The Firm as a Socialization Device

Abhijit Ramalingam
School of Economics
University of East Anglia
Norwich NR4 7TJ UK
a.ramalingam@uea.ac.uk

Michael T. Rauh
Kelley School of Business
Indiana University
Bloomington, IN 47405-1701 USA
mtrauh@indiana.edu

Abstract

Why do firms exist? What is their function? What do managers do? What is the role, if any, of social motivation in the market? In this paper, we address these questions with a new theory of the firm which unites some major themes in management, principal-agent theory, and economic sociology. We show that while the market is a superior incentive mechanism, the firm has a comparative advantage with respect to social motivation. We then show that the market is efficient in environments that favor the provision of incentives, such as when subjective risk is low and performance is easy to measure. The firm is efficient in other environments where incentives are costly and/or ineffective. We compare our model and results with the views of Durkheim and Granovetter.

Keywords: embeddedness, endogenous preferences, institutions, multi-tasking, norms, principal-agent theory, theory of the firm, trust.

JEL Classification Numbers: L22, M5, Z13.

> The extent of vertical integration and the reasons for the persistence of small firms operating through the market are not only narrow concerns of industrial organization; they are of interest to all students of the institutions of advanced capitalism.
>
> Granovetter (1985, p. 507)

1. Introduction

Why do firms exist? What are their strengths and weaknesses relative to the market? What do managers do? Is the market an impersonal institution devoid of "sentiment" as described by Weber (1978) or is it inextricably embedded in society as asserted by Durkheim (1984), Polanyi (2001), and Granovetter (1985)? Do *homo economicus* or *homo sociologicus* exist? In this paper, we provide formal answers to these questions by drawing together several major themes from economics, management, and sociology.

Our starting point is the *multi-task* theory of the firm in Holmström and Milgrom (1991) (HM91), which consists of one asset, a principal (she), and an agent (he). The agent performs two tasks, production and asset maintenance. In accordance with much of the literature, HM91 define the *firm* to be the institution where the principal owns the asset, while the *market* (non-integration or independent contracting) is the one where the agent owns it. Since asset values are difficult to assess, HM91 assume asset maintenance is non-contractible and therefore cannot be incentivized. It follows that the firm cannot offer incentives for production because the agent would then devote zero effort to asset maintenance. This is called the *multi-tasking problem*, which implies weak (zero) incentives in firms. To make the firm viable with zero incentives, HM91 assume an exogenous amount of intrinsic motivation in the sense that the agent's least-cost total effort is positive and the same across institutions, so the agent is willing to provide some effort for free.

In this paper, we develop a new theory of the firm as a *socialization device*, a term that distinguishes our approach based on *values-based trust* from the alternative approach based on *calculative trust* in the literature on repeated games and relational contracts, including Baker, Gibbons, and Murphy (2002). Although both kinds of trust seem important in actual organizations, in this paper we focus on the *clan* form in Ouchi (1980), "social conditioning" and the *relational team* in Williamson (1981), *commitment* in Kalleberg and

Reve (1993), and *identification* and *docility* in Simon (1991). Similar ideas are widespread in management, organizational theory, and sociology.

Our model combines the theory of the firm in HM91 with the theory of endogenous social preferences in Casadesus-Masanell (2004, section 4) (CM04). In CM04, institutional culture is an equilibrium phenomenon where the principal establishes the *work ethic* and the agent chooses the degree to which he internalizes it. An agent who truly does so experiences *guilt* when his actions fall short of the ideal. Although other models of intrinsic or social motivation are available, we chose CM04 because it effectively endogenizes the agent's least-cost action and hence preserves the structure and completes the results in HM91. It also allows us to overcome Granovetter's (1985) classic *oversocialization* critique because social relations are not exogenous but rather strategically determined based on individual and institutional factors. Finally, the work ethic in CM04 is closely related to several *organizational citizenship behaviors* which have been shown to be positively correlated with organizational motivation and success [e.g., see Podsakoff *et al.* (2000)].

The combined model is strikingly similar to the informal framework in Kalleberg and Reve (1993), which integrates insights from principal-agent theory, transaction cost economics, and sociology. It also shares important aspects with Callon (1998, Introduction) in the *performativity* literature in economic sociology. In the economics literature, Rob and Zemsky (2002) model the development of endogenous social preferences in the firm and Tabellini (2008) in the prisoner's dilemma. The closest paper to ours is Rotemberg (2006), who considers the formation of altruism in firms and markets.

In our model, the fundamental disadvantage of the firm is the same as in HM91: the multi-tasking problem prevents the use of incentives. Since incentives are *balanced* (zero) across tasks, the agent is indifferent between them. This creates a *zone of indifference* or *acceptance* which allows the principal to use her *discretion* to choose the first best ratio of efforts as in Coase (1937), Barnard (1968), Simon (1951), and HM91. In contrast with the oversocialized agent in HM91, in our model the agent *chooses* a positive degree of internalization, which explains (rather than assumes) the existence and viability of the firm. Internalization is less than complete, however, which is consistent with Granovetter's (1985) critique of Williamson (1975, 1981) to the effect that the moral hazard problem does

not disappear inside the firm. The main strength of the market derives from its ownership structure, which solves the multi-tasking problem: the agent has maximum (ownership) incentives for asset maintenance, so the principal can offer strong production incentives such that the agent performs both tasks. The market is therefore good at providing incentives. Its weaknesses are that incentives impose risk when the agent's performance is subject to random shocks and that overall incentives are "unbalanced" in the sense that the market ratio of efforts is generally different from the first best.

We then turn to the role which social motivation serves in each institution. This is an old question which concerned Smith, Marx, Weber, and many other classical writers. In his influential survey, Bowles (1998) gathers an extensive body of experimental and field evidence drawn from across the social sciences to support his contention that preferences are both endogenous and shaped by institutions.

> Walrasian grocery markets support personal interactions quite distinct from the long term relationship characteristic of a lifelong employment firm; and the differences in the structure of these exchanges appear to have effects on preferences...
>
> Bowles (1998, p. 78)

As Bowles observes, subjects tend to behave in a self-interested manner in experiments that mimic market and auction environments, but exhibit social preferences in other contexts like the dictator, ultimatum, and gift exchange games.[1] One potential explanation is that preferences are exogenous and independent of institutions, but that certain institutions restrict the kind of preferences which can be observed.[2] For example, agents with social preferences will nevertheless appear purely self-interested in a perfectly competitive market because they cannot influence the market price or trade volume. As summarized in the title of Sobel (2008), "Markets Make People Appear Selfish."

In this paper, we develop the alternative position in Bowles (1998) that preferences are endogenous and shaped by institutions. The degree of social motivation is naturally captured by the agent's least-cost action (what he does with zero economic incentives) which is exogenous in HM91 but endogenous in our model. A comparison of the agent's

[1] E.g., see Fehr and Gächter (2000), Fehr and Fischbacher (2002), Sobel (2005), Fehr, Klein, and Schmidt (2007), and Ostrom and Walker (2007).

[2] See Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Falk and Fischbacher (2006), Dufwenberg, Heidhues, Kirchsteiger, Riedel, and Sobel (2008), and Sobel (2008).

least-cost action across institutions reveals that social motivation is (i) positive in the market, consistent with Granovetter's (1985) embeddedness position; (ii) higher in the market when subjective risk (objective risk scaled by the agent's degree of risk aversion) is extremely low or performance is easily measured; and (iii) otherwise higher in the firm, consistent with its characterization as a socialization device. In case (ii), the market dominates the firm with respect to both economic and social incentives, but this only holds when the moral hazard problem is negligible. The market is therefore primarily an incentive system, whereas the firm is a socialization device. This contrasts markedly with the theory of the firm in Alchian and Demsetz (1972) and Holmström (1982), where the firm provides stronger incentives for team production than the market, and "The Firm as an Incentive System" in Holmström and Milgrom (1994).

Institutions influence preferences, but are themselves endogenous in the sense that people choose which institutions to create and join. An environment where subjective risk is low and performance is easy to measure is conducive for the provision of incentives which is the strength of the market. In that case, the market will be chosen because it is efficient. The firm will be chosen in other environments where incentives are costly and/or ineffective, where the viability of the firm rests on its comparative advantage with respect to institutional culture. As Henderson (2008) and Akerlof and Kranton (2010) emphasize, this is one of the central functions of management. These results formalize ideas in Ouchi (1980), Williamson (1981), and generalize those in HM91 to the case where the agent's least-cost action is endogenous.

The plan for the rest of the paper is as follows. In section 2 below, we lay out the primitives of the model. We analyze the firm in section 3 and the market in section 4. In section 5, we consider the relative efficiency of each institution and the role that social motivation serves in each. We discuss some related literatures in economics, management, and sociology in section 6. Section 7 concludes. All proofs are in the Appendix.

2. Model Primitives

We consider a simplified version of HM91 (their section 3), which includes one asset, a principal, and an agent. The agent performs two tasks, production and asset maintenance.

