Organizational Purpose and the Dynamics of Motivation*

Timothy Besley, Maitreesh Ghatak, and Linchuan Xu
LSE

April 22, 2025

Abstract

This paper builds a framework for understanding the dynamics of motivation and the missions adopted by profit-maximizing firms. The approach is useful for thinking about the dynamic consequences of adopting missions such as Diversity, Equity, and Inclusion (DEI) or Environmental, Social, and Governance (ESG) goals. By embedding these ideas in a model of cultural evolution via workplace socialization, we explore when, in the long-run, such goals can become consistent with profit-maximization even if they involve pecuniary costs. But the incentives of a firm's owners can be important, with more patient and mission-oriented ownership likely to making a mission sustainable in the long-run. However, we show that there is the possibility of hysteresis, whereby how a firm behaves in its early years can have long-run consequences that are robust to subsequent changes of ownership. Throughout the paper we focus on cases where mission choice is voluntary, but we also discuss the case for regulations that impose requirements to adopt such missions at the firm level.

*We thank Patrick Francois, Robert Gibbons, Jay Lee, Paul Niehaus, Rob Oxoby, and participants in several conferences and seminars for helpful feedback. Besley is grateful for support from the Programme for Cohesive Capitalism at the LSE. Email address: t.besley@lse.ac.uk, m.ghatak@lse.ac.uk, and l.xu25@lse.ac.uk

1 Introduction

Shaping the social purpose of business and other organizations has become hugely controversial in recent times. After a period when such things as Diversity, Equity, and Inclusion (DEI) and Environmental, Social, and Governance (ESG) goals were apparently embraced almost universally, there is now a backlash that is seemingly reaching torrential proportions following the re-election of Donald Trump as President in the US. From an economic point of view, a key question is whether introducing such goals into the workplace increases or reduces economic efficiency and/or reduces profitability. If the latter is true, their adoption will be inherently fragile when they are part of a voluntary code of practice.

Such considerations raise the wider question of whether DEI and ESG goals are incentive-compatible without external enforcement. One important consideration that it is easy to overlook, is whether such social goals serve a dynamic purpose, i.e., whether they are intended to change the workplace, having permanent consequences for the nature of the business and society. Believing in their long-term value makes sense if the values of workers are malleable, are influenced by workplace experiences and are ultimately internalized into workers' preferences. This raises the prospect that exposing them to different purposes in the firm's objectives can lead to alignment between the mission of the firm and its workers, increasing both efficiency and profits.

In this paper, we explore a setting where the workplace becomes an engine of social change by influencing workers' motivation. We develop a model that makes sense of the idea that the motivational capital of the firm is endogenous and can be a source of productivity. Since motivational capital changes over time, the set of incentive compatible social goals that motivate workers rest on how firms and organizations look at the future anticipating how far the motivations of workers change over time. A key insight is that these dynamic considerations also affect how remuneration policies are designed so that the organization of the firm co-evolves with workplace motivation and its purpose.

The model developed in the paper is a dynamic version of Besley and Ghatak (2005).[1] We allow for some workers to be motivated, i.e., are willing

[1]The key mechanism for creating pro-social firms in their framework is sorting. In a similar vein, Henderson and Van den Steen (2015) show that firms with a clear sense of

to work harder when they are in sympathy with the mission of the organization, such as having DEI or ESG goals. Firms incentivize workers but can economize on workplace incentives when workers are aligned with organizational goals. The novel feature of the model is the possibility of cultural evolution within the organization. We describe a form of "Darwinian" dynamics whereby adopting values that increase the payoffs of workers propagate faster. Thus, the organization's remuneration policy and mission choice feeds back into the process of cultural evolution such that workers have an incentive to align their values with the purpose of the firm.

We assume that adopting social goals, whether of the DEI or ESG variety, is costly. Otherwise, it is hard to understand why these goals were not always part of firms' objectives. These costs could in some cases be due to re-engineering the goods or services that organizations produce so that they reflect these goals. But it could also be the cost of hiring compliance officers within the organization to implement these goals and or increasing the number of committee or staff meetings that are costly to the organization. Whatever the source, this creates a trade-off. The firm may take a hit on the bottom line and reduce profit. But if these can be offset by motivational benefits. Moreover, the trade-off is lessened over time making a pro-social goal the more profitable strategy. Reaping these dynamic benefits may require patience as the motivational capital of the firm builds up over time. Otherwise, the dynamic path is fragile. It could even be that introducing a regulation that commits a firm to a more social stance actually increases profit and makes it less likely that it will renege on a commitment to a pro-social stance.

For descriptive purposes, we refer to DEI and ESG as "pro-social" objectives. But we do not prejudge whether it is *socially* optimal to adopt such goals. Such issues are contentious. But we are clear about what must go into a welfare analysis that could resolve this. Any costs incurred by firms to maintain these goals are real and need to be considered. But in a utilitarian framework, creating a workforce that is motivated by such goals can also be a source of direct benefit to workers as well as an indirect benefit by increasing worker effort and firm productivity. We show that profit considerations do not fully internalize the benefits of pro-social goals even without adding

purpose can be more profitable than others by enhancing their workers' sense of identity and reputation, as then employees are willing to accept lower wages in exchange for the meaningful experience of working at such firms.

wider societal benefits. Our framework provides a way of weighing up both static and dynamic welfare effects of such goals in a world where the goals affect the evolution of motivation.

Our paper contributes to the large and growing literature that has been debating the mission of the firm and the role of corporate social responsibility (for example, Bénabou and Tirole, 2010; Kitzmueller and Shimshack, 2012; Kotler and Lee, 2008; Mayer, 2018). Bénabou and Tirole (2010) list three broad reasons why profit-maximizing firms may violate Friedman's famous dictum that the social responsibility of business is to increase its profits (Friedman, 1970): (i) firms could be adopting a longer-term outlook in the presence of externalities, (ii) they could be acting on behalf of external stakeholders, or (iii) they could be influenced by stakeholders within the company such as workers, owners and shareholders.[2] The approach that we take falls under the third category and encompasses a wide range of missions that have been stressed by commentators and economists who study this. For example, it could encompass the ideas in Henderson (2023) which thinks in terms of a more moral approach to doing business based on how customers and workers are treated. We stress that even if, as Friedman suggests, firms are bound to maximize profits, pro-sociality and profit-orientation need not be in conflict, especially in the long run when the dynamics of motivation are considered.

Even though it has not been connected to models of incentives with motivated agents, the importance of workplace socialization is well-understood in more sociological discussions of the workplace. Thus, in a classic account Van Maanen and Schein (1979) say:

> "..organizational socialization refers ... to the fashion-in which an individual is taught and learns what behaviors and perspectives are customary and desirable within the work setting as well as what ones are not." (page 4)

Our approach takes such ideas seriously and studies their implications. From the start, organizational psychologists have emphasized the importance of group dynamics in shaping cultural change (for example, Schein and Schein, 1970). But there is no canonical model of how such socialization processes

[2] In Broccardo et al. (2022) some agents are socially motivated and firms generate externalities. They explore the relative effectiveness of exit and voice options in achieving the socially desirable outcome.

occur and how the dynamics proceed.[3] For this part of the model, we borrow ideas from models of cultural evolution which originated in anthropology (e.g. Boyd and Richerson, 1988; Cavalli-Sforza and Feldman, 1981). Such ideas have been brought into economics by Bisin and Verdier (2001).[4] More generally, our paper is related to an increased interest in how culture matters for economic outcomes (see, for example, Guiso et al., 2006).

The paper also makes a contribution to studying more psychologically informed theories of human motivation as discussed, for example, in Lazear (1991) and Kamenica (2012). The specific approach that we build on is Besley and Ghatak (2005) who study competition and incentives when workers can be motivated by non-pecuniary "mission" goals. The empirical relevance of this approach has now been shown in a variety of contexts. For example, Carpenter and Gong (2016) uses a lab experiment to confirm that motivated workers will produce higher output, and financial incentives can largely substitute for mission motivation when workers and employers are mismatched in mission preferences. In similar vein, Hiller and Verdier (2014) explores how market structure affects firms' investment in corporate culture, i.e. the cultural homogeneity that align workers to firm's objectives and can help to substitute monetary incentives: a larger product market size and higher competition for managers on the labour market induces firms to invest more in corporate culture and reduce financial incentives. Brekke and Nyborg (2008) stresses how social goals in firms can affect the recruitment of highly motivated employees securing socially responsible firms' survival even in a highly competitive environment.

The approach that we take has similarities with the identity-approach of Akerlof and Kranton (2000) who argue that people are moved to act because they associate a particular way of behaving with adopting a particular identity. They stress that identities can change over time and may vary according to location and culture. Such ideas have been influential in the or-

[3]Dessein and Prat (2022) study a model of the firm where the dynamics are driven not by socialization but by organizational capital—an intangible, slow-evolving, and productive asset that requires the direct involvement of the firm's leadership to develop.

