comparison costs imposed by dispersion in the marginal productivity of activities is operative at both ends of the distribution. Managers who internalize activities with low average marginal productivity activities find it very costly and difficult to maintain low pay. The envy felt and the resulting comparison costs are reduced by placing the activity in an organization composed of activities with more similar average marginal productivity and more similar pay. Indeed, much outsourcing occurs because managers are unable to adopt or maintain ‘market’ wages for low productivity activities. Instead, social comparison costs in the form of influence activities or reduced effort drive managers to elevate these wages. On the other end of the spectrum, managers who internalize activities with very high marginal productivity also find it costly to adopt market wages for these high productivity activities because low marginal productivity workers impose social comparison costs. Thus, our prediction is that social comparison costs associated with providing market-like incentives increase as the activity scope of the firm increases. Consequently, the more divergent is an activity’s average individual output from the average individual output of all other activities within the firm, the more likely that activity will be outsourced or the wages of those within the activity vertically compressed to match the other activities of the firm.

Interactions between scale and scope

To this point in our discussion of scale and scope, we have examined two simplifications of the firm. In examining scale, we limited the firm to a single activity. In examining scope, we limited the
firm to rewarding average marginal compensation for each activity, that is, horizontally compressing wages within activities. However, firm scale and scope can interact in an important way to shape firm boundaries. In particular, we contend that the social comparison costs that accompany an increase in the firm’s activity scope can increase with the aggregate scale of the firm. Thus, consider two single activity firms, one that is rather large with 10,000 workers and another that is quite small with 50 workers. When the small firm seeks to add a new activity with a higher marginal product of labor and to reward those engaged in that activity, based either on average marginal product or individual marginal product of labor, the 50 lower paid workers engage in comparison, feel envy, and impose comparison costs. Of course, admittedly some of this comparison is indirect, because not all employees will have direct contact with the newly added employees. Indeed, it is likely that a greater percentage of employees in the small firm will have direct contact with the newly added activity. However, with the previously discussed, the managerial hierarchy, as well as informal networks within the firm, will ensure that comparison costs, though perhaps at a diminished level, extend beyond the boundaries of those who directly observe these individuals. Thus, as long as there are positive comparison costs imposed by all within the boundaries of the firm, the comparison costs of adding this additional activity rise with the aggregate scale of the firm. Consequently, larger firms in terms of the number of employees, whether they are engaged in common or disparate activities, confront higher comparison costs in attempting to increase the activity scope of the firm than smaller firms. As firm size increases, vertical pay compression becomes more likely. Similarly, as firm size increases, management is more likely to either constrain the boundary of the firm, and forego integration, or compromise on production technology by isolating geographically this newly added activity.

DISCUSSION

Our theory suggests that social comparison costs play a pivotal role in shaping the boundaries and internal design of firms. We see our theory of social comparison costs as playing a role in organizational failure that is analogous to the role that transaction costs play in market failure. While transaction cost economics views hierarchy as a device to attenuate high transaction costs associated with market failures, social comparison cost economics views the market as a device to attenuate high social comparison costs associated with organizational failures. Thus, while market failures create a type of centrifugal force for moving activities out of the market and into a firm, our theory explains when and how organizational failures create a centrifugal force for moving activities out of the firm and into the market.