The final value of the asset is

$$f(a) = \begin{cases} a - a^2 & \text{if } 0 \leq a \leq 1/2 \\ 1/4 & \text{if } a > 1/2, \end{cases} \tag{1}$$

where $a \geq 0$ is the agent's effort maintaining (or investing in) the asset. The *firm* is defined as the institution where the principal owns the asset and receives its final value, whereas the *market* is the institution where the agent owns it. The principal's benefit (excluding asset value) is

$$B(e, a) = \begin{cases} ea + e & \text{if } a > 0 \\ 0 & \text{if } a = 0, \end{cases} \tag{2}$$

where e is the agent's production effort.[3] A central feature of the HM91 model is that the principal's benefit B drops discontinuously to zero at zero asset maintenance. As we will see, this is one of the basic ingredients of the multi-tasking problem. Again following HM91, we assume asset value f and the principal's benefit B are *non-contractible* (i.e., cannot be incentivized) because they are difficult to measure and verify for third-party enforcement purposes. Instead, the principal observes a contractible signal $y = e + \epsilon$ and offers a linear contract $I = \alpha + \beta y$, where ϵ is a normally distributed random variable with mean zero and variance σ^2, I is the agent's income, α the salary, and β the incentive or piece rate.[4]

For example, a could be the agent's marketing effort which increases the value f of the brand, both of which are difficult to assess, especially for the courts. The brand is an asset whose value accrues to the principal if she owns the brand and the agent otherwise. The agent's production effort e affects both the quantity and quality produced and increases sales B except when no marketing is done. If the sales function is performed by the principal or other agents under her direction then B is manipulable and therefore non-contractible. All the courts observe is an imperfect signal y of the agent's production outcome, which reflects both quantity and quality. The variance σ^2 of ϵ represents *objective risk* or the difficulty of measuring performance especially with respect to quality.

[3] Since the agent is risk averse, we can prevent the principal from selling the store (both f and B) by adding a sufficiently risky shock to (2). The principal is risk neutral so our results would be unaffected.

[4] Given λ and v, we can formally re-interpret (8) as the agent's cost function, so linear contracts are optimal in the sense of Holmström and Milgrom (1987). In this interpretation, institutional culture affects the agent's cost function, the main source of goal incongruence between the two parties.

The agent's utility function is negative exponential

$$-\exp\left\{-r\left[I-(1/2)t^2-G^i\right]\right\},$$

where r is the agent's *coefficient of absolute risk aversion* (the higher r is, the more risk averse he is), $t=e+a$ his total effort, and the quadratic term the cost of total effort. In this paper, we consider the case of *ethical standards* in CM04 where the principal chooses the *work ethic* v to maximize her expected profit and the agent chooses his *degree* λ *of internalization* to maximize his *material* payoff. An agent who truly internalizes the principal's work ethic experiences *guilt* G^i when his actions fall short of the ideal v, which depends on the institution i. We discuss these ideas at length below.

In the firm, guilt is defined by

$$G^f=(1/2)\lambda(v-t)^2,$$

which is a function of total effort t because both tasks are subject to moral hazard.[5] The agent's certainty equivalent payoff is

$$U^f=\alpha+\beta e-(1/2)t^2-G^f-(1/2)k\beta^2, \tag{3}$$

where $k=r\sigma^2$ is *subjective risk* (objective risk σ^2 scaled by the agent's degree r of risk aversion).[6] The first two terms are the agent's expected income and the last is his *risk premium*, which is the disutility of risk. Note that the risk premium is increasing in subjective risk k as well as the incentive β because stronger incentives expose the agent to more risk (his income is tied more closely to the stochastic signal y). The agent's *material payoff* is defined as $M^f=U^f+G^f$, which is (3) without the guilt term. It is therefore the same as his certainty equivalent payoff in HM91 (i.e., his standard economic payoff).

[5] As in Akerlof (1982, Example 2) and CM04, the work ethic serves as a directive to exert a specific level v of total effort, so the agent experiences guilt even when $t>v$. This seems reasonable because deviations from v in either direction can reduce the principal's expected profits and expected total surplus. In any event, $t>v$ in the firm and $e>v$ in the market never occur along the equilibrium path. Note that v corresponds to the norm e_n in Akerlof (1982) and not the minimum standards e^+_{min} and e^-_{min}, so the fact that the poster girls at Eastern Utilities exceeded the minimum work standard is not necessarily inconsistent with our results.

[6] To derive (3), we first calculate the agent's expected utility with respect to the normal density with mean zero and variance σ^2. We then apply a certain monotonic transformation to obtain (3). See Bolton and Dewatripont (2005, Chapter 4) for the details.

In the market, guilt is defined by

$$G^m = (1/2)\lambda(v - e)^2,$$

which depends solely on e because in the market the agent owns the asset and experiences no guilt with respect to its maintenance. The agent's certainty equivalent payoff is

$$U^m = \alpha + \beta e + f(a) - (1/2)t^2 - G^m - (1/2)k\beta^2, \tag{4}$$

where $f(a)$ now appears because he owns the asset. His material payoff is $M^m = U^m + G^m$.

We make the standard assumption that a party who is indifferent over a set of choices is willing to choose as the other party directs. The timing of the game is as follows. (i) The principal and agent choose the efficient institution, which is equivalent to an assignment of asset ownership. As per the Coase theorem, this may entail cash transfers which we ignore. (ii) Given the institution, the agent chooses his degree λ of internalization to maximize his material payoff. The agent is free to choose $\lambda = 0$, in which case he remains *homo economicus*. We assume the agent can commit to his choice of λ (i.e., he cannot change it for the rest of the game). (iii) The principal observes λ and either shuts down or chooses a contract (α, β) and work ethic v. (iv) The agent decides whether or not to accept the contract and, if so, how much effort to devote to each task. If the agent rejects the contract or the principal does not offer one then both parties receive their outside options. (v) If the agent is indifferent, the principal chooses the agent's efforts.

We normalize the principal's outside option to be zero. In the market, the agent's outside option $\overline{U}$ is the value of the asset outside the relationship. If the asset is *specific* [e.g., see Williamson (1981)] then $\overline{U} = 0$. If the asset is non-specific then

$$\overline{U} = \max_{a \geq 0} f(a) - (1/2)a^2 = 1/6 > 0. \tag{5}$$

In general, $0 \leq \overline{U} \leq 1/6$ depending on the degree of asset specificity. In the firm, the principal owns the asset and the agent's outside option is zero.

We now discuss these assumptions in detail. The agent's choice of internalization λ in stage (ii) corresponds to the following dual-self conception:

> one can think of an "inner" self that is selfish and relinquishes control of actions to an "outer" self. What the inner self can do, however, is to mold the preferences that guide the outer self's actions. Thus the inner self can make the outer self altruistic, and this altruism becomes genuine because the inner self cannot change the outer self's preferences too rapidly.
>
> Rotemberg (1994, p. 690)

As CM04 notes, such ideas are common in the social sciences, as well as in economics and game theory.[7] In our context, the assumption that the inner self chooses λ to maximize his material payoff M^i rather than his overall certainty equivalent payoff U^i in (3) or (4) was introduced by Akerlof (1983, p. 54):

> Most persons attempt to choose values for their children (and perhaps also for themselves) according to their economic opportunities that allow them to get along economically... not only the wealthy... but also the poorest of the poor... consciously teach their children values aimed at leading them best to survive economically.

Likewise, Simon (1991, p. 35) stresses the role of *identification* in organizations and argues that *docile* individuals who internalize social norms "may do better at earning a living." The conception of character as an object of individual choice is central in the field of *virtue ethics* in moral philosophy which dates back to Aristotle [see Hursthouse (1999)].

A fundamental insight of CM04 is that prosocial preferences enable the selfish inner self to extract additional payments from the principal. If the agent can commit to a higher λ and credibly signal its value, the principal will have to increase the fixed component α or the agent will reject the contract. Seabright (2004, p. 93) describes a similar mechanism:

> in order to exchange with strangers people need a way to signal their trustworthiness... one of the most effective ways to do this is to create an identity for yourself, a set of internal rules in which you yourself believe and by which you live, and which will make you unhappy if you fail to honor them.

Akerlof (1983, p. 57) and Bowles (1998) discuss evidence that preferences are endogenous and fairly stable. As Rotemberg notes, such stability and commitment may reflect the inability of the inner self to change the preferences of the outer self too rapidly. A potential explanation is *cognitive dissonance*: if only a limited subset of the agent's attitudes, beliefs, and values can be changed at any particular moment then excessive changes will lead to disconcerting inconsistencies.

[7] For example, the *impartial spectator* in Smith's *Theory of Moral Sentiments*, Akerlof (1983), Frank (1987), Raub (1990), Rabin (1993), and more recently Akerlof and Kranton (2005, 2008, 2010), Rotemberg (2006), and Tabellini (2008).

As for credible signals, Seabright (p. 5 and Chapter 6) suggests several characteristics which derive from education, training, and an extended period of commitment to the task:

> almost all occupations in a modern society embody an ethic, a code. For trust requires an assurance of reliability, and some of the most effective policemen are internal, lodged in the surveillance mechanisms of the individual personality. The fiercest external vigilance will rarely be enough to ensure the honesty of a really determined cheat, so what better to deal with people whose character, training, or upbringing leads them not to want to cheat even when they have the chance? Those who can convince others of their intrinsic honesty may thereby prosper, and it may be easier for the genuinely honest to be thus convincing — the more so if honesty, or at least the true and honorable performance of a certain trade or skill, requires a degree of style, confidence, even grace, built up over a long period of commitment to the task, that are hard for an opportunist to feign.
>
> *ibid.,* p. 690.

Granovetter (2005, p. 42) and Ouchi (1980, p. 131) express similar views. Like CM04, we assume that only the "genuinely honest" can credibly signal such honesty and, for simplicity, omit the costs of such signaling. Frank, Gilovich, and Regan (1993, p. 170) discuss evidence which suggests that such traits are at least partially observable: "even on the basis of brief encounters involving strangers, experimental subjects are adept at predicting who will cooperate and who will defect in prisoner's dilemma games."

As the case studies in Akerlof (1982, 1983), Roberts (2004), and Akerlof and Kranton (2005, 2008, 2010) show, many institutions expend considerable resources developing and maintaining an effective culture. Such investments would be worthless, however, without sufficient employee "buy-in" and CM04 captures both sides of this exchange through λ (chosen by the agent) and v (chosen by the principal). This allows us to overcome the *oversocialization* critique in Granovetter (1985, p. 483) because λ and v are strategically chosen to match the social and institutional environment.