[4]Besley and Persson (2024) apply these ideas to studying how decentralizing power in an organization interacts with cultural evolution. There is an emerging literature which looks at these phenomena in political economy applications. For example, Besley and Persson (2019); Bisin and Verdier (2024) study the coevolution of culture and institutions while Besley (2020) applies these ideas to the emergence of the social contract. See Bisin and Verdier (2023) for a recent review of the burgeoning literature.

ganizational sociology literature following on the analysis of bureaucracy in Weber (1922). Related also is the work of Ashraf et al. (2024) who consider factors that give workers a sense of meaning and how this can affect their well-being and productivity using a field experiment. Workplace motivation can also be thought of as intrinsic motivation, developing inherent enjoyment and satisfaction from performing certain tasks. In a well-known experiment (Deci, 1975), college students were either paid or not paid to solve an interesting puzzle, and it was found that those who were not paid spent more time on it and also reported greater interest in the task. Such ideas have been modeled in the literature on incentives by, for example, Bénabou and Tirole (2006, 2003).[5]

The remainder of the paper is organized as follows. The next section develops the core model. Section 3 asks when pro-social firms emerge endogenously even if it is costly to adopt a pro-social mission. We show that there is an important distinction by looking at this dynamically. Section 4 and 5 discuss the findings and return to contemporary debates about DEI and ESG, bringing out the insights that the model offers on these. Section 6 offers some concluding remarks.

2 The Core Model

2.1 Basics

Workers and Firms There is a continuum of workers and a single firm. We suppose that there is high level of firm-specific human capital so that individuals join a firm for life with turnover being due to death or illness. A worker earns a payoff of zero if she leaves the firm. A firm has a workforce size one, comprising a continuum of agents indexed by $i \in [0, 1]$.

There are two types of workers indexed by $\tau \in \{m, s\}$ where m stands for "motivated" and s standards for "selfish", with $\mu \in [0, 1]$ being the fraction of motivated workers.[6] All workers incur a disutility of effort from $e = 1$ denoted by ψ.[7] We assume that $\psi \in \left[0, \bar{\psi}\right]$ with distribution function $F(\psi)$

[5]Besley and Ghatak (2018) reviews a range of relevant literature on pro-social motivation and incentives.

[6]We could extend our framework to consider $\mu \in \left[\underline{\mu}, \bar{\mu}\right]$ where $\underline{\mu} > 0$ and $\bar{\mu} < 1$ so there are upper and low bounds on each type of workers in the population determined by cultural processes outside the model. Our results will not be affected.

[7]The basic model of moral hazard that we use, where both the employer and the

and each worker receives an idiosyncratic draw from the distribution each period, with associated strictly positive density $f(\cdot)$.

Although both types of worker like consumption and dislike effort, motivated workers also care about the the nature of the output produced by the firm, which we refer to as the firm's *mission* denoted by $\sigma \in \{0, 1\}$. We will refer to $\sigma = 1$ as the "prosocial mission". This could be striving to help low-income customers, caring for the environment or diversifying the workforce. To capture this formally, let $v_\tau(\sigma)$, with $\tau \in \{s, m\}$, be a non-pecuniary pay-off that can partially offset the disutility of effort where for selfish workers, $v_s(\sigma) = 0$ for $\sigma \in \{0, 1\}$, while for motivated workers:

$$v_m(\sigma) = \begin{cases} \theta & \sigma = 1 \\ -\varepsilon & \text{otherwise.} \end{cases} \tag{1}$$

The assumption that motivated workers earn some disutility when the firm sets $\sigma = 0$ reflects the distaste that they feel from working for a for-profit firm when it does not pursue a pro-social mission.

Worker utility is linear in both consumption and the cost of effort:

$$U^\tau(z, e) \equiv z + e\left[v_\tau(\sigma) - \psi\right], \quad \tau \in \{s, m\} \tag{2}$$

where z is private consumption.

Output Each worker chooses whether to put in effort $e \in \{0, 1\}$, and individual output is $x(e)$ where $x(0) = 0$ for all $i \in [0, 1]$, and

$$x(1) = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p. \end{cases}$$

This says that, conditional on setting $e = 1$, the *expected* output of a worker is p.[8] We assume that the output realization of workers are independent. Let λ be the proportion of agents in the organization who set $e = 1$.[9] Expected

workers are risk neutral and there is a limited liability constraint is the same as in Besley and Ghatak (2005) and is based on Innes (1990), which has been widely used in the context of financial contracting, managerial incentives, labor markets, and tenancy.

[8] It would be straightforward to introduce the possibility that a worker can produce some baseline output even with low effort without affecting any of the main results.

[9] Since there is a continuum of workers, formally we have

$$\lambda = \left(\int_0^1 e(i)\, di\right).$$

aggregate firm output is $X(\lambda) = \lambda p$. The firm earns a revenue of y per unit of output net of the cost of non-labor inputs.

Worker Effort Since output per worker is either 0 or 1, and realizations are *i.i.d.* across workers, we can focus without loss of generality on wage contracts that have a flat wage component $\omega \geq 0$ and a bonus component $\beta \in [0, y]$ for high output. The optimal choice of effort for an agent with disutility draw ψ is:

$$e\left(\beta p + v_\tau(\sigma) - \psi\right) = \arg\max_{e\in\{0,1\}} \left\{e\left[\beta p + v_\tau(\sigma) - \psi\right]\right\}. \tag{3}$$

This defines a cutoff level of $\hat{\psi}$ below which an agent chooses $e = 1$:

$$\hat{\psi}_\tau(\beta, \sigma) = \begin{cases} \beta p + v_\tau(\sigma) & \tau = m \\ \beta p & \tau = s. \end{cases}$$

Average effort across the firm is given by:

$$\hat{\lambda}(\sigma, \mu, \beta) \equiv \mu F\left(\beta p + v_m(\sigma)\right) + (1-\mu) F(\beta p). \tag{4}$$

Expected output is therefore $\hat{\lambda}(\sigma, \mu, \beta)\, p$. Notice that $\hat{\lambda}(\sigma, \mu, \beta)$ is always increasing in β as we would expect. Higher output can also be achieved by setting $\sigma = 1$ via its impact on motivated agents. For $\sigma = 1$, $v_m(\sigma) = \theta$ and $\hat{\lambda}$ is increasing in μ and θ. Otherwise, $\hat{\lambda}$ is decreasing in μ.

We make the following regularity assumption on the distribution function $F(.)$:

Assumption R: *For all $\psi \in \mathbb{R}$: (i) $F(\psi)$ is a log concave distribution; (ii) $F'(0)$ is bounded.*

This holds for many standardly used distributions such as the uniform, exponential and Pareto. Crucially for the analysis that follows, this implies that $h(\psi) = \frac{F'(\psi)}{F(\psi)}$ is a decreasing function of ψ.

Following Judd (1985), we assume that this integral is well-defined. This will be the case if we use a Pettis integral (Al-Najjar, 2004).

Firms Firms cannot observe the type of each worker and therefore have to offer the same employment contract, $\{\beta, \omega, \sigma\}$ to all workers. The mission choice of the firm, σ, is specified in the contract. Firms observe the output produced by a worker, $x(e)$, but not the individual effort choices of workers. Profit per worker is therefore

$$(y - \beta)\, ep - \omega, \tag{5}$$

and expected total profit of the firm, aggregating over all workers, is given by:

$$[y - \beta]\, \lambda p - \sigma c - \omega. \tag{6}$$

where we are assuming that adopting a prosocial mission is costly. This captures the idea that modifying products and production processes to make a good or service that conforms to a pro-social mission requires human and material resources.

Below, we will study the optimal determination of $\{\beta, \omega, \sigma\}$ depending on the payoffs of the firm's management.

Worker Turnover and Socialization Each period an exogenous fraction, ρ, of workers in the firm leave and are replaced. We assume that all workers who enter the firm are selfish but can be socialized into being motivated workers by the existing cadre of workers with some fraction becoming motivated workers. The socialization process depends on the material "fitness advantage" of being a motivated worker (which could be positive or negative).[10] We have in mind a situation in which workers are initially "impressionable" and subject to peer influence during their first period in the firm, learning the culture and its associated costs and benefits. The influence of the stock of motivated and selfish workers depends on the material benefit from belong to each group. Thus, it is the payoffs when they enter the firm that drive the socialization process. Once socialization has taken place, this does not change for the duration of their careers.

The expected payoff of an agent of type τ given the employment contract on offer is:

$$Y^{\tau}(\beta, \omega, \sigma) = F\left(p\beta + v_{\tau}(\sigma)\right)\left[p\beta + v_{\tau}(\sigma)\right] + \omega,\ \tau = m, s. \tag{7}$$

[10]The framework that we use for this is similar to the forward-looking socialization models with overlapping (or sequential) generations in Bisin and Verdier (2001), Tabellini (2008) and Besley and Persson (2023).

The material fitness advantage of the motivated type is then defined as:

$$\Delta = Y^{m}(\beta,\omega,\sigma) - Y^{s}(\beta,\omega,\sigma). \tag{8}$$

This can be positive or negative depending on the material and psychological payoffs as well as the mission choice of the firm.

The fraction of motivated workers is assumed to evolve in line with the relative payoffs of motivated and selfish types. Specifically:

$$\frac{\mu_{t+1}-\mu_t}{\mu_t} = \rho\kappa(1-\mu_t)\Delta \tag{9}$$

for some constant $\kappa > 0$. The Appendix shows that this can be viewed as a first order approximation to micro-founded socialization model based on contact and influence within the firm. The workforce becomes more prosocially motivated over time if and only $\Delta > 0$. Higher turnover of workers speeds up transition as there are more workers whose type can be change as long as there is material benefit from being pro-socially motivated. The growth rate is convergent as it is decreasing in μ_t.