Our theory of social comparison costs also has direct implications for the scope and scale of the firm. We argued that the larger the scale of an activity within a firm involving an intermediate degree of team production, the more likely the firm is to horizontally compress wages with the accompanying reduction in incentives. Consequently, the larger the activity, the more likely the firm is to partially outsource the marginal addition to that activity. Thus, a small firm is better able to offer workers their marginal product of labor. These incentives yield a more productive set of workers than larger firms. Our toy factory illustration may provide a case in point. It might have been wiser for the toy manufacturer to outsource the toy painting activity and benefit from the high level of worker productivity through a market interface rather than keep the activity inside the firm, which ultimately led to compressed wages, a loss of productivity, and the departure of specially trained workers. A firm may also benefit from outsourcing when its necessary activities are diverse in scope. Our theory predicts that the firm may benefit by outsourcing those activities that are most distant in terms of their individual marginal product from the average activity. For instance, Tenneco’s acquisition of HOMC internalized a set of activities with a very different set of individual marginal products compared to other activities within the firm. The resulting horizontal and vertical wage compressions experienced by HOMC employees led to a high departure rate and, ultimately, an efficiency loss for Tenneco. Structuring a relationship with HOMC through a contractual interface might have led to a superior outcome. Of course, moving an activity outside the boundaries of the firm may
imply additional costs. Market failures that provide the impetus for placing the activity within the firm are confronted once again should an activity be outsourced. For instance, Larry Summers, Harvard University’s president, was keenly aware that outsourcing Harvard’s investment portfolio would potentially lead to lower performance, but certainly higher fees. In the short term, Harvard decided that outsourcing the activity was too costly and instead compressed wages by restricting the maximum level of compensation. When this caused the departure of key fund managers, increased outsourcing became the more efficient option. Note that while our illustrations have focused on difficulties in internally organizing and rewarding activities with high individual marginal products, similar issues arise when attempting to internally organize and reward activities with low individual marginal products. In such instances, internally organizing low marginal product activities and rewarding based on marginal product generates envious employees who are less productive than they would be if they were bundled with employees earning more similar wages. Moreover, such envy and social comparison costs may cause the firm to elevate wages for the low marginal product activity. While this action may attenuate social comparison costs, simply increasing pay imposes its own obvious costs on the firm. Thus, in determining the optimal boundaries of the firm, management balances the costs that accrue to market failures when an activity is outsourced with the costs that accrue to organizational failures, which we argue are managerial diseconomies of scale and scope that derive from social comparison costs.

Our discussion assumes the decision to internalize an activity depends on a comparison of governance costs, including social comparison costs, between internal and external sourcing. In making this comparison for the marginal activity, we must consider the magnitude of diseconomies of scale and scope imposed by social comparison. The magnitude of these costs depends on the set of activities already organized within the boundaries of the firm and their relation to the activity considered for addition. Therefore, the costs and benefits of integrating a marginal activity will depend not only on the attributes of the exchange, as Williamson has maintained (Williamson, 1985), but also on the attributes of the firm. Firms engaged in different sets of activities or a different scale of activities will experience different levels of governance costs in considering further additions to the firm. Because the firm’s best response policies depend on managerial diseconomies of scale and scope from activities already internalized, the boundaries of the firm are path dependent. Thus, the marginal transaction, as Coase (1937) puts it, depends not only on the attributes of the exchange but also on characteristics of the organizations on both sides of the transaction. Those marginal activities that are more distant in terms of their marginal products from the acquiring firm’s activities will be more costly to integrate. For instance, in our illustration of Syntex and Varian’s joint venture, the differentiated incentive structure offered to the head of the new venture might have survived if Syntex had been a smaller, entrepreneurial firm. Since the marginal productivity of the joint venture appears to have been distant from the rest of Syntex’s activities, the failure of selective intervention appears to be consistent with our theory.

Several limitations pervade our framework. Our theory admittedly focuses on the ‘negative’ effects of envy. Social comparisons lead to workers adopting behavioral strategies that impose costs on the organization. Yet, envy may also stimulate a competitive response. Invidious comparison may cause workers to expend greater effort to increase their output if they expect doing so will lead to higher income. Or, workers might (over) invest in capabilities if they have the expectation that doing so leads to higher income. We anticipate that within limits management may be able to shape expectations so that workers pursue these organizationally ‘beneficial’ behavioral strategies to reduce envious emotions. Incorporating these possibilities into our framework might lead to second order effects. For instance, workers might be less willing on the margin to impose social comparison costs on the firm if they believe they can reduce envy by increasing investment in their capabilities, which would impact the degree to which managers could offer differentiated incentives. While managing worker expectation might affect management’s best structural response to invidious comparisons, we maintain that the basic set of relationships illuminated in our framework should still hold.