> culture is not a once-and-for-all influence but an ongoing process, continuously constructed and reconstructed during interaction. It not only shapes its members but also is shaped by them, in part for their own strategic reasons.
>
> Granovetter (1985, p. 486)

The particulars of the work ethic construct in CM04 and our paper are similar to those in Akerlof (1982, Example 2) and Kandel and Lazear (1992). The work ethic is also closely related to several *organizational citizenship behaviors* in the management and psychology literatures, such as *organizational loyalty*, *organizational compliance*, *individual initiative*,

and *civic virtue*. Podsakoff *et al.* (2000) survey this literature and cite evidence that such behaviors are positively correlated with monetary rewards rather than personality or dispositional factors. Kalleberg and Reeve (1993, p. 1105) provide similar evidence for their hypothesis that "employers exchange money and other job rewards to employees in return for their efforts and sometimes loyalty" which is consistent with the transmission mechanism in CM04.

3. The Firm

We begin with the firm, the institution where the principal owns the asset. We solve the game using backward induction, starting from stage (iv) where the agent decides whether or not to accept the contract and, if so, how much effort to devote to each task. Given the institution, internalization $\lambda > 0$, the contract (α, β), and the work ethic $v \geq 0$, the agent chooses $e \geq 0$ and $a \geq 0$ to maximize his certainty equivalent payoff in (3).[8]

Lemma 1. *(i) If $\beta > 0$ the solution to the agent's problem is $a = 0$ and*

$$e = \frac{\beta + \lambda v}{1 + \lambda}. \tag{6}$$

(ii) If $\beta = 0$ the solution is

$$t = \frac{\lambda v}{1 + \lambda} \tag{7}$$

and the agent is indifferent as to how (7) is allocated across tasks.

The proof is straightforward and omitted. In (i), the agent devotes all of his effort to production, which is rewarded, and none to asset maintenance, which is not because it is non-contractible. In that case, the principal prefers to shut down because her benefit B and asset value f are both zero. This is called the *multi-tasking problem,* which prevents the use of incentives in firms. To make the firm viable with zero incentives, HM91 assume the agent's cost of effort is U-shaped, where the minimum occurs at some positive level of total effort. This implies an exogenous amount of intrinsic motivation in the sense that the

[8] Consider stage (ii) in the firm, where the agent chooses λ to maximize his material payoff M^f. In what follows, we show that $\lambda^f > 0$ and $M^f > 0$ along the equilibrium path. Since $M^f \leq 0$ in all subgames where $\lambda \leq 0$, we may assume $\lambda > 0$ without loss of generality.

agent is willing to supply a positive level of effort (the cost-minimizing level) even with zero incentives. The HM91 theory of the firm is therefore vulnerable to Granovetter's (1985) oversocialization critique because the agent's prosocial behavior is automatic, rigid, and independent of the institutional and social environment. In (ii), the agent is indifferent between the two tasks because incentives are equal (zero) for both. All the agent cares about is his total effort t, whose solution in (7) minimizes the U-shaped expression

$$(1/2)t^2 + G^i \tag{8}$$

with $i = f$, the analogue in our model of the U-shaped cost of effort in HM91. The difference is that we use CM04 to make the least-cost total effort (7) endogenous and socially determined. If $v > 0$ and $\lambda > 0$ the agent supplies positive total effort because he has truly internalized the principal's work ethic. We observe that (7) is increasing in both the work ethic v and the degree λ of internalization.[9]

We now consider the principal's problem in stage (iii). Since the principal owns the asset and $\beta = 0$, her profits are

$$B(e, a) + f(a) - I = B(e, a) + f(a) - \alpha. \tag{9}$$

Given the institution and internalization $\lambda > 0$, the principal chooses the lump sum α, the work ethic v, production effort e, and asset maintenance a to maximize profit (9) subject to four constraints. The first is the *participation constraint* $U^f \geq 0$, which must be satisfied or the agent will reject the contract. Since the agent is indifferent as to how his total effort t in (7) is allocated, the principal chooses e and a subject to the second constraint $e + a \leq t$. This discretionary role of the principal and the region defined by $e + a \leq t$ are similar to the *zone of acceptance* or *indifference* in Coase (1937, p. 391), Barnard (1968), and Simon (1951, 1991).[10] The other two constraints are $e \geq 0$ and $a \geq 0$.

It is clear from (9) that the principal will choose the minimum α consistent with the agent accepting the contract

$$\alpha = \frac{v^2\lambda}{2(1+\lambda)}, \tag{10}$$

[9] Note that comparative statics results with respect to endogenous variables (in this case v and λ) should not be observed. Only those in terms of exogenous variables (e.g., k or $\overline{U}$) are empirically testable.

[10] We use the term "discretion" rather than "authority" as in Coase and Simon because the latter term is better captured in Van den Steen (2010).

which is the solution to $U^f = 0$. Note that α is increasing in internalization and the work ethic, consistent with the empirical evidence in Kalleberg and Reeve (1993) and Podsakoff *et al.* (2000) that employers offer financial rewards in exchange for commitment from their employees. In CM04 and our model, such rewards are the source of employee buy-in and explain why the principal sets a non-infinite work ethic. Since $U^f = 0$,

$$M^f = \alpha - (1/2)t^2 = G^f$$

and the agent's material payoff equals his guilt. Propositions 1 and 2 below are the main results of this section. Proposition 1 characterizes the principal's optimal contract and work ethic and Proposition 2 the agent's optimal degree of internalization.

Proposition 1. *If $\lambda > 0$ the principal offers a unique contract (α^f, β^f) and work ethic v^f with zero incentives $\beta^f = 0$ and*

$$v^f = \frac{4(1+\lambda)}{4+3\lambda}. \tag{11}$$

The agent provides total effort

$$t^f = \frac{4\lambda}{4+3\lambda}, \tag{12}$$

which the principal allocates as follows

$$e^f = \frac{3\lambda}{4+3\lambda} \qquad a^f = \frac{\lambda}{4+3\lambda}. \tag{13}$$

The principal's profit and the agent's material payoff are

$$\Pi^f = \frac{2\lambda}{4+3\lambda} \qquad M^f = \frac{8\lambda}{(4+3\lambda)^2}. \tag{14}$$

Given the agent's degree of internalization, a stronger work ethic induces him to work harder (7) but also requires a higher lump sum payment (10). The optimal work ethic in (11) balances these two effects. To gain some insight into the principal's choice of efforts in (13), we consider the *first best problem*

$$\max_{e\geq 0, a\geq 0} V_{fb} \equiv B(e,a) + f(a) - (1/2)t^2, \tag{15}$$

where it is assumed that the principal can observe the agent's efforts and severely punish him if he does not choose as the principal directs. The first best solution is $e^{fb} = 1$ and $a^{fb} = 1/3$, so Proposition 1 reveals that the principal uses her discretion to choose the same *ratio* of efforts $e^f/a^f = e^{fb}/a^{fb} = 3$ as the first best.

Given the institution, in stage (ii) the agent chooses λ to maximize his material payoff M^f in (14).

Proposition 2. *The agent's optimum is* $\lambda^f = 4/3 > 0$, *so in equilibrium*

$$e^f = 1/2, \quad a^f = 1/6, \quad \beta^f = 0, \quad v^f = 7/6, \quad M^f = 1/6, \quad \textit{and} \quad \Pi^f = 1/3. \tag{16}$$

A higher λ increases the payment α but also requires more work, which explains why internalization is neither zero nor infinite. If the agent fully internalized the principal's work ethic in the sense that $\lambda = \infty$, the principal could achieve the first best by setting $\beta = 0$ and $v = 4/3$. The agent would be forced to choose $t^{fb} = 4/3$ to avoid infinite guilt and the principal could then use her discretion to choose $e^{fb} = 1$ and $a^{fb} = 1/3$. In equilibrium, however, internalization is only partial $0 < \lambda^f < \infty$ and the moral hazard problem persists within the firm. This is consistent with Granovetter's (1985, p. 500-1) criticism of Williamson's (1975, 1979, 1981) "oversocialized view that orders within a hierarchy elicit easy obedience and that employees internalize the interests of the firm, suppressing any conflict with their own." Although Granovetter draws on case studies where networks of agents effectively colluded against their superiors, we have shown that less than full internalization can occur with only one agent.[11]

We summarize these results as follows. In our model, the firm is a stylized institution which operates solely on the basis of institutional culture without any monetary incentives. Its strengths include balanced (zero) incentives across tasks, which creates the indifference necessary for managerial discretion, which in turn is used to achieve the same ratio of efforts as the first best. In the next section, we will show that incentives are generally unbalanced in the market. Another advantage of the firm is that the agent bears zero risk (a zero incentive implies a zero risk premium). On the other hand, the multi-tasking

[11] We thank Olav Sorenson for making this connection.

problem prevents the firm from providing incentives, which means it can never achieve the first best.

4. The Market

We now turn to the market, where the agent owns the asset. We again proceed by backward induction starting from stage (iv), where the agent chooses his efforts $e \geq 0$ and $a \geq 0$ to maximize his certainty equivalent payoff (4) given the institution, internalization $\lambda > 0$, the contract (α, β), and the work ethic v.

Lemma 2. *The agent's optimal efforts in the market are given by*

(Region 1) $e = 0$ *and* $a = 1/3$ *when*

$$\beta + \lambda v \leq 1/3. \tag{17}$$

(Region 2) e *equals (6) and* $a = 0$ *when*

$$\beta + \lambda v \geq 1 + \lambda. \tag{18}$$

(Region 3)

$$e = \frac{3(\beta + \lambda v) - 1}{2 + 3\lambda} \qquad a = \frac{1 + \lambda - (\beta + \lambda v)}{2 + 3\lambda} \tag{19}$$

otherwise.