2.2 Model Timing and Equilibrium

Time is infinite and indexed by t. Firms are assumed to be unable to commit to future employment contracts with timing as follows:

1. At the beginning of each period, an organization inherits a fraction of motivated workers μ_t.
2. The firm chooses organizational form, $\sigma \in \{0,1\}$, and a wage contract wage contract $\{\beta,\omega\}$.
3. Agents choose their effort level, $e \in [0,1]$
4. Output and payoffs are realized.
5. A fraction ρ of workers are replaced, new workers are socialized according to Δ, and μ_{t+1} is determined.

We study Markov perfect equilibrium in which the state variable μ_s evolves over time. The problem has a recursive structure with the only state variable being μ_s. Incentive contracts $\left\{\hat{\omega}_\sigma(\mu), \hat{\beta}_\sigma(\mu)\right\}$ depend on σ given μ. We then determine $\hat{\sigma}(\mu_t)$ optimally which depends on μ_s. And the evolution of μ depends on these choices.

2.3 Incentive Contracts

Suppose that incentives are set to maximize short term profits:

$$\hat{\Pi}(\mu, \sigma) = \max_{\{\omega,\beta\}} \left\{ [y - \beta] \hat{\lambda}(\sigma, \mu, \beta) p \right\} - \omega. \quad (10)$$

It is clear that profits are decreasing in ω hence it is optimal to set the fixed wage as low as as possible. Henceforth, we will therefore set $\omega_\sigma(\mu) = 0$ for $\sigma \in \{0, 1\}$ and $\mu \in [0, 1]$. Hence, the only optimization decision for firms is what level of bonus β to choose. This involves balancing the marginal cost of providing incentives in terms of lower profit against the incentive benefits from rewarding agents for high output. Using (4), and assuming an interior solution, the first-order condition for β, define $B_\sigma(\mu)$ from:

$$\frac{1}{y - B_\sigma(\mu)} = \frac{\partial \log\left(\mu F\left(B_\sigma(\mu) p + v_m(\sigma)\right) + (1-\mu) F\left(B_\sigma(\mu) p\right)\right)}{\partial \beta}. \quad (11)$$

These vary with the mission choice so there will be two different bonus levels depending on whether the firm chooses the prosocial mission.

We have the following characterization of optimal bonuses $\left\{ \tilde{\beta}_0(\mu), \tilde{\beta}_1(\mu) \right\}$:

Proposition 1 *Suppose that Assumption R holds. Optimal bonuses are then as follows:*

1. *With a for-profit mission,* $\tilde{\beta}_0(\mu) = B_0(\mu) > 0$ *and increasing in* μ.
2. *With a pro-social mission*

$$\tilde{\beta}_1(\mu) = \begin{cases} B_1(\mu) > 0 & \text{if } \theta < h^{-1}(\frac{1}{py}) \text{ or } \mu \leq \frac{pyF'(0)}{F(\theta) - py[F'(\theta) - F'(0)]} \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, $B_1(\mu)$ *is decreasing in* μ. *For all* $\theta > 0$, $\tilde{\beta}_1(\mu)$ *is (weakly) decreasing in* θ *and when the firm chooses a pro-social mission* ($\sigma = 1$) *incentives are flatter, i.e.,* $\tilde{\beta}_0(\mu) > \tilde{\beta}_1(\mu) \geq 0$.

This results shows that agent motivation and financial incentives are substitutes and so using bonuses is less attractive, all else equal, in a world of motivated workers as in Besley and Ghatak (2005).[11] However, whether there is a pro-social mission is chosen by the firm. We now turn to this.

[11]The Proposition also says that if $\theta \geq h^{-1}(\frac{1}{py})$ and $\mu \geq \frac{pyF'(0)}{F(\theta) - py[F'(\theta) - F'(0)]}$), then bonuses in pro-social firms are zero, i.e. a flat wage. However, even if they are not actually zero, they will be lower in firms that adopt the pro-social mission.

2.4 Mission Choice

The choice of σ is made in each period but the objective of the firm is to maximize the discounted present value of profits. Hence:

$$\hat{\sigma}\left(\mu_s\right) = \arg\max_{\sigma\in\{0,1\}} \left\{ \left[\hat{\Pi}\left(\mu_s, \sigma_t\right) - c\sigma_s\right] + \sum_{t=s}^{\infty} \gamma^{t+1-s} \left[\hat{\Pi}\left(\mu_t, \sigma_t\right) - c\sigma_t\right] \right\}. \tag{12}$$

A key observation is that profits vary with μ_t. This implies that a forward-looking owner must consider how μ_t will evolve in future due to socialization and how the mission choice will be made over time. Even though it cannot choose the future mission, equation (12) requires that the current owner *predicts* the future time path of mission choices when it chooses the mission in period s. When we study the optimal mission choice below, we will require that the predicted future mission choices must be time-consistent, i.e. driven by the profit-maximizing choice at each future date.

2.5 Socialization

To study the dynamics of motivation, we have to understand how the relative material payoffs of the two types of workers evolves. To this end, we can write down the fitness advantage (or disadvantage) of the motivated type as

$$\hat{\Delta}\left(\mu\right) = \begin{cases} \left[F\left(p\beta_1\left(\mu\right)+\theta\right) - F\left(p\beta_1\left(\mu\right)\right)\right] p\beta_1\left(\mu\right) > 0 & \sigma = 1 \\ \left[F\left(p\beta_0\left(\mu\right)-\varepsilon\right) - F\left(p\beta_0\left(\mu\right)\right)\right] p\beta_0\left(\mu\right) < 0 & \text{otherwise.} \end{cases} \tag{13}$$

This is just the material payoff difference between being a motivated and selfish worker. It is a function of whether the organization chooses a pro-social mission.

Given a contract $\left\{\hat{\omega}_{\hat{\sigma}(\mu_t)}\left(\mu_t\right), \hat{\beta}_{\hat{\sigma}(\mu_t)}\left(\mu_t\right), \hat{\sigma}\left(\mu_t\right)\right\}$, we have the following expression for the expected material payoff of each type as a function of the mission in place:

$$Y^{\tau}\left(\mu_t\right) = F\left(p\hat{\beta}_{\hat{\sigma}(\mu_t)}\left(\mu_t\right) + v_\tau\left(\hat{\sigma}\left(\mu_t\right)\right)\right) p\left[\hat{\beta}_{\hat{\sigma}(\mu_t)}\left(\mu_t\right) v_\tau\left(\hat{\sigma}_t\right)\right] + \hat{\omega}_{\hat{\sigma}(\mu_t)}\left(\mu_t\right), \ \tau = m, s,$$

Using this, we can then write down the process that governs the dynamics of motivation as:

$$\mu_{t+1} = \mu_t\left[1 + \rho\kappa\left(1-\mu_t\right)\hat{\Delta}\left(\mu_t\right)\right] = H\left(\mu_t\right) \tag{14}$$

where $\hat{\Delta}(\mu) = Y^m(\mu) - Y^s(\mu)$. Thus dynamic path of σ_t is given by (12) when μ_t follows (14).

It is clear from (14), that pro-social motivation among the stock of workers will expand (contract) between t and $t+1$ as long as $\hat{\Delta}(\mu_t)$ is strictly positive (negative).

3 The Emergence of Pro-social Firms

One of the core issues in the adoption of DEI and ESG goals is how these can be sustained in profit-maximizing firms. One route is to have external regulations or changes in internal governance arrangements that enforce them. Our dynamic model creates an important subtlety because of dynamically evolving changes in motivation since what is optimal in the short-run may diverge from the long-run optimal mission of the firm. We now explore when pro-sociality becomes self-enforcing dynamically given the dynamics of motivation and requiring that mission choices at each future date are time-consistent.

We begin with the observation that a *necessary* condition for a pro-social firm to emerge is that profits would be higher if *all* workers were motivated. This is because, conditional on adopting a pro-social mission, profits are higher when more workers are motivated to work in pro-social firms. Formally, this is captured with the following assumption, which we will refer to as the *viability* condition:

Assumption V:

$$p\left[F\left(p\tilde{\beta}_1(1)+\theta\right)\left[y-\tilde{\beta}_1(1)\right]-F\left(p\tilde{\beta}_0(1)-\varepsilon\right)\left[y-\tilde{\beta}_0(1)\right]\right]>c \tag{15}$$

Naturally enough, this condition will hold as long as c is small enough, i.e. costs of a pro-social mission are not too high and/or θ is large enough, the motivational benefits from pro-social workers who work with a pro-social mission are large enough.