Another concern with our framework is that it relies on the conjunction of two distinct perspectives of human nature—two distinct models of man. Scholars comfortable with an emotional model of man where individuals engage in
social comparison and experience negative affect in response to inequity may be uncomfortable with adjoining to this model a rational calculus in which behavioral strategies are chosen to reduce envious emotions. In contrast, scholars who work with a rational model of man may feel uncomfortable with adjoining an emotional model of man. While fully exploring this combination is beyond the scope of our current study, we find that conjoining these models of man is both realistic and fruitful for theory development. Obviously, our assumption will evoke debate and we hope stimulate further discussions about various elements that should be incorporated into the organization theorist’s model of man.

Our predictions focus on the use of three structural levers for mitigating social comparison costs. At present, our theory does not offer specific predictions on when each lever is used or in what order they might be used. We also do not theorize about the extent to which these levers are substitutes or complements. Such questions await further theoretical and empirical development.

Our framework also offers several implications that we mention below but that are beyond the scope of this study to fully develop. Our framework potentially opens up what may be a new area of research about the location (i.e., production technology choices) of workers. For instance, why is it that some firms locate R&D distantly from manufacturing or locate manufacturing distantly from marketing and sales? Or, why do some firms locate workers distantly from other firms? Our theory suggests that these location choices are strategic choices in the manager’s tool kit for shaping salient referents. Knowledge spillover externalities and agglomeration economies provide two approaches for making activity location choices. The consideration of envy provides another explanation for determining these choices.

Mainstream labor economics focuses on asymmetric information and tournaments to explain why promotions are accompanied by a substantial jump in income (Baker, Gibbons, and Murphy, 1994; Medoff and Abraham, 1980). While asymmetric information and tournament models have been used to explain this phenomena (e.g., Lazear and Rosen, 1981; Gibbons and Waldman, 1999), we believe that wage compression within the same level and job classification, that is, horizontal wage compression, primarily results from social comparison costs. The extant literature on envy indicates that social comparison and the resulting costs are likely to be the greatest among those workers who are the ‘closest’ salient referents. Envy should be greatest among peers suggesting that the best response for management may be to define or frame an employee’s set of peers and then within this set dramatically compress wages. This approach leads to discontinuities in income that are linked to promotion or changes in level. Such discontinuities distance the promoted worker as a salient referent thereby attenuating envy and social comparison costs.

These applications suggest that theories built on social comparison costs and envy may lead to a theoretical synthesis that connects several fragmented and disconnected theories of human behavior inside of organizations. While such connections are presently only suppositions, the connections established by our theory suggest potential inquiries.

CONCLUSION

Our goal has been to develop a more complete theory of the firm by highlighting an important source of organizational failure. To do so, we have explored how social comparison and envy impose social comparison costs within firms. Managers must make governance decisions cognizant of both social comparison costs and transaction costs. Management’s best response to social comparison costs may involve a range of options: vertically or horizontally compressing compensation, adopting increasingly inefficient production technology, or shifting entirely the boundaries of the firm.

An economic theory of the firm must explain both when firms arise and when markets arise. Most economics theories of firm boundaries, like transaction cost economics, predict when hierarchies supplant markets. Like a centrifugal force, market failures increase the cost of using markets and push activities within firm boundaries. Until now, the countervailing centripetal force that pushed activities outside of the firm has largely been assumed. Acknowledgment of social comparison costs and envy, we argue, provide at least one theoretical underpinning to explain these diseconomies that shape a firm’s boundaries and design.
ACKNOWLEDGEMENTS

We would like to thank seminar and conference participants at the Academy of Management Meetings, Dartmouth, INSEAD, University of Southern California, University of Pittsburgh, University of Washington, the BYU/University of Utah Strategy Conference, and the Organization Science Winter meeting. We owe special thanks to Stuart Bunderson, Kurt Dirks, Li Ma, Hideo Owan, Brian Silverman, and two anonymous referees.

REFERENCES