From (4), we observe that the agent's production effort is motivated by the incentive β, internalization λ, and the work ethic v, while asset maintenance is motivated by the returns f to asset ownership. According to Lemma 2, the agent focuses exclusively on asset maintenance when $\beta + \lambda v$ is relatively low in the sense of (17), on production when $\beta + \lambda v$ is relatively high (18), and both tasks when $\beta + \lambda v$ is intermediate. The principal will never choose a contract in region 2 because her benefit B is zero when asset maintenance is zero. From (7) and (19) we observe an important qualitative difference between the two institutions: a stronger work ethic increases *total* effort in the firm, whereas in the market a stronger work ethic (or an increase in incentives) increases production effort at the *expense* of asset maintenance and therefore serves an allocative role.

We now consider the principal's problem. The principal's profits and expected profits are $B - I$ and

$$B(e, a) - \alpha - \beta e, \tag{20}$$

respectively. Given the institution, internalization λ, and the agent's efforts in Lemma 2, the principal chooses the contract (α, β) and the work ethic v to maximize (20) subject to the participation constraint $U^m \geq \overline{U}$, where $\overline{U}$ is the agent's outside option. As before, the principal chooses the smallest α consistent with the agent accepting the contract. The agent's material payoff is therefore

$$M^m = G^m + \overline{U} = \alpha + \beta e + f(a) - (1/2)t^2 - (1/2)k\beta^2.$$

Propositions 3 and 4 below are the main results of this section.

Proposition 3. *Let*

$$\lambda^q = (16/27) - (1/k). \tag{21}$$

(i) If $\lambda \geq \lambda^q$ the unique solution to the principal's problem entails

$$\beta^m = \frac{5}{2\left[3 + k(1 + 3\lambda)\right]} \qquad v^m = \frac{9 + k(8 + 9\lambda)}{4\left[3 + k(1 + 3\lambda)\right]}, \tag{22}$$

the agent's optimal efforts are

$$e^m = \frac{9 + k(9\lambda - 2)}{4\left[3 + k(1 + 3\lambda)\right]} \qquad a^m = \frac{1 + k(2 + \lambda)}{4\left[3 + k(1 + 3\lambda)\right]}, \tag{23}$$

and

$$\Pi^m = \frac{13 + k(13\lambda - 4)}{8\left[3 + k(1 + 3\lambda)\right]} - \overline{U} \qquad M^m = \frac{25k^2\lambda}{8\left[3 + k(1 + 3\lambda)\right]^2} + \overline{U}. \tag{24}$$

(ii) If $0 \leq \lambda < \lambda^q$ the principal sets $\beta^m = v^m = 0$, the agent's optimal efforts are in region 1 of Lemma 2, $\Pi^m = (1/6) - \overline{U} \geq 0$, and $M^m = \overline{U}$.

The expression in (21) is the level of internalization which equalizes the principal's expected profits in cases (i) and (ii). Since Π^m in (24) is increasing in λ, the principal chooses (i) when $\lambda \geq \lambda^q$ and (ii) otherwise.[12] In (ii), internalization is too low to support

[12] If $\lambda = \lambda^q$ the principal is indifferent between (i) and (ii) but the agent strictly prefers (i) for all $\lambda > 0$. Given our assumption about indifference in section 2, the principal will agree to choose (i).

production, so the principal sets $\beta^m = v^m = 0$ and production effort is zero.[13] This case never occurs in equilibrium, however, because the agent's material payoff is strictly higher in (i) for all $\lambda > 0$. It follows that the agent always chooses $\lambda \geq \lambda^q$ and we focus exclusively on (i) from now on.

Proposition 3 highlights two important qualitative differences between the firm and the market. The first is that the incentive in (22) is strictly positive rather than zero as in the firm. This follows directly from the difference in ownership structures which distinguishes the two institutions. Since asset maintenance is non-contractible, economic incentives can only be provided via ownership. In the firm, the agent has no direct interest in the asset and will not maintain it when production is rewarded. In contrast, the agent owns the asset in the market and will therefore continue to maintain it with positive production incentives as long as they are not too high as in region 2 of Lemma 2. The other qualitative difference is that overall incentives in the market (by this we mean the combined effect of λ, v, the returns f to asset ownership, and β) are generally unbalanced in the sense that the ratio e^m/a^m of efforts generically differs from the first best $e^{fb}/a^{fb} = 3$.

The following comparative statics results are immediate from Propositions 3 and 4. We record them here because of their importance in the next section. We will discuss (iv) after Proposition 4.

Lemma 3. *(i) λ^q is increasing in k. (ii) β^m in (22) is decreasing in both k and λ. (iii) v^m in (22) is increasing in k and decreasing in λ. (iv) λ^m is decreasing in k.*

The intuition is as follows. An increase in subjective risk increases the agent's risk premium, which increases the necessary payment α for the agent to accept the contract. The principal therefore lowers incentives in (ii) to reduce the risk premium and increases the work ethic in (iii) to partially offset the loss of economic incentives. The decline in expected profits raises the minimum threshold λ^q necessary to make the principal indifferent between (i) and (ii) in Proposition 3. An increase in internalization λ increases the agent's overall incentives for production. This allows the principal to lower production incentives in (ii)

[13] From (20), $\alpha = \overline{U} - (1/6) < 0$. In effect, the principal rents the firm to the agent and extracts the difference, due to asset specificity, between the surplus created within the relationship (5) and the agent's outside option. Grossman and Hart (1986, p. 693) discuss a similar contract in the context of more general transaction costs.

and the work ethic in (iii) to reduce the risk premium and guilt, which in turn reduces the payment α. The result that production incentives are decreasing in subjective risk is called the *risk-reward tradeoff* in the principal-agent literature. All the other results, which address the interaction between economic and social incentives, are new.

Given the institution, the agent's problem is to choose λ to maximize his material payoff M^m in (24) subject to the constraint $\lambda \geq \lambda^q$ that internalization is high enough to support production in (i) of Proposition 3.

Proposition 4. *(i) (pure market) If* $0 \leq k \leq 54/7$ *the agent's optimal* λ *is*

$$\lambda^m = (1/6) + (1/k) \tag{25}$$

and

$$\begin{array}{ccccc} e^m = \frac{18+k}{8(3+k)} & & a^m = \frac{6+7k}{24(3+k)} & & \beta^m = \frac{5}{4(3+k)} \\ v^m = \frac{18+11k}{8(3+k)} & M^m = \frac{25k}{96(3+k)} + \overline{U} & & \Pi^m = \frac{78+k}{48(3+k)} - \overline{U}. & \end{array} \tag{26}$$

(ii) (quasi-firm) If $k > 54/7$ *the optimal* λ *is* $\lambda^q > 0$ *and*

$$\begin{array}{ccccc} e^q = \frac{3}{10} & & a^q = \frac{7}{30} & & \beta^q = \frac{9}{10k} \\ v^q = \frac{6}{5} & M^q = \frac{6}{25} - \frac{81}{200k} + \overline{U} & & \Pi^q = \frac{1}{6} - \overline{U}. & \end{array} \tag{27}$$

The expression in (25) maximizes M^m ignoring the constraint $\lambda \geq \lambda^q$. As recorded in Lemma 3, λ^m is decreasing in subjective risk k because a decrease in k reduces the work ethic, the agent's guilt, and his material payoff. If the constraint $\lambda \geq \lambda^q$ is not an issue, the agent offsets this with an increase in internalization λ^m. Figure 1 below depicts λ^i for $i = f, m, q$ as a function of k. We observe that $\lambda^q \leq \lambda^m$ when subjective risk is relatively low $k \leq 54/7$ so the agent chooses $\lambda = \lambda^m$, while $\lambda^m < \lambda^q$ when subjective risk is relatively high $k > 54/7$ and the agent is essentially forced to choose $\lambda = \lambda^q$.

Figure 1. *Internalization* λ^i*, risk premium* RP^i*, effort ratio* e^i/a^i*, productive efficiency* V^i_{fb}*, work ethic* v^i*, and social motivation* l^i *as functions of subjective risk* k*.*

In (ii), the incentive β^q declines to zero as subjective risk increases without bound, while internalization λ^q increases up to its maximal level 16/27. In this case, the market resembles the *quasi-firm* characterized by Eccles (1981) because it relies almost exclusively

on social motivation. An example is the construction industry, where relations between general contractors and subcontractors (who tend to own their own tools) are "stable and continuous over fairly long periods of time" (p. 340). As Eccles notes, such stability seems inconsistent with competitive bidding and other forms of market contracting based on price. Furthermore, 71.6% of the observed contracts were fixed-price contracts (p. 350) with "obvious incentives for shirking performance requirements" (p. 340). In our model, the quasi-firm contract takes the form of a fixed price α and negligible incentive β^q.

In summary, the main advantage of the market is that its ownership structure solves the multi-tasking problem so that economic (as well as social) incentives can be provided. Its disadvantages are that overall incentives are generally unbalanced across tasks and production incentives generate risk.

5. Efficiency

In the first stage of the game, the principal and agent choose the institution with the highest expected total surplus

$$V^i = \Pi^i + M^i = V_{fb} - (1/2)k\beta^2, \tag{28}$$

where $i = f, m, q$ and V_{fb} is defined in (15). Note that the agent's material payoff cancels in (28) because it is just a transfer. As shown in Proposition 4, the choice is between the firm and the market (i) when $0 \leq k \leq 54/7$ and the firm and the quasi-firm (ii) when $k > 54/7$. The efficient institution is therefore determined on the basis of productive efficiency V_{fb} (value created minus the cost of effort) and the risk premium, where V_{fb} is a function of efforts in both tasks, which in turn are functions of economic incentives (production β and ownership f incentives) and social incentives (internalization λ and the work ethic v).