Given the viability condition, we can now explore the optimal path of mission choices that maximizes long run profits for any value of μ_0, i.e. solve (12). We have the following core result for the firm's optimal mission choice which shows that beyond some initial stock of motivated worker, the social mission will become entrenched when the viability and regularity conditions hold:[12]

Proposition 2 *Suppose Assumptions R and V hold. There exists $M \in (0,1)$ such that always choosing $\sigma = 1$ will be the optimal path of firm mission choice if and only if $\mu_0 \geq M$. Otherwise, the firm always chooses $\sigma = 0$. Furthermore, if $\mu_0 \geq M$, then $\lim_{t\to\infty} \mu_t = 1$ while if $\mu_0 < M$, then $\lim_{t\to\infty} \mu_t = 0$.*

This says that, in spite of the underlying dynamics of motivation, the optimal mission choice is stationary, i.e. the current and any future owner of the firm will pick the same mission. Which will be chosen boils down to whether choosing a pro-social mission from period 0 will yield a higher value for the discounted profits for the firm. This depends, in turn on the initial stock of pro-social workers; when the initial stock μ_0 is high enough, the firm will always choose a pro-social mission $\sigma = 1$, and the share of motivated workers will increase along the dynamic path, i.e. $\mu_{t+1} > \mu_t$. The limit of this process is that all workers will become the pro-social type. And, if the firm never chooses a pro-social mission, then all workers will eventually become selfish even if they are motivated to start with.[13]

The proof of the result is somewhat involved but the key to understanding this result is recognizing the complementarity between having a higher stock of motivated workers and profits once a pro-social mission has been adopted. So if being pro-social at t, then *a fortiori* being pro-social will be optimal forever more. Moreover, this means that the firm will have even more motivated workers at every date in future.

The same forces imply a form of hysteresis since the initial stock of pro-social workers matters to the long-run outcome. Thus, the firm effectively becomes locked into a particular mission which depends on this. An implication of this is that the optimal mission choice does not directly depend on whichever the long term profits $\hat{\Pi}(1,1) - c$ and $\hat{\Pi}(0,0)$ are larger. The firm may end up choosing $\sigma = 0$ even though profits from a social mission are

[12]The proof of this proposition is in the appendix.

[13]This stark result could be avoided if we were to assume that there are some workers who cannot be influenced by the socialization process and so remain selfish or motivated regardless of what they are exposed to in the workplace. In this case, in the limit there would be convergence to an outcome where only these non-malleable types remain in one of the categories.

higher in the long run, i.e. the per period profit $\hat{\Pi}(1,1) - c$ is larger than $\hat{\Pi}(0,0)$ (so that the firm should have chosen $\sigma = 1$).

This observation makes it important to think about where μ_0 comes from. We can think of this depending on the societal fractions of motivated workers, due to how they are brought up by parents, influenced by educational institutions and their peers. In this sense, the kind of pro-social businesses that might emerge become a function of the wider pool of workers. Of course, if firms have a selection process at their inception which can identify pro-social workers from among the pool of applicants, this will help to sustain a pro-social business. But we would expect the capacity to do that to be depending on the prevalence of pro-social workers in the talent pool.

Our framework provides an interesting link between the organization of the firm and the cultural context in which it is operating, something that has been stressed in studies such as Hofstede (2001). Our analysis contributes two things, first to show why there may be a link between culture and how firms behave. But, going beyond this, we are stressing that this may have a dynamic component with a feedback loop onto the culture of the firm, something which does not seem have been stressed so much in the literature on how cultural differences matter but is a key feature of dynamic models of cultural evolution (Boyd and Richerson, 1988; Cavalli-Sforza and Feldman, 1981) who stress that culture is dynamically evolving.

It is straightforward to link the critical threshold value of the initial stock to the fundamentals of the model which we do in:

Corollary 1 *The threshold value M is decreasing in θ and ϵ, and increasing in c.*

This corollary links the predictions of the model to drivers of the profit functions: $\hat{\Pi}(1,1) - c$ and $\hat{\Pi}(0,0)$. The initial stock of motivated workers needed to create a pro-social firm, μ_0, is larger if the motivated workers obtain a higher non-pecuniary payoff θ when $\sigma = 1$ is chosen or higher disutility ϵ if instead $\sigma = 0$, or the cost of setting up the pro-social mission c is lower.

This finding stresses that technological fundamentals still matter when it comes to the adoption of pro-social mission by firms and organizations. If there is little scope for creating motivation towards a pro-social objective of the firm and/or it is too costly to for the bottom line, creating a culture that supports such objectives will be an uphill struggle and even doomed to failure.

4 Does Capitalist Motivation Matter?

We have focused on the case where long-run profits are the driver of the firm or organization's objectives. This makes sense for a private ownership economy. We have shown that, even then with such motives, a model where motivation is dynamic can support a pro-social business model without apology. But does that mean founders of firms, and their owners more generally, have no role to play in the emergence of pro-social firms? We now give two examples of when the preferences of owners matter. The first stresses the distinction between patient and impatient capital while the second allows for mission-oriented owners.

4.1 Patient Capitalists

One of the arguments that has raged about the nature of capitalism, particularly shareholder capitalist is that it can encourage the pursuit of short-term goals at the expense of long-term goals even when profits are the objective (for example, Bebchuk and Stole, 1993). We show that such debates are especially relevant in our framework where an evolving asset of the firm is the cultural capital of the workforce which is more likely to be built when the firm is more patient.

To see this, observe that it is implicit in (12) which governs the firm's mission choice that the patience of the owners of the firm matters since there can be an underlying trade-off between short term and long-term profit. Even if a firm cannot make more profit today, it may have faith in its capacity to socialize a workforce that will ultimately make it profitable to have a pro-social mission. This patience is a virtue when it comes to creating pro-social firms and organizations in a world where there is a dynamic evolution of motivation. We now show that this intuitive argument is formally correct. the following results shows how γ affects the mission choice:

Proposition 3 *Suppose Assumptions R and V hold. Then there exists a* $\hat{\mu}$ *such that* $\hat{\Pi}(\mu, 1) - c \geq \hat{\Pi}(\mu, 0)$ *if and only if* $\mu \geq \hat{\mu}$*. Moreover, the threshold value* M *depends on* γ *as follows:*

(i) $\lim_{\gamma\to 0} M = \hat{\mu}$*;*

(ii) If $\hat{\Pi}(0, 0) > \hat{\Pi}(1, 1) - c$*,* $\lim_{\gamma\to 1} M = 1$*. Furthermore, for sufficiently large* γ*,* M *is increasing in* γ*.*

(iii) If $\hat{\Pi}(0,0) < \hat{\Pi}(1,1) - c$, $\lim_{\gamma\to 1} M = 0$. *Furthermore, for sufficiently large* γ, M *is decreasing in* γ.

The first part says that even an impatient principal will choose $\sigma = 1$ if μ is high enough. The second part shows how the choice depends on γ. To understand this intuitively, note that the long term (latent) profits depends on whether $\hat{\Pi}(1,1) - c$ or $\hat{\Pi}(0,0)$ is higher. If the firm is sufficiently patient, such long term (latent) profits will play a larger role (i.e. the firm cares enough about future profits). Then if $\hat{\Pi}(0,0) > \hat{\Pi}(1,1) - c$, the initial threshold M will be higher if γ is higher, i.e. the firm is willing to allow a higher threshold level of μ_0 to choose $\sigma = 0$. At the extreme case $\gamma = 1$, firm will always choose $\sigma = 0$ regardless of μ_0. On the other hand, if $\hat{\Pi}(0,0) < \hat{\Pi}(1,1) - c$, then the initial threshold M will be lower if γ is higher, i.e. the firm will tolerate a lower μ_0 to choose $\sigma = 1$. When $\gamma = 1$, the firm will always choose $\sigma = 1$ starting from period 0 in order to obtain the latent long term profit $\hat{\Pi}(1,1) - c$. If the firm is relatively impatient, its mission choice will be more dependent on short term profits and, *in extremis* with $\gamma = 0$, the firm mission choice reduces to a problem of maximizing only current period profits, and the threshold M should converge to $\hat{\mu}$.

To summarize, this result illustrates how patience of the firm owner affects mission choice in our framework, i.e. by changing the threshold level of the initial stock of pro-sociality.

4.2 Motivated Capitalists

Although we have so far worked with the case where the owner is profit-oriented, another way to think about how pro-social firms could emerge is to believe that founders of such firms have pro-social objectives.[14] One of the most famous examples that earned the founder a Nobel Prize was Grameen Bank where the aim was to give credit market opportunities to those who had been excluded. But the dynamic consequences of such founders is less clear. When the founders of Ben and Jerry's Ice Cream, another widely cited example of motivated founders, sold out to Unilever, there was cynicism about the depth of such motivation and whether the more pro-social business model would ultimately be abandoned.

[14]This is similar in spirit to the idea of "stake-holder" capitalism (see, for example, Fleurbaey and Ponthière, 2023; Zingales and Hart, 2022).