Proposition 5. *(i) In equilibrium,*

$$V^f = 1/2, \quad V^m = \frac{52 + 9k}{32(3 + k)}, \quad \textit{and} \quad V^q = \frac{61}{150} - \frac{81}{200k}. \tag{29}$$

(ii) The market is the efficient institution for all $0 \leq k < 4/7$ and the firm for all $k > 4/7$.
(iii) The quasi-firm is never efficient.

The expressions in (29) follow directly from (16), (26), and (27). A straightforward comparison reveals that $V^m > V^f$ for all $0 \leq k < 4/7$, $V^f > V^m$ for all $4/7 < k < 54/7$, and $V^f > V^q$ for all $k > 54/7$ (with equality at the cut-offs). As in Ouchi (1980), Williamson (1981), and HM91, the firm is the efficient institution when subjective risk (or the difficulty of measuring performance in Williamson) is high, while the market is efficient otherwise.

To explain these results, Figure 1 depicts the agent's risk premium

$$RP^i = (1/2)k\beta^{i\,2}$$

for $i = m, q$ as a function of subjective risk k. An increase in subjective risk increases the risk premium directly but also reduces the incentive β^m because of the risk-reward tradeoff. As a result of these two opposing forces, RP^m at first rises and then declines. The market becomes the quasi-firm for all $k > 54/7$ (at the kink in Figure 1) and RP^q declines to zero. Since the risk premium is zero in the firm, it dominates the market on this dimension.

Figure 1 also exhibits the ratio e^i/a^i of efforts for $i = f, m, q$ from Propositions 2 and 4. While the firm achieves the first best ratio $e^f/a^f = 3$ because of balanced incentives and discretion, the market is unbalanced in favor of production when subjective risk k is low and asset maintenance when k is high. These imbalances partly explain why productive efficiency V_{fb} is highest in the firm when k is sufficiently high. From (28), the market is efficient when its advantage in terms of productive efficiency is sufficiently great to compensate for its risk premium. In contrast, the quasi-firm is never efficient because the firm dominates it on both dimensions.

We now consider the degree of social motivation in each institution. As Muller (2003) shows, the issue as to whether the market reflects anything more than self-interest has a long and distinguished intellectual history. Classical authors such as Marx and Weber often depicted it as an impersonal institution which actively dissolves personal ties and traditional values.

> The market community as such is the most impersonal relationship of practical life into which humans can enter with one another... Where the market is allowed to follow its own autonomous tendencies, its participants do not look toward the persons of each other but only toward the commodity; there are no obligations of brotherliness or reverence, and none of those spontaneous human relations that are sustained by personal unions. They all would just obstruct the free development of the bare market relationship, and its specific interests serve, in their turn, to weaken the sentiments on which these obstructions rest.
>
> Max Weber, *Economy and Society* (1978, p. 636).

Other writers, including Durkheim (1984), Polanyi (2001), and Granovetter (1985), argue that markets are always embedded in society and that pure contractual relations cannot exist or be studied separately from social considerations.

> economic phenomena cannot be adequately studied in the manner of classical economic theory, as if these were separate from the moral norms and beliefs which govern the life of individuals in society. There is no society (nor could there conceivably be a society) where economic relationships are not subject to customary and legal regulation. That is to say, as Durkheim was later to express the matter in *The Division of Labor*, "a contract is not sufficient unto itself."
>
> Giddens (1971, p. 69)

In our model, the degree of social motivation is measured by the agent's least-cost action l^i, which is what he would choose without any economic (production or ownership) incentives. In HM91, l^i is assumed to be exogenous, the same in both institutions, and positive to make the firm viable with zero incentives. In this paper, we combine HM91 and CM04 to endogenize l^i and make it the product of social forces. In the firm, (8) is minimized by $l^f = t$ in (7). In the market, (8) is minimized by $a^i = 0$ and $e^i = t$, so

$$l^i = \frac{\lambda^i v^i}{1 + \lambda^i}$$

for $i = f, m, q$. The least-cost action therefore has the same functional form in all three institutions, but its arguments λ^i and v^i generally differ.

Lemma 4. *(i) The agent's least-cost actions are*

$$l^f = 2/3, \quad l^m = \frac{18 + 11k}{8(3 + 4k)}, \quad \textit{and} \quad l^q = \frac{6(27 - 16k)}{5(27 - 43k)}.$$

(ii) $l^m > l^f$ *for all* $0 \leq k < 6/31$ *but otherwise the least-cost action is highest in the firm.*

In Figure 1, we observe that social motivation l^i is strictly positive in the market for all levels of subjective risk, which is consistent with the embeddedness view, but generally

higher in the firm, consistent with its characterization as a socialization device. The exception ($l^f < l^m$) occurs when subjective risk is extremely low $0 \leq k < 6/31$ and the moral hazard problem is negligible. In our model, these differences in social motivation reflect choices within different institutions with different ownership structures. In the firm, social incentives are necessary to partially complete a maximally incomplete economic contract (zero incentives in both tasks) crippled by the multi-tasking problem. In the market, social incentives may be less central because the corresponding ownership structure allows for more complete market contracting in the form of ownership incentives f for asset maintenance and production incentives β for production effort. Indeed, in a world of complete and perfect contracts (e.g., classical general equilibrium theory in economics) there would be little or no role for either firms or ethics.[14]

We can now provide a complete account of the strengths and weaknesses of each institution. When subjective risk is extremely low $0 \leq k < 6/31$, the market is superior because it provides strong economic incentives in the form of ownership incentives for asset maintenance as well as strong production incentives β^m, whose cost in terms of the agent's risk premium is low. Given strong production incentives, the market does not need a strong work ethic v^m [see Lemma 3 and Figure 1] and the principal institutes a weak one to reduce the fixed payment α. A weak work ethic reduces the agent's material payoff, who chooses a high degree λ^m of internalization and we obtain the novel result that the market dominates the firm in terms of both economic and social incentives.

At higher levels of subjective risk, the principal reduces incentives and increases the work ethic v^m to partially offset the loss of economic motivation. A stronger work ethic increases the agent's material payoff, who reduces internalization λ^m to lower his cost of effort. The combined effect is that social motivation l^m falls in Figure 1. As economic and social incentives (and therefore productive efficiency) decline in the market and the risk premium RP^m rises, the firm eventually $4/7 < k < 54/7$ becomes the efficient institution. When $k > 54/7$ the market takes the form of the quasi-firm, but production incentives and social motivation are too low to overcome its disadvantages relative to the firm in terms of unbalanced incentives and a positive risk premium.

[14] This discussion paraphrases insightful remarks made by an anonymous referee.

6. Related Literatures

Durkheim

Another parallel, in addition to his embeddedness position discussed earlier, is Durkheim's emphasis on the two-way interaction between the economic and the social. In contrast, Bowles (1998) and other authors emphasize the one-way effects of economic institutions on preferences.

> One can understand nothing of the rules of morality that govern property, contract, work, etc., if one does not know the economic causes which underlie them; and, conversely, one would arrive at a completely false notion of economic development if one neglected the moral causes which influenced it.
>
> Durkheim quoted in Giddens (1973, p. 69).

In our model, it may appear from the timing of the game that institutions are one of the "economic causes" of "the rules of morality that govern... contract" (i.e., internalization and the work ethic) but this is misleading because institutions are themselves chosen in a forward-looking manner. The relationship is therefore two-way in the sense that both the economic (institutions and contracts) and the social are endogenous variables determined in equilibrium and therefore cannot legitimately be studied in isolation. Durkheim also stresses the positive attachment (internalization) to ideals (the work ethic) and obligation (guilt) as opposed to norms enforced by punishments as in the literature on repeated games (*ibid.,* p. 88). Indeed, this is the crucial distinction between CM04 and our paper and the literature on relational contracts discussed below. In our model, "morality is a collective property and must be studied as such" (*ibid.*, p. 69) in the sense that institutional culture is the product of strategic interaction and is therefore a *social* rather than individual phenomenon. In the statement "moral rules are shaped by society, under the pressure of collective needs" (*ibid.*, p. 69), we interpret "collective needs" in terms of efficiency.

Callon (1998)

A central issue in Callon (1998, Introduction) and the *performativity* literature concerns the emergence of *calculativeness* or rational economic action. As in this paper, preferences are endogenous but Callon rejects the view that institutions determine preferences because

institutions are also endogenous (*ibid.*, p. 6) and constructed through strategic interaction (*ibid.*, p. 44). Rationalization occurs through a process involving lock-in (*ibid.*, p. 48-50) where the agents' choices successively "frame" the network.[15] The outcome of this process is *homo economicus* (*ibid.*, p. 51) in a state similar to the game-theoretic concept of common knowledge of rationality and the rules of the game (*ibid.*, p. 50). As in Durkheim, once the network is established individuals cannot be conceptually separated from it (*ibid.*, p. 8-9). A similar lock-in process occurs in our model, where the contracting stage (iii) (the subject of standard principal-agent theory) is successively framed by the choice of institution and the agent's degree of internalization. The difference is that Callon describes the evolution of bounded rationality into rational and self-interested *homo economicus*, whereas we show how rational *homo economicus* consciously evolves into *homo sociologicus* and how that decision depends on factors such as economic institutions and subjective risk.

Theories of the Firm

We have already mentioned Coase (1937) in connection with Proposition 1. According to Williamson (1975, 1981, 1985), the three main dimensions of transactions are asset specificity, frequency, and uncertainty. In an employment context, Williamson (1981) adds a fourth dimension: the difficulty of measuring performance as in this paper and the principal-agent literature. The *relational team* (which he likens to Ouchi's clan form) is the efficient employment relation when asset specificity and the difficulty of measuring performance are high. At present, our model cannot account for frequency or uncertainty (which Williamson links to ex post dispute resolution) and our result is more limited because asset specificity only affects the agent's fixed payment α.

The property rights theory of the firm includes Grossman and Hart (1986) and Hart and Moore (1990). In this model, the parties first assign asset ownership and then make non-contractible investments similar to asset maintenance in our model. It is assumed that the surplus created by those investments cannot be divided ex ante and is subsequently allocated through Nash bargaining (the Shapley value with more than two parties). The

[15] The role of the economics profession and other disciplines in the framing process does not concern us here.

ownership structure determines the parties' outside options and therefore their payoffs from Nash bargaining, so the role of asset ownership is to provide efficient investment incentives. Although the exact mechanisms differ, asset ownership serves a similar role in HM91 and our model, with the added complexity that the agent performs two tasks which compete for his attention.

Baker, Gibbons, and Murphy (2002) add *relational contracts* to the property rights theory of the firm. The model consists of one asset, a principal, and an agent, whose efforts influence the value of output and the outside value of the asset (the agent's outside option in the market). The parties choose between spot (Nash bargaining) or relational (repeated game) contracts which entail non-binding promises to pay a bonus or fine depending on the outcome. The punishment for reneging is permanent reversion to spot contracting. Instead of the classical risk-reward tradeoff, there is a feasibility-reward tradeoff because large bonuses or punishments increase the temptation to renege on the relational contract. This temptation is greater in the firm where the principal owns the asset and will therefore be in a stronger position in the Nash bargaining following reneging. As in our model, this implies weak incentives in firms.

Henderson (2008) argues that hard-to-imitate relational contracts form the basis for important organizational competencies such as turnaround times at Southwest Airlines. Like Durkheim, she distinguishes between "calculative trust" based on punishments as in Baker *et al.* and "values-based trust" as in CM04 and this paper. Henderson also argues that both kinds of trust are important in actual organizations and that managing relational contracts is one of (perhaps the main) functions of management.

As discussed in Dal Bó (2005), the experimental evidence for calculative trust is mixed. Although Dal Bó's findings are consistent with standard predictions for the Prisoner's Dilemma with random continuation, cooperation rates were generally well below 50%. In contrast, the experimental results on fairness and reciprocity [see footnote 1] have been widely replicated and show that other-regarding preferences can achieve cooperation even in anonymous one-shot settings. Since calculative trust requires an infinite or indefinite horizon, values-based trust should be fostered in contexts where the "shadow of the future" is weak (e.g., for managers later in their careers). Furthermore, subjects cast in the

role of principals preferred bonus contracts supported by reciprocity over monitoring and punishments in Fehr, Klein, and Schmidt (2007) and instead of incentive contracts in the multi-task experiments of Fehr and Schmidt (2004). After reviewing several case studies, Akerlof and Kranton's (2010, p. 58) assessment of the logic of calculative trust is that "this reasoning does not reflect what we see and hear on the shop floor or in the trenches."

Vohs, Mead, and Goode (2006) conduct a series of experiments designed to elicit the effects of money on social behavior. In one experiment, subjects filled out a computer questionnaire. A third of them were randomly assigned to computers with screensavers involving money, a third to screensavers involving fish, and the final third to computers whose screens went blank. Afterwards, subjects were asked to pull two chairs together ostensibly to meet another participant. Those in the money screensaver treatment kept the two chairs furthest apart. In another experiment, a confederate of the experimenters pretended to spill a box of pencils. Those who were "primed" with money helped pick up the fewest. All nine experiments had similar outcomes: "participants primed with money preferred to play alone, work alone, and put more physical distance between themselves and a new acquaintance" (p. 1154).

It seems difficult to reconcile these findings with the explanation in Dufwenberg *et al.* (2008) and Sobel (2008) because the experimental setting did not constrain behavior. Nor is it clear how the theory of repeated games applies. A natural interpretation, in line with Bowles (1998), is that the same individual can have different sets of preferences which are adapted to different institutions and that sensory cues involving money activated those preferences associated with money and markets. This interpretation is broadly consistent with the approach in this paper.

Other Economic Models

Tabellini (2008) considers similar issues within a version of the prisoner's dilemma where players have an endogenous preference for cooperation chosen partly by their parents. Under certain conditions, there are two steady-states (see his proposition 9) — one with a strong endogenous enforcement mechanism (the probability of detecting and punishing defectors) and where the majority of players have a strong preference for cooperation, and

another steady-state with the opposite properties. Unlike this paper, incentives and social motivation are therefore positively related within equilibrium institutions.

The closest paper to ours is Rotemberg (2006), who considers the formation of altruism in firms and markets. In his paper, firms have an inherent advantage over markets because monitoring can also lead to product improvements. Given this advantage, altruism is less valuable in the firm and Rotemberg provides sufficient conditions (see his theorems 2 and 3) such that the firm prefers to deal with altruistic independent contractors rather than selfish employees.

Rob and Zemsky (2002) develop a dynamic model of the firm (the sole exogenous institution) with a continuum of risk neutral agents. In each period, agents allocate effort between an individual and a cooperative task, where the latter is more profitable but the former more easily measured. The form of social motivation is similar to that in CM04, except the degree of internalization is increasing in past cooperation (a version of reciprocity). An increase in incentives increases individual effort at the expense of cooperative effort in the current period and therefore reduces cooperation in the future. There are two potential steady-states: a "good" equilibrium with low incentives, high cooperation, and high profits, and a "bad" equilibrium. As in our model, incentives and social motivation are therefore negatively related.

7. Conclusion

In this paper, we combined the multi-task theory of the firm in Holmström and Milgrom (1991) with the theory of endogenous social preferences in Casadesus-Masanell (2004). We then showed how the combined model captures several seemingly disparate themes in economics, management, and sociology. These include Durkheim's conception of solidarity founded upon obligation and the positive attachment to ideals and the two-way interaction between the economic and the social; embeddedness and the oversocialization critique in Granovetter (1985); theories of the firm based on identification and commitment like the clan form in Ouchi (1980), the relational team in Williamson (1981), Kalleberg and Reve (1993), and Simon (1991); and standard principal-agent theory. We then used the model to provide answers to such classical questions as the character of the firm, the nature and

role of market relations, and the impact of economic institutions on preferences.

In our model, the market is good at providing incentives and is the superior institution when performance is easily specified and measured and subjective risk is low. In contrast, the firm has a comparative advantage in terms of social motivation. Individuals choose their own social relations and values based on their material self-interest, in exchange for financial and perhaps other rewards. Institutions and their associated forms of motivation (economic and/or social) are collectively chosen to serve the general interest. We therefore stress the instrumentality of institutions and social relations over the ethical considerations often emphasized in the classical literature.

Our model also provides an alternative to the theory of the firm based on calculative trust in Baker *et al.* (2002) with different managerial implications. In Casadesus-Masanell (2004) and our model, the role of the manager is to set appropriate normative standards and foster their internalization, as opposed to monitoring and punishing. In particular, managers should assess both objective performance (the signal y) and internalization and reward them *separately* through incentives β and the salary α, respectively. Podsakoff *et al.* (2000) discuss evidence which suggests that this is already common practice. Our model takes this a step further and recommends its extension to market relationships (e.g., with external suppliers) albeit perhaps to a lesser extent.

We conclude with some avenues for future research. (i) As a robustness check, it would be useful to consider other social preferences such as fairness and reciprocity in Fehr and Schmidt (1999). (ii) In our model, the efficient institution is determined solely by subjective risk, whereas asset specificity is also important in Williamson (1981) and the empirical literature on transaction costs. (iii) The quasi-firm is an important real-world institution, but is never efficient in our model. It can only arise when subjective risk is high (so the market takes the form of the quasi-firm) but the principal cannot acquire the asset for some reason. An example would be capital market imperfections which prevent the principal from borrowing the necessary funds. (iv) Our model is essentially static and involves only one agent. In contrast, Harrison and Carroll (2002) simulate a more realistic model with multiple agents and dynamic cultural influence processes to study the effects of organizational size, hiring selectivity, turnover rates, and other factors on institutional

culture. The advantage of our approach is that institutional culture is endogenous, whereas Harrison and Carroll assume exogenous cultural transmission mechanisms disconnected from the rest of the governance structure (e.g., incentive mechanisms).

Appendix

Proof of Proposition 1

Substituting $U^f = 0$ into (9),

$$\Pi^f = B(e,a) + f(a) - (1/2)t^2 - G^f. \tag{A.1}$$

The principal therefore chooses $e \geq 0$, $a \geq 0$, and $v \geq 0$ to maximize (A.1) subject to $e + a \leq t$, where t is given in (7). We first consider the region where $a \geq 1/2$. On that region, the principal maximizes (A.1) with $f(a) = 1/4$ [see (1)] subject to the constraints $e \geq 0$, $a \geq 1/2$, and $e + a \leq t$. The last constraint clearly binds, so substituting $e = t - a$ into (A.1), we obtain the negative definite quadratic form

$$\Pi^f = \frac{1}{4} + \left(\frac{\lambda}{1+\lambda}\right) v - a + \left(\frac{\lambda}{1+\lambda}\right) av - a^2 - \frac{1}{2}\left(\frac{\lambda}{1+\lambda}\right) v^2. \tag{A.2}$$

The unconstrained maximizer for (A.2) is

$$a = -\frac{1}{2+\lambda} \qquad v = \frac{1+\lambda}{2+\lambda} \tag{A.3}$$

so the constrained solution occurs on the boundary $a = 1/2$. We can therefore restrict attention to the region $0 \leq a \leq 1/2$ where $f(a) = a - a^2$. Again substituting $e = t - a$ into (A.1), we obtain the negative definite quadratic form

$$\Pi^f = \left(\frac{\lambda}{1+\lambda}\right) v + \left(\frac{\lambda}{1+\lambda}\right) av - 2a^2 - \frac{1}{2}\left(\frac{\lambda}{1+\lambda}\right) v^2. \tag{A.4}$$