We now see how having a motivated founder can have a difference both in the short-run and long-run.[15] To model this very simply, suppose that the founder receives a utility of $\sigma\Theta$ from choosing a pro-social mission which offsets the cost c, then the optimal path of mission choice is given by

$$\max_{\{\sigma_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \gamma^t \left[\hat{\Pi}(\mu_t, \sigma_t) - (c - \Theta)\sigma_t \right].$$

The observation that this would influence in a pro-social direction in a static world is trivial. It is akin to a reduction of c, which then leads to lower M by Corollary 1 above which yields:

Proposition 4 *There exists $\hat{\Theta}$ such that $M < \hat{\mu}$ if and only if $\Theta > \hat{\Theta}$.*

The dynamic implications of this finding are, however, much more interesting if one thinks about the possibility of a firm being sold by its founder to someone who does not share their commitment to pro-sociality. In the case of Ben and Jerry's ice-cream to Unilever, the big question was whether this would lead to a change in the business model driven by a desire to cut costs. Analytically, a change of ownership can be thought of as reset of the value of μ_0 in Proposition 2 as the new owner will inherit a stock of motivated workers which could depend on the mission set by the founder. Then we have the possibility that a firm run by a motivated founder will have built up a motivated workforce which guarantees that the mission persist even if the new owner does not share their direct preference for a pro-social mission. This legacy effect of motivated founders can therefore be a source of hysteresis in this dynamic setting. So an optimistic view of a case like the Ben and Jerry's takeover would be that it happened at a point where the pro-social mission was safe for the period beyond their ownership of the firm because they had built up a motivated workforce.[16]

5 Well-being and Welfare

We now consider two further implications of the model that requires us to think more deeply about the well-being of workers and the welfare of having

[15] As in Oehmke and Opp (2025) this could also come from social pressure from an external investor.

[16] Of course, such effects would be reinforced by sorting into pro-social firms alongside workplace socialization.

a pro-social firm.

Well-being Versus Materialism We have taken a materialistic view of fitness as a driver of motivation. So workers look at the material success of their co-workers when socialization is on-going. We now consider what happens when the dynamics of motivation is driven instead by comparisons of well-being across worker types. The main difference now is that utility includes the cost of effort associated with striving to achieve rewards not just the payoffs.

In this case (13) is replaced by

$$\Delta(\mu) = \begin{cases} \nu_1(\mu) & \text{if } \sigma = 1 \\ \nu_0(\mu) & \text{if } \sigma = 0, \end{cases} \tag{16}$$

where

$$\nu_1(\mu) = \int_0^{p\hat{\beta}_1(\mu)} \theta dF(\psi) + \int_{p\hat{\beta}_1(\mu)}^{p\hat{\beta}_1(\mu)+\theta} \left[p\hat{\beta}_1(\mu) + \theta - \psi\right] dF(\psi) > 0$$

and

$$\nu_0(\mu) = -\varepsilon F\left(p\hat{\beta}_0(\mu) - \varepsilon\right) - \int_{p\hat{\beta}_0(\mu)-\varepsilon}^{p\hat{\beta}_0(\mu)} \left[p\hat{\beta}_0(\mu) - \psi\right] dF(\psi) < 0.$$

So the core dynamics are essentially the same in this case and the model exhibits the complementarity that drives the dynamics. Which model of motivation dynamics is more realistic is moot. Arguably, material remuneration is more visible to co-workers than their psychological state but the satisfaction levels of different workers could still be picked up in social interactions. Fortunately, the results are robust to whichever is used.

Welfare Implications We now consider the implications for welfare of having more motivated agents in an organization. The exact welfare criterion to use is not entirely clear when motivation is changing. But, in the previous sub-section, we have derived an expression for worker welfare rather than just material rewards. We will use this to to work with a social surplus criterion which adds worker welfare and profits.

It is debatable whether there should be additional considerations in a welfare analysis if society also values objectives such as DEI and ESG in firms. If the mission is genuinely pro-social and creates a benefit to society of $S > c$, then there can be an overwhelming case for having more workers choose this. Of course, in such cases, it would be natural for $\sigma = 1$ to be imposed by regulation rather the leaving it to voluntary decisions. We will set aside such considerations for now and study welfare only within an organization.

We have supposed that there is a utility from being motivated of θ so worker's well-being is higher and since they choose their effort optimally, this must still be true when the cost of effort is considered. Hence, the fact that workers have higher well-being means that having pro-social firms must increase the welfare of workers. The firm will also only pick $\sigma = 1$ if it is more profitable for it to do so. To explore long run steady state welfare, we consider, the total surplus in the firm, i.e. the sum of profits and worker utility.

It is straightforward to compute welfare when $\sigma = 1$ and $\sigma = 0$ and compare the two. The result is given in:

Proposition 5 *Suppose Assumptions R and V hold. The motivated steady state will deliver higher welfare than the selfish steady state if and only if*

$$\overline{c} \equiv \hat{\Pi}(1,1) - \hat{\Pi}(0,0) + \int_{p\hat{\beta}_0(0)}^{p\hat{\beta}_1(1)+\theta} F(\psi)d\psi > c, \tag{17}$$

where $\overline{c} > 0$ *and is increasing in* θ*.*

This says that whenever firm profit at the steady state of $\mu = 1$ is larger than that at steady state $\mu = 0$, i.e. $\hat{\Pi}(1,1) - c > \hat{\Pi}(0,0)$, the welfare of the former steady state will also be higher than that of the latter. Otherwise the motivated steady state will still deliver higher welfare as long as $c < \overline{c}$.

However, there is not full congruence between what a firm chooses to do and what maximizes welfare. This is not surprising since the firm does not internalize gains in worker welfare when it makes its mission choice. In the case of Proposition 3 where the firm is sufficiently patient, i.e. $\gamma \to 1$, then there is congruence between welfare and what the firm chooses. So if $\hat{\Pi}(1,1) - c > \hat{\Pi}(0,0)$, then a patient enough firm will always choose the pro-social mission regardless of the initial stock of motivated workers μ_0 and the pro-social mission also delivers higher welfare.

But the exact divergence between the decentralized case and welfare optimum is however somewhat more subtle as can be seen in the case of Proposition 3 where $\hat{\Pi}(1,1) - c < \hat{\Pi}(0,0)$. To see this, suppose first of all that

$$\hat{\Pi}(1,1) - \hat{\Pi}(0,0) < c \leq \hat{\Pi}(1,1) - \hat{\Pi}(0,0) + \int_{p\hat{\beta}_0(0)}^{p\hat{\beta}_1(1)+\theta} F(\psi)d\psi.$$

For such intermediate range of c, the pro-social steady state still delivers higher welfare than the selfish one. However, for the sufficiently patient firm, it will not now choose the pro-social mission (regardless of the initial stock of motivated workers μ_0) because of its long run profit $\hat{\Pi}(1,1) - c < \hat{\Pi}(0,0)$. Thus the profit-maximizing and forward-looking firm is on a dynamic path that results in a payoff below the social optimum. Mandating $\sigma = 1$ can then increase welfare.

Finally, consider the case where c is sufficiently high and/or θ is low, i.e.

$$c > \hat{\Pi}(1,1) - \hat{\Pi}(0,0) + \int_{p\hat{\beta}_0(0)}^{p\hat{\beta}_1(1)+\theta} F(\psi)d\psi.$$

A sufficiently patient firm, according to Proposition 3, choose $\sigma = 0$ from the beginning (again, regardless of the initial condition of μ), and this will lead to the selfish steady state which is congruent with the welfare optimum. This finding shows that mandates that force DEI or ESG policies on unwilling firms are not necessarily optimal at least in the case where the government cares about the stakeholders in the firm – shareholders and workers. However, this would change if S is extremely high relative to c. But even then this would almost certainly require something other than a blanket mandate and would have to tailored to the specific benefits and costs.

But note also that temporary regulations based on welfare calculations may be sufficient since once firms have cross the threshold at which pro-social motivations and profits are aligned, they regulation can be relaxed as the pro-social mission will be consistent with profit maximization. Just how long this will take depends of course on the strength of the dynamic socialization process that we have described.

6 Concluding Comments

We have analyzed a specific way of bending the arc of capitalism by recognizing that the missions of firms and the motivation of workers can be intertwined in important ways. Pro-social objectives in firms can have dynamic consequences and there are sources of hysteresis that models of cultural evolution are well-placed to articulate. The paper feeds into a wider idea that the motivations of workers (and firms' reactions to them) are shaped by their exposure to workplace practices. But this is not something that organizational economics seems to have built into its understanding of incentives. But the factors that shape motivation can be responsive to incentives to the extent that motivation is malleable.

Our model of workplace cultures is organization-specific. But it would be interesting to connect this to wider themes around how capitalist *systems* work and, in particular, the influence of the market system on citizens' sensibilities (Bowles, 2016; Hirschman, 1992; Polanyi, 1944). Although these ideas have a long history, but have not been fully embraced in modern mainstream economics. For example Adam Smith had the notion of a *commercial society* which recognized the endogeneity of motivation. But the nexus of institutions which shape motivation is wider than just work. A famous example is Bowles et al. (1976) which discusses the role of schooling in this process. Extending our approach to allow for a wider context and the possibility of spillovers across organizations would be allow a more systematic analysis of the dynamics of motivation at a societal level.

The paper has focused on intra-firm dynamics of motivation. The questions of how these issues play out in a world of competition for workers is an important topic for the future as are the implications of product market competition which has been studied in a related context by Dewatripont and Tirole (2024). They argue that there could be a trade-off between intense competition and creating pro-social missions in firms.

The analysis allows us to reflect on current debates on DEI and ESG where the pressure to impose such requirements are currently being relaxed. The model presented here emphasizes that the impact of doing so will likely be heterogeneous according the way that workplace cultures have already evolved as such measures have been introduced in the past. If an organization is far along the dynamic path that we have outlined, then the effect could be minimal as it could already be optimal for firms to maintain such missions. If not, they could put an organization on a different dynamic path. Either way, having a way of framing these issues as a problem of dynamic cultural evolution enriches the frame that can be useful for engaging with such debates.

References

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The quarterly journal of economics*, 115(3):715–753.

Al-Najjar, N. I. (2004). Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47(1):1–35.

Ashraf, N., Bandiera, O., Minni, V., and Zingales, L. (2024). Meaning at work. *Unpublished Working Paper, London School of Economics.*

Bebchuk, L. A. and Stole, L. A. (1993). Do short-term objectives lead to under-or overinvestment in long-term projects? *The Journal of Finance*, 48(2):719–729.

Bénabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *The review of economic studies*, 70(3):489–520.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American economic review*, 96(5):1652–1678.

Bénabou, R. and Tirole, J. (2010). Individual and corporate social responsibility. *Economica*, 77(305):1–19.

Besley, T. (2020). State capacity, reciprocity, and the social contract. *Econometrica*, 88(4):1307–1335.

Besley, T. and Ghatak, M. (2005). Competition and incentives with motivated agents. *American economic review*, 95(3):616–636.

Besley, T. and Ghatak, M. (2018). Prosocial motivation and incentives. *Annual Review of Economics*, 10(1):411–438.

Besley, T. and Persson, T. (2019). Democratic values and institutions. *American Economic Review: Insights*, 1(1):59–76.

Besley, T. and Persson, T. (2023). The political economics of green transitions. *The Quarterly Journal of Economics*, 138(3):1863–1906.

Besley, T. and Persson, T. (2024). Organizational dynamics: culture, design, and performance. *The Journal of Law, Economics, and Organization*, 40(2):394–415.

Bidner, C. and Francois, P. (2011). Cultivating trust: Norms, institutions and the implications of scale. *The Economic Journal*, 121(555):1097–1129.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319.

Bisin, A. and Verdier, T. (2023). Advances in the economic theory of cultural transmission. *Annual Review of Economics*, 15(1):63–89.

Bisin, A. and Verdier, T. (2024). On the joint evolution of culture and political institutions: Elites and civil society. *Journal of Political Economy*, 132(5):1485–1564.

Bowles, S. (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.

Bowles, S., Gintis, H., et al. (1976). *Schooling in capitalist America*, volume 57. New York: Basic Books.

Boyd, R. and Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago press.

Brekke, K. A. and Nyborg, K. (2008). Attracting responsible employees: Green production as labor market screening. *Resource and Energy Economics*, 30(4):509–526.

Broccardo, E., Hart, O., and Zingales, L. (2022). Exit versus voice. *Journal of Political Economy*, 130(12):3101–3145.

Carpenter, J. and Gong, E. (2016). Motivating agents: How much does the mission matter? *Journal of Labor Economics*, 34(1):211–236.

Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach*. Number 16. Princeton University Press.

Deci, E. (1975). *Intrinsic Motivation*. New York, Platinum Press.

Dessein, W. and Prat, A. (2022). Organizational capital, corporate leadership, and firm dynamics. *Journal of Political Economy*, 130(6):1477–1536.

Dewatripont, M. and Tirole, J. (2024). The morality of markets. *Journal of Political Economy*, 132(8):2655–2694.

Fleurbaey, M. and Ponthière, G. (2023). The stakeholder corporation and social welfare. *Journal of Political Economy*, 131(9):2556–2594.

Friedman, M. (1970). of business is to increase its profits. *New York Times Magazine, September*, 13:122–126.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic perspectives*, 20(2):23–48.

Henderson, R. (2023). Moral firms? *Dædalus*, 152(1):198–211.

Henderson, R. and Van den Steen, E. (2015). Why do firms have "purpose"? the firm's role as a carrier of identity and reputation. *American Economic Review*, 105(5):326–330.

Hiller, V. and Verdier, T. (2014). Corporate culture and identity investment in an industry equilibrium. *Journal of Economic Behavior & Organization*, 103:93–112.

Hirschman, A. O. (1992). *Rival views of market society and other recent essays*. Harvard University Press.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of economic theory*, 52(1):45–67.

Judd, K. L. (1985). The law of large numbers with a continuum of iid random variables. *Journal of Economic theory*, 35(1):19–25.

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annu. Rev. Econ.*, 4(1):427–452.

Kitzmueller, M. and Shimshack, J. (2012). Economic perspectives on corporate social responsibility. *Journal of economic literature*, 50(1):51–84.

Kotler, P. and Lee, N. (2008). *Corporate social responsibility: Doing the most good for your company and your cause*. John Wiley & Sons.

Lazear, E. P. (1991). Labor economics and the psychology of organizations. *Journal of Economic perspectives*, 5(2):89–110.

Mayer, C. (2018). *Prosperity: Better business makes the greater good*. Oxford University Press.

Oehmke, M. and Opp, M. M. (2025). A theory of socially responsible investment. *Review of Economic Studies*, 92(2):1193–1225.

Polanyi, K. (1944). *The Great Transformation*. Boston: Beacon Press.

Schein, E. H. and Schein, E. H. (1970). *Organizational psychology*. Prentice-Hall Englewood Cliffs, NJ.

Tabellini, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, 123(3):905–950.

Van Maanen, J. E. and Schein, E. H. (1979). Toward a theory of organizational socialization. *Research in organizational behavior*, 1:209–264.

Weber, M. (1922). *Economy and society: An outline of interpretive sociology*. University of California press.

Zingales, L. and Hart, O. (2022). The new corporate governance. Technical report, CEPR Discussion Papers.

Appendix

A Microfoundations of Evolution

We assume that the evolutionary dynamics is "Darwinian" in the sense that the increase in the proportion of motivated agents is driven by their fitness advantage.[17] For our analysis we assume that there is a well-behaved function $Q(\mu, \Delta)$ that is *increasing* in Δ (and can depend on μ in a number of ways including being independent of it) :

$$\mu_{t+1} - \mu_t = Q(\mu_t, \Delta(\mu_t)). \tag{18}$$

We will work with a specific formulation which delivers such dynamics.[18]

All newly hired agents are assumed to be selfish but can be socialized on arrival by being mentored by an existing worker chosen at random. If she is mentored by a motivated agent, which happens with probability μ_t, we assume that she may become motivated depending on the relative psychological fitness of motivated and selfish types. In other words, socialization is based on material rewards received by the two types of agents. Moreover, we assume that this is something that can be observed by workers within an organization and it is also comparable across types, i.e. money has similar worth to both motivated and selfish types.

A randomly selected new agent is matched with an existing agent who is motivated with probability μ_t. Such an agent becomes motivated through mentoring by a motivated agent if:

$$\Delta(\mu_t) + \eta \geq 0,$$

where η is a mean-zero, *symmetrically* distributed idiosyncratic shock with a continuous distribution function $G(\cdot)$. We assume that $G(0) = \frac{1}{2}$, $G(\Delta(\mu_t)) > \frac{1}{2}$for $\Delta(\mu_t) > 0$, and $G'(\Delta(\mu_t)) > 0$ for all $\Delta(\mu_t)$.[19]

Let $g(\cdot)$ be the density function corresponding to $G(\cdot)$. The probability that a new recruit mentored by a motivated type becomes motivated is the

[17] From a philosophical point of view, this assumes that it is possible to compare utility across types for the same individual. In particular, we assume that an individual can figure out what their utility would be, if they were of a different type.

[18] Note that we have μ_{t+1} depending on μ_t and so this is a form of adaptive expectations where the fitness is measured for the contemporaneous value of μ_t.

[19] An example would be the logistic distribution where the probability of a randomly

probability that $\eta \geq -\Delta(\mu_t)$, which is $1 - G(-\Delta(\mu_t))$. Given the symmetry assumption, this is equal to $G(\Delta(\mu_t))$.

Despite being matched with an existing agent who is motivated, if such direct socialization fails, the new recruit may still be indirectly socialized by observing and learning from other workers.[20] The probability of indirectly becoming a motivated type depends monotonically on the average fraction of such types in the organization, a kind of social learning postulated in much of the cultural-evolution literature (Bisin and Verdier, 2001). Assuming a linear relation, the probability of indirect socialization becomes $(1 - G(\Delta(\mu_t)))\,\mu_t$ where μ_t is the fraction of motivated agents in the existing workforce at the beginning of period and $1 - G(\Delta(\mu_t))$ is the fraction of new agents for whom $\eta < -\Delta(\mu_t)$.

Adding these expressions, the overall probability that a new recruit who is matched with a motivated agent becomes motivated is:

$$G(\Delta(\mu_t)) + \{1 - G(\Delta(\mu_t))\}\mu_t. \tag{19}$$

If a new worker is matched with and mentored by a selfish worker instead, which happens with probability $1 - \mu_t$, there are two possibilities. First, she can be socialized into being selfish if

$$\Delta(\mu_t) + \eta \leq 0.$$

Thus, $G(-\Delta(\mu_t)) = 1 - G(\Delta(\mu_t))$ is the proportion of selfish workers coming from such matches. Second, she can indirectly become motivated (as above) depending on the aggregate fraction of motivated agents (μ_t) in the organization. The resulting probability of becoming motivated is therefore:

$$G(\Delta(\mu_t))\,\mu_t. \tag{20}$$

The probability of a new agent being matched with a motivated or selfish agent being μ_t and $1 - \mu_t$, multiplying (19) by μ_t and (20) by $1 - \mu_t$, and

selected new agent to become motivated through mentoring is:

$$G(\Delta(\mu_t)) = \frac{\exp[\Delta(\mu_t)]}{1 + \exp[\Delta(\mu_t)]}.$$

It is easy to verify that the listed properties are satisfied.

[20]This parallels the approach taken in Bidner and Francois (2011).

adding, we get the fraction of new agents who become motivated agents. The overall fraction of motivated agents in the next period, μ_{t+1}, is therefore:

$$\mu_{t+1} = (1-\rho)\,\mu_t + \rho\left[2\mu_t\,(1-\mu_t)\,G\,(\Delta\,(\mu_t)) + \mu_t^2\right]$$

Simplifying the resulting expressions yields the following the equation of motion for the share of motivated types:

$$\mu_{t+1} = \mu_t + \rho\mu_t\,(1-\mu_t)\,[2G\,(\Delta\,(\mu_t)) - 1]\,. \tag{21}$$

From this it is clear that studying the evolutionary dynamics of motivation requires studying the properties of $\Delta\,(\mu_t)$, in particular, its sign and how it changes with respect to μ_t.

Following Besley and Persson (2023), we now consider a linearized approximation of (21) around the value $\hat{\mu}$, i.e. where $\Delta\,(\hat{\mu}) = 0$. Then we have

$$\frac{\mu_{t+1} - \mu_t}{\mu_t} \simeq \rho\,(1-\hat{\mu})\,[g\,(\Delta\,(\hat{\mu}))]\,\Delta\,(\mu_t) = \kappa(1-\mu_t)\Delta\,(\mu_t)\,.$$

The approximation is exact when $G\,(\cdot)$ is a uniform distribution.

B Proof of Proposition 1

We begin the analysis with a preliminary result which is if $F\,(\psi)$ is logconcave then so is $\mu F\,(\beta p + v_m\,(\sigma)) + (1-\mu)\,F\,(\beta p)$. To see this, note that one way to characterize a function $F(x)$ that is logconcave is

$$\log F\,(\alpha x_0 + (1-\alpha)\,x_1) \geq \alpha \log F(x_0) + (1-\alpha) \log F(x_1).$$

Now consider the function

$$H(x) = \lambda \log F(x) + (1-\lambda) \log F(x+a).$$

We show that if $F(x)$ is logconcave then so is this function. We can write

$$\begin{aligned}
& \log H\left(\alpha x_0+(1-\alpha) x_1\right) \\
= & \log \left[\lambda \log F(\alpha x_0+(1-\alpha) x_1)+(1-\lambda) \log F(\alpha x_0+(1-\alpha) x_1+a)\right] \\
\geq & \log [\lambda[\alpha \log F(x_0)+(1-\alpha) \log F(x_1)]+ \\
& (1-\lambda)[\alpha \log F(x_0+a)+(1-\alpha) \log F(x_1+a)]] \\
= & \log [\alpha[\lambda \log F(x_0)+(1-\lambda) \log F(x_0+a)]+ \\
& (1-\alpha)[\lambda \log F(x_1)+(1-\lambda) \log F(x_1+a)]] \\
\geq & \alpha \log [\lambda \log F(x_0)+(1-\lambda) \log F(x_0+a)] \\
& +(1-\alpha) \log [\lambda \log F(x_1)+(1-\lambda) \log F(x_1+a)] \\
= & \alpha \log H(x_0)+(1-\alpha) \log H(x_1)
\end{aligned}$$

The first inequality follows from the assumption that $F(x)$ is logconcave, and the second inequality follows from the fact that $\log(x)$ is strictly concave.

Turning to the proof of our result, recall the first order condition yields

$$(y-\beta)\lambda'-\lambda=0.$$

We can check that the second order condition holds by plugging in the first order condition:

$$\frac{\partial^2 \hat{\Pi}}{\partial \beta^2}=(y-\beta)\lambda''-2\lambda'=\frac{1}{\lambda'}[\lambda''\lambda-2\lambda'^2].$$

We know this is negative from log concavity of λ in β.

To show the comparative static results, we can obtain from the first order condition that

$$\frac{\partial \beta(\mu, v)}{\partial \mu}=-\frac{(y-\beta) \frac{\partial^2 \log \lambda(v, \mu, \beta)}{\partial \beta \partial \mu}}{\frac{\partial^2 \log \lambda(v, \mu, \beta)}{\partial \beta^2}-\frac{\partial \log \lambda(v, \mu, \beta)}{\partial \beta}},$$

and

$$\frac{\partial \beta(\mu, v)}{\partial v}=-\frac{(y-\beta) \frac{\partial^2 \log \lambda(v, \mu, \beta)}{\partial \beta \partial v}}{\frac{\partial^2 \log \lambda(v, \mu, \beta)}{\partial \beta^2}-\frac{\partial \log \lambda(v, \mu, \beta)}{\partial \beta}}.$$

From log concavity of λ, we know $\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial \beta^2}<0$. What remains is to determine the signs of $\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial\mu}$ and $\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial v}$. We have

$$\begin{aligned}
\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial\mu} &= \frac{p}{\lambda^2}[[F'(\beta p+v)-F'(\beta p)][\mu F(\beta p+v)+(1-\mu)F(\beta p)] \\
&\quad -[F(\beta p+v)-F(\beta p)][\mu F'(\beta p+v)+(1-\mu)F'(\beta p)]] \\
&= \frac{p}{\lambda^2}\left[F'(\beta p+v)F(p\beta)-F'(\beta p)F(\beta p+v)\right].
\end{aligned}$$

Hence $\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial\mu} < 0$ if and only if

$$\frac{F'(\beta p+v)}{F(\beta p+v)} < \frac{F'(\beta p)}{F(p\beta)}.$$

From the log concavity of $F(.)$, we know $\frac{F'(\beta p+v)}{F(\beta p+v)}$ is decreasing in v. Hence the above inequality holds if and only if $v > 0$. Therefore, when $v > 0$, $\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial\mu} < 0$, and when $v < 0$, $\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial\mu} > 0$. This implies that when $v > 0$, $\beta_1(\mu, v)$ is decreasing in μ, and when $v < 0$, $\beta_0(\mu, v)$ is increasing in μ.

To show $\beta_1(\mu, v)$ is decreasing in v (as $\sigma = 1$, in this case $v = \theta > 0$), again we examine

$$\begin{aligned}
\frac{\partial^2 \log \lambda(v,\mu,\beta)}{\partial\beta\partial v} &= \frac{p}{\lambda^2}[[\mu F''(\beta p+v)][\mu F(\beta p+v)+(1-\mu)F(\beta p)] \\
&\quad -[\mu F'(\beta p+v)][\mu F'(\beta p+v)+(1-\mu)F'(\beta p)]] \\
&\leq \frac{p\mu}{\lambda^2}[\frac{F'(\beta p+v)^2}{F(\beta p+v)}[\mu F(\beta p+v)+(1-\mu)F(\beta p)] \\
&\quad -[F'(\beta p+v)][\mu F'(\beta p+v)+(1-\mu)F'(\beta p)]] \\
&= \frac{p\mu}{\lambda^2}F'(\beta p+v)(1-\mu)F(\beta p)\left[\frac{F'(\beta p+v)}{F(\beta p+v)}-\frac{F'(\beta p)}{F(\beta p)}\right] < 0,
\end{aligned}$$

given log-concavity of $F(.)$, since $v = \theta > 0$. Then it follows that $\beta_1(\mu, v)$ is decreasing in v.

To show that $\beta_0(\mu, v) > 0$ for all $\mu \in [0, 1]$, note first that $\beta_0(\mu, v) \geq \beta_0(0, v) = \beta_0(0)$. From the first order condition, we know that $\beta_0(0)$ is characterized by

$$\frac{1}{y-\beta} = p\frac{F'(\beta p)}{F(\beta p)}$$

As $\beta \to 0$, the left hand side goes to $\frac{1}{y}$ and the right hand side goes to infinity given that $F(0) = 0$. By Assumption 1 $F'(0)$ is bounded. As $\beta \to y$, the left hand side goes to infinity, and the right hand side goes to a constant, i.e. $p\frac{F'(py)}{F(py)} > 0$. Hence there must exist a unique $\beta \in (0, y)$ that solves the above equation.

To show that $\beta_1(\mu, v) > 0$ for all $\mu \in [0, 1]$, again notice that $\beta_1(\mu, v) \geq \beta_1(1, \theta)$. From the first order condition, we know that $\beta_1(1, \theta)$ is characterized by

$$\frac{1}{y - \beta} = p\frac{F'(\beta p + \theta)}{F(\beta p + \theta)}$$

Similar to the argument above, as β goes from 0 to y, the left hand side of the above equation increases from $\frac{1}{y}$ to infinity, whereas the right hand side decreases from $p\frac{F'(\theta)}{F(\theta)}$ to $p\frac{F'(py+\theta)}{F(py+\theta)}$. Hence the above equation yields a positive solution to $\beta_1(1, \theta)$ if and only if

$$\frac{1}{py} < \frac{F'(\theta)}{F(\theta)}.$$

Denote $h(a) = \frac{F'(a)}{F(a)}$ and from the log-concavity of $F(.)$, the above inequality is equivalent to

$$\theta < h^{-1}\left(\frac{1}{py}\right).$$

Hence we know

$$\begin{aligned} \beta_0(0) \leq \beta_0(\mu) &\leq \beta_0(1) \\ \beta_1(1) \leq \beta_1(\mu) &\leq \beta_1(0). \end{aligned}$$

Observe that $\beta_0(0) = \beta_1(0)$. To see this, from the first order condition, we take $\mu = 0$ and we know that $\beta_0(0)$ and $\beta_1(0)$ are characterized by the same equation:

$$\frac{1}{y - \beta_\sigma(\mu)} = \frac{\partial \log(F(p\beta_\sigma(\mu)))}{\partial \beta}.$$

Therefore we have $\beta_1(\mu) \leq \beta_1(0) = \beta_0(0) \leq \beta_0(\mu)$.

In our analysis we so far we assumed that motivated agents are not too motivated to the point that the firm might give no incentives at all, i.e., $\beta_1 = 0$. Suppose θ can be large enough so that, in principle, the firm can set $\hat{\beta}_1(\mu) = 0$. To extend our analysis we can write down the Lagrangian for the firm's maximization problem (by choosing $\sigma = 1$) as

$$L = (y - \beta)p\lambda + \delta\beta,$$

where δ is the Lagrangian multiplier. The first order condition can be now rewritten as

$$(y-\beta)p[\mu F'(p\beta+\theta)+(1-\mu)F'(p\beta)]-[\mu F(p\beta+\theta)+(1-\mu)F(p\beta)]+\delta=0.$$

We would also require

$$\delta \geq 0, \beta \geq 0, \delta\beta = 0.$$

We have discussed the interior solution for which $\delta = 0$ (i.e. limited liability constraint is slack). Now we consider the case where $\delta > 0$, which implies $\beta = 0$. Then from the first order condition $\delta > 0$ is equivalent to

$$\mu > \hat{\mu} = \frac{pyF'(0)}{F(\theta) - py[F'(\theta) - F'(0)]}.$$

One can further show $\hat{\mu} \leq 1$ if and only if

$$\theta \geq h^{-1}\left(\frac{1}{py}\right),$$

where $h(a) = \frac{F'(a)}{F(a)}$. Then we have if $\theta \geq h^{-1}\left(\frac{1}{py}\right)$,

$$\hat{\beta}_1(\mu) = \begin{cases} 0 & \text{if } \mu > \hat{\mu}; \\ \beta_1(\mu) & \text{if } \mu \leq \hat{\mu}. \end{cases}$$

If instead $\theta < h^{-1}\left(\frac{1}{py}\right)$, $\hat{\beta}_1(\mu) = \beta_1(\mu)$. We then define $B_0(\mu) \equiv \beta_0(\mu)$ and $B_1(\mu) \equiv \beta_1(\mu)$. This completes the proof for Proposition 1.

C Proof of Proposition 2

The proof has three steps:

Step 1 We show that the profit function is convex in μ. We first show the convexity of $\hat{\Pi}(\mu,\sigma)$ in μ. By envelope theorem, we have

$$\frac{\partial\hat{\Pi}(\mu,1)}{\partial\mu} = (y-\beta)p[F(p\beta+\theta) - F(p\beta)] > 0$$

Given $\beta_1(\mu)$ is the optimal bonus under $\sigma = 1$, we have

$$\frac{\partial^2\hat{\Pi}(\mu,1)}{\partial\mu^2} = [-[F(p\beta_1+\theta) - F(p\beta_1)] + (y-\beta_1)[F'(p\beta_1+\theta) - F'(p\beta_1)]p]\, p\frac{\partial\beta_1(\mu)}{\partial\mu}$$

From the first order condition characterizing β_1, we know that $(y-\beta_1)p[\mu F'(p\beta_1 + \theta) + (1-\mu)F'(p\beta_1)] = \mu F(p\beta_1+\theta) + (1-\mu)F(p\beta_1)$. Then we know that we could rewrite

$$\frac{\partial^2\hat{\Pi}(\mu,1)}{\partial\mu^2} = [F(p\beta_1) - p(y-\beta_1)F'(p\beta_1)]\,\frac{1}{\mu_1}\frac{\partial\beta_1(\mu)}{\partial\mu}$$

We then know that

$$\frac{1}{y-\beta_1} < p\frac{F'(p\beta_1)}{F(p\beta_1)}.$$

To see this, note that we have already established that $\beta_1(1)$ is characterized by $\frac{1}{y-\beta_1(1)} = p\frac{F'(p\beta_1(1)+\theta)}{F(p\beta_1(1)+\theta)}$. We have also established that $\frac{F'(p\beta+\theta)}{F(p\beta+\theta)} < \frac{F'(p\beta)}{F(p\beta)}$ as long as $\theta > 0$. We also know that the right hand side of the above inequality is decreasing in β_1, and the left hand side is increasing in β_1. Then it follows that for any $\beta_1(\mu)$,

$$\frac{1}{y-\beta_1(\mu)} < \frac{1}{y-\beta_1(1)} = p\frac{F'(p\beta_1(1)+\theta)}{F(p\beta_1(1)+\theta)} < p\frac{F'(p\beta_1(1))}{F(p\beta_1(1))} < p\frac{F'(p\beta_1(\mu))}{F(p\beta_1(\mu))}.$$

Since we have also established that $\frac{\partial\beta_1(\mu)}{\partial\mu} < 0$, then we know that $\frac{\partial^2\hat{\Pi}(\mu,1)}{\partial\mu^2} > 0$, i.e. $\hat{\Pi}(\mu,1)$ is convex in μ.

Similarly, we show $\hat{\Pi}(\mu,0)$ is convex in μ. Again, by envelope theorem, we have

$$\frac{\partial\hat{\Pi}(\mu,0)}{\partial\mu} = (y-\beta)p[F(p\beta-\epsilon) - F(p\beta)] < 0.$$

We further evaluate

$$\frac{\partial^2\hat{\Pi}(\mu,0)}{\partial\mu^2} = [-[F(p\beta_0-\epsilon)-F(p\beta_0)]+(y-\beta_0)p[F'(p\beta_0-\epsilon)-F'(p\beta_0)]]p\frac{\partial\beta_0(\mu)}{\partial\mu}$$

By the same logic, we substitute in the first order condition characterizing β_0, we have

$$\frac{\partial^2 \hat{\Pi}(\mu,0)}{\partial \mu^2} = [F(p\beta_0) - p(y-\beta_0)F'(p\beta_0)]\frac{1}{\mu_0}\frac{\partial \beta_0(\mu)}{\partial \mu}.$$

We then know that for any $\beta_0(\mu)$

$$\frac{1}{y-\beta_0(\mu)} > \frac{1}{y-\beta_0(0)} = p\frac{F'(p\beta_0(0))}{F(p\beta_0(0))} > p\frac{F'(p\beta_0(\mu))}{F(p\beta_0(\mu))}.$$

Since we have also established $\frac{\partial \beta_0(\mu)}{\partial \mu} > 0$, then we know $\frac{\partial^2 \hat{\Pi}(\mu,0)}{\partial \mu^2} > 0$, i.e. $\hat{\Pi}(\mu,0)$ is also convex in μ.

We can also show that $\hat{\Pi}(\mu,1) - \hat{\Pi}(\mu,0)$ is increasing in μ. This simply follows from the fact that $\hat{\Pi}(\mu,1)$ is increasing in μ and $\hat{\Pi}(\mu,0)$ is decreasing in μ.

Step 2 We now show if at any time T such that for all future time period $t > T$, the same σ is being chosen, i.e. $\sigma_{T+1} = \sigma_{T+2} = ... = \sigma_t = \sigma_{t+1} = ...$, then it is optimal to choose the same σ at time T. To simplify notation, let us denote $\Pi(\mu,\sigma) = \hat{\Pi}(\mu,\sigma) - c\sigma$. Without loss of generality, let us suppose $\sigma_{T+1} = \sigma_{T+2} = ... = 0$, and we show it is optimal that $\sigma_T = 0$ as well.

Suppose not, i.e. the firm chooses $\sigma_T = 1$ instead, which implies that

$$\Pi(\mu_T,1)+\gamma\Pi(\mu_{T+1},0)+\gamma^2\Pi(\mu_{T+2},0)+... \geq \Pi(\mu_T,0)+\gamma\Pi(\tilde{\mu}_{T+1},0)+\gamma^2\Pi(\tilde{\mu}_{T+2},0)+... \tag{22}$$

However, at period $T+1$, we know

$$\Pi(\mu_{T+1},0) + \gamma\Pi(\mu_{T+2},0) + ... \geq \Pi(\mu_{T+1},1) + \gamma\Pi(\hat{\mu}_{T+2},0) + ... \tag{23}$$

Note that (22) implies

$$\Pi(\mu_T,1)-\Pi(\mu_T,0) \geq \gamma[\Pi(\tilde{\mu