The unconstrained maximizer for (A.4) is given in (11) and (13). Since $a^f < 1/2$ in (13), this is the unique global maximum for the principal's problem. The rest of the statement follows from straightforward substitutions. ∎

Proof of Proposition 2

Since the partial derivative

$$M_{\lambda}^{f} = \frac{8(4-3\lambda)}{(4+3\lambda)^{3}}, \tag{A.5}$$

$M^f = 0$ at $\lambda = 0$, increases to its maximum at $\lambda^f = 4/3$, and decreases thereafter. The rest follows from straightforward substitutions. ∎

Proof of Lemma 2

The agent never chooses $a > 1/2$ [see (1)] because effort is costly. After replacing $f(a)$ in (4) with $a - a^2$, we obtain a negative definite quadratic form. The first-order conditions are

$$U_{e}^{m} = \beta - e - a + \lambda(v - e) \leq 0 \tag{A.6}$$

$$U_{a}^{m} = 1 - e - 3a \leq 0 \tag{A.7}$$

with complementary slackness. If $e = 0$ and $a > 0$ (Region 1) then (A.7) implies $a = 1/3$ and (A.6) implies (17). If $e > 0$ and $a = 0$ (Region 2) then (A.6) implies (6) and (A.7) implies $e \geq 1$ or (18). If $e > 0$ and $a > 0$ (Region 3) then (A.6) and (A.7) imply (19). ∎

Proof of Proposition 3

Substituting $U^m = \overline{U}$ into (20),

$$\Pi^{m} = B(e,a) + f(a) - (1/2)t^{2} - G^{m} - (1/2)k\beta^{2} - \overline{U}. \tag{A.8}$$

In region 2 of Lemma 2, $a = 0$ and $\Pi^m < 0$. The agent chooses the same efforts for all contracts in region 1, so $\beta^m = 0$ and $v^m = 0$ are optimal in that region. In that case, the principal sets $\alpha^m = (1/6) - \overline{U}$ to extract all the surplus and (ii) holds. We now consider region 3. Substituting (19) into (A.8),

$$\Pi^{m} = -Av^{2} + Bv - C\beta v + D\beta - E\beta^{2} - F, \tag{A.9}$$

where

$$A = \tfrac{2\lambda(1+3\lambda)}{(2+3\lambda)^2} \qquad B = \tfrac{\lambda(8+9\lambda)}{(2+3\lambda)^2}$$
$$C = \tfrac{6\lambda}{(2+3\lambda)^2} \qquad D = \tfrac{2(5+6\lambda)}{(2+3\lambda)^2} \tag{A.10}$$
$$E = \tfrac{12+9\lambda+k(2+3\lambda)^2}{2(2+3\lambda)^2} \quad F = \tfrac{4+3\lambda-3\lambda^2}{2(2+3\lambda)^2} + \overline{U}.$$

Note that A-E are positive and

$$\Pi^m_{\beta v} = -C < 0, \tag{A.11}$$

so incentives and the work ethic are strategic substitutes for the principal. Since

$$\Pi^m_{\beta\beta} = -2E < 0 \tag{A.12}$$

and

$$\Pi^m_{\beta\beta}\Pi^m_{vv} - \left(\Pi^m_{\beta v}\right)^2 = \frac{4\lambda[3+k(1+3\lambda)]}{(2+3\lambda)^2} > 0, \tag{A.13}$$

(A.9) is a negative definite quadratic form in (β, v). The first-order conditions for (A.9) are linear, so (22) follows from straightforward calculations. Evaluating (18) at (22),

$$\beta + \lambda v - (1+\lambda) = -\frac{(2+3\lambda)\,[1+k(2+\lambda)]}{4\,[3+k(1+3\lambda)]} < 0, \tag{A.14}$$

so the unconstrained maximizer (22) is never in region 2. Evaluating (17) at (22),

$$\beta + \lambda v - (1/3) = \frac{(2+3\lambda)[9+k(9\lambda-2)]}{12[3+k(1+3\lambda)]}, \tag{A.15}$$

so $\lambda \leq \lambda^1$ implies (22) is in region 1, where

$$\lambda^1 = (2/9) - (1/k). \tag{A.16}$$

In that case, the constrained solution is on the boundary with region 1 and (ii) is the global optimum. If $\lambda > \lambda^1$ then (22) is in the interior of region 3. Evaluating (A.9) at (22), we obtain Π^m in (24), which is increasing in λ. Since $\lambda^q > \lambda^1$ and $\lambda = \lambda^q$ equates the expressions for expected profits in (24) and (ii), it follows that (22) is the global optimum when $\lambda \geq \lambda^q$. The other expressions in (23) and (24) follow from straightforward substitutions. ∎

Proof of Proposition 4

The agent always chooses $\lambda \geq \lambda^q$ because his material payoff (24) in (i) of Proposition 3 exceeds that in (ii). From (24),

$$M_\lambda^m = \frac{25k^2[3 + k(1 - 3\lambda)]}{8[3 + k(1 + 3\lambda)]^3}, \tag{A.17}$$

so the agent's unique unconstrained maximizer is λ^m in (25). It follows that the agent chooses λ^m when $\lambda^m \geq \lambda^q$ and λ^q otherwise. Since

$$\lambda^m - \lambda^q = \frac{2}{k} - \frac{7}{27}, \tag{A.18}$$

he chooses λ^m when $0 \leq k \leq 54/7$ and λ^q otherwise. The expressions in (26) and (27) follow from straightforward substitutions for λ in Proposition 3. ∎

Acknowledgments

We thank the Department Editor Olav Sorenson for helping us make the paper truly interdisciplinary. An Associate Editor and two anonymous referees made suggestions which led to major changes (and improvements) in the paper. We thank Dan Bernhardt, Ramon Casadesus-Masanell, Dorothea Kübler, Jimmy Walker, Birger Wernerfelt, and audiences at Illinois, the Workshop in Political Theory and Policy Analysis at Indiana University, the IIOC (2009), and the Social Science Research Centre Berlin (WZB) for discussions and comments.

References

Alchian, A.A., H. Demsetz. 1972. Production, information costs, and economic organization. *Amer. Econ. Rev.* 62(5) 777-795.

Akerlof, G.A. 1982. Labor contracts as partial gift exchange. *Quart. J. Econ.* 97(4) 543-569.

Akerlof, G.A. 1983. Loyalty filters. *Amer. Econ. Rev.* 73(1) 54-63.

Akerlof, G.A., R.E. Kranton. 2005. Identity and the economics of organizations. *J. Econ. Persp.* 19(1) 9-32.

Akerlof, G.A., R.E. Kranton. 2008. Identity, supervision, and workgroups. *Amer. Econ. Rev.* 98(2) 212-217.

Akerlof, G.A., R.E. Kranton. 2010. *Identity Economics. How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton University Press, Princeton, NJ.

Baker, G., R. Gibbons, K.J. Murphy. 2002. Relational contracts and the theory of the firm. *Quart. J. Econ.* 117(1) 39-84.

Barnard, C.I. 1968. *The Functions of the Executive.* Harvard University Press, Cambridge, Mass.

Bó, P.D. 2005. Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *Amer. Econ. Rev.* 95(5) 1591-1604.

Bolton, G.E., A. Ockenfels. 2000. ERC: A theory of equity, reciprocity, and competition. *Amer. Econ. Rev.* 90(1) 166-93.

Bolton, P., M. Dewatripont. 2005. *Contract Theory.* MIT Press, Cambridge, Mass.

Bowles, S. 1998. Endogenous preferences: the cultural consequences of markets and other economic institutions. *J. Econ. Lit.* 36(1) 75-111.

Casadesus-Masanell, R. 2004. Trust in agency. *J. Econ. Man. Strat.* 13(3) 375-404.

Callon, M. 1998. Introduction: the embeddedness of economic markets in economics. M. Callon, ed. *The Laws of the Markets.* Wiley-Blackwell, Malden, Mass. 1-57.

Coase, R.H. 1937. The nature of the firm. *Economica* 4(16) 386-405.

Dufwenberg, M., P. Heidhues, G. Kirchsteiger, F. Riedel, J. Sobel. 2008. Other-regarding preferences in general equilibrium. Working paper.

Durkheim, E. 1984. *The Division of Labor in Society.* The Free Press, New York.

Eccles, R.G. 1981. The quasifirm in the construction industry. *J. Econ. Behav. Org.* 2(4) 335-357.

Falk, A., U. Fischbacher. 2006. A theory of reciprocity. *Games Econ. Behav.* 54(2) 293-315.

Fehr, E., U. Fischbacher. 2002. Why social preferences matter – the impact of non-selfish motives on competition, cooperation, and incentives. *Econ. J.* 112(478) C1-C33.

Fehr, E., S. Gächter. 2000. Fairness and retaliation: the economics of reciprocity. *J. Econ. Persp.* 14(3) 159-181.

Fehr, E., A. Klein, K.M. Schmidt. 2007. Fairness and contract design. *Econometrica* 75(1) 121-154.

Fehr, E., K.M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quart. J. Econ.* 114(3) 817-68.

Fehr, E., K.M. Schmidt. 2004. Fairness and incentives in a multi-task principal-agent model. *Scand. J. Econ.* 106(3) 453-474.

Frank, R.H. 1987. If homo economicus could choose his own utility function, would he want one with a conscience? *Amer. Econ. Rev.* 77(4) 593-604.

Frank, R.H., T. Gilovich, D.T. Regan. 1993. Does studying economics inhibit cooperation? *J. Econ. Persp.* 7(2) 159-171.

Giddens, A. 1971. *Capitalism and Modern Social Theory: An Analysis of the Writings of Marx, Durkheim, and Max Weber.* Cambridge University Press, Cambridge, UK.

Granovetter, M. 1985. Economic action and social structure: the problem of embeddedness. *Amer. J. Soc.* 91(3) 481-510.

Granovetter, M. 2005. The impact of social structure on economic outcomes. *J. Econ. Persp.* 19(1) 33-50.

Grossman, S.J., O.D. Hart. 1986. The costs and benefits of ownership: a theory of vertical and lateral integration. *J. Polit. Econ.* 94(4) 691-719.

Harrison, J.R., G.R. Carroll. 2002. The dynamics of cultural influence networks. *Comp. Math. Org. Theory* 8(1) 5-30.

Hart, O., J. Moore. 1990. Property rights and the nature of the firm. *J. Polit. Econ.* 98(6) 1119-58.

Henderson, R. 2008. What do managers do (to build competitive advantage)? The development of relational contracts and the origins of organizational capability. Working paper.

Holmström, B. 1982. Moral hazard in teams. *Bell J. Econ.* 13(2) 324-340.

Holmström, B., P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55(2) 303-328.

Holmström, B., P. Milgrom. 1991. Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *J. Law Econ. Org.* 7(1) 24-52.

Holmström, B., P. Milgrom. 1994. The firm as an incentive system. *Amer. Econ. Rev.* 84(4) 972-991.

Hursthouse, R. 1999. *On Virtue Ethics.* Oxford University Press, Oxford.

Kalleberg, A.L., T. Reve. 1993. Contracts and commitment: economic and sociological perspectives on employment relations. *Human Relations* 45(9) 1103-1132.

Kandel, E., E.P. Lazear. 1992. Peer pressure and partnerships. *J. Polit. Econ.* 100(4) 801-17.

Muller, J.Z. 2003. *The Mind and the Market: Capitalism in Western Thought.* Anchor Books, New York.

Ostrom, E., J. Walker. 2007. Trust and reciprocity as foundations for cooperation: individuals, institutions, and context. Working paper W07-3, Workshop in Political Theory and Policy Analysis, Indiana University.

Ouchi, W.G. 1980. Markets, bureaucracies, and clans. *Admin. Sci. Quart.* 25(1) 129-141.

Podsakoff, P.M., S.B. MacKenzie, J.B. Paine, D.G. Bachrach. 2000. Organizational citizenship behaviors: a critical review of the theoretical and empirical literature and suggestions for future research. *J. Man.* 26(3), 513-563.

Polanyi, K. 2001. *The Great Transformation: the Political and Economic Origins of Our Time.* Beacon Press, Boston, Mass.

Rabin, M. 1993. Incorporating fairness into game theory and economics. *Amer. Econ. Rev.* 83(5) 1281-1302.

Raub, W. 1990. A general game-theoretic model of preference adaptations in problematic social situations. *Rationality and Soc.* 2(1) 67-93.

Rob, R., P. Zemsky. 2002. Social capital, corporate culture, and incentive intensity. *RAND J. Econ.* 33(2) 243-257.

Roberts, J. 2004. *The Modern Firm: Organizational Design for Performance and Growth.* Oxford University Press, New York.

Rotemberg, J.J. 1994. Human relations in the workplace. *J. Polit. Econ.* 102(4) 684-717.

Rotemberg, J.J. 2006. Endogenous altruism in buyer-seller relations and its implications for vertical integration. Working paper.

Seabright, P. 2004. *The Company of Strangers: A Natural History of Economic Life.* Princeton: Princeton University Press.

Simon, H.A. 1951. A formal theory of the employment relationship. *Econometrica* 19(3) 293-305.

Simon, H.A. 1991. Organizations and markets. *J. Econ. Persp.* 5(2) 25-44.

Sobel, J. 2005. Interdependent preferences and reciprocity. *J. Econ. Lit.* 43(2) 392-436.

Sobel, J. 2008. Markets make people appear selfish. Working paper.

Tabellini, G. 2008. The scope of cooperation: values and incentives. *Quart. J. Econ.* 123(3) 905-950.

Van den Steen, E. 2010. Interpersonal authority in a theory of the firm. *Amer. Econ. Rev.* 100(1) 466-490.

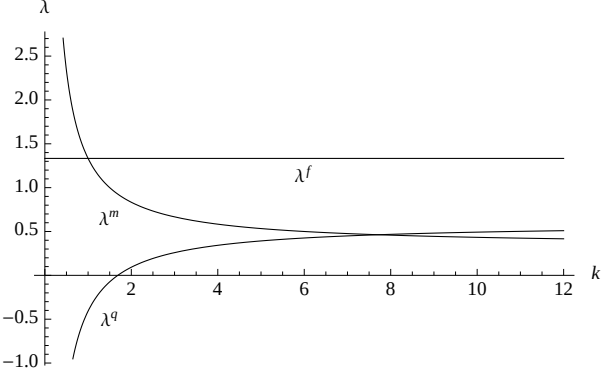
Vohs, K.D., N.L. Mead, M.R. Goode. 2006. The psychological consequences of money. *Science* 314(5802) 1154-1156.

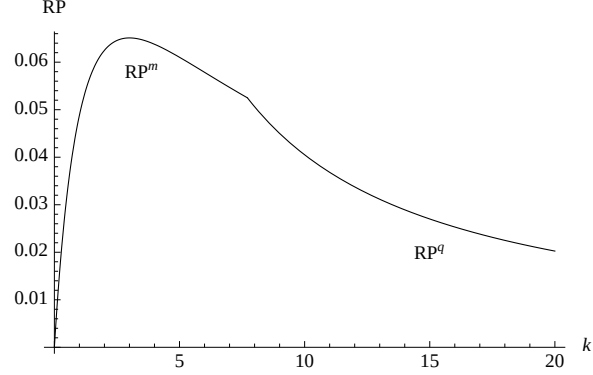
Weber, M. 1978. *Economy and Society: An Outline of Interpretive Sociology.* Roth, G., C. Wittich, eds. Berkeley: University of California Press.

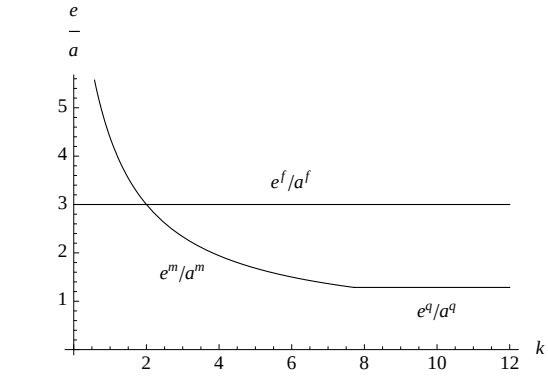
Williamson, O.E. 1975. *Markets and Hierarchies.* Free Press, New York.

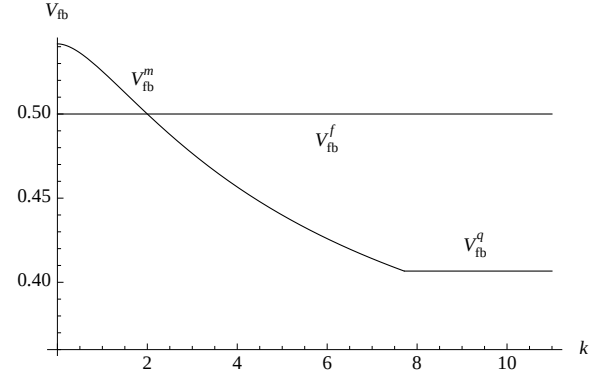
Williamson, O.E. 1981, The economics of organization: the transaction cost approach. *Amer. J. Soc.* 87(3) 548-577.

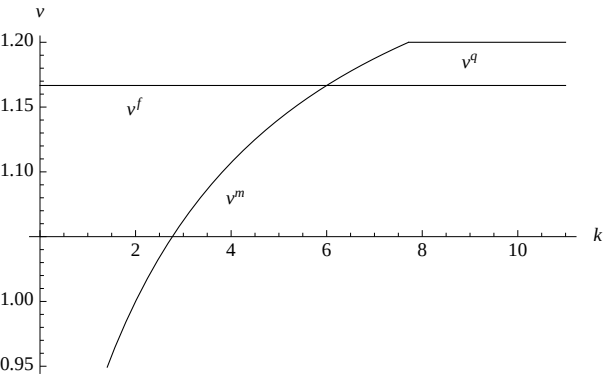
Williamson, O.E. 1985. *The Economic Institutions of Capitalism.* Free Press, New York.

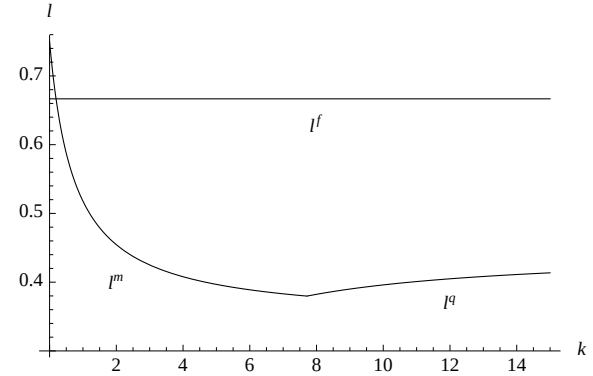

λ
2.5
2.0
1.5
1.0
0.5
−0.5
−1.0
2
4
6
8
10
12
k
λ^f
λ^m
λ^q


RP
0.06
0.05
0.04
0.03
0.02
0.01
5
10
15
20
k
RP^m
RP^q


e/a
5
4
3
2
1
2
4
6
8
10
12
k
e^f/a^f
e^m/a^m
e^q/a^q


V_fb
0.50
0.45
0.40
2
4
6
8
10
k
V_fb^m
V_fb^f
V_fb^q


v
1.20
1.15
1.10
1.00
0.95
2
4
6
8
10
k
v^f
v^m
v^q


l
0.7
0.6
0.5
0.4
2
4
6
8
10
12
14
k
l^f
l^m
l^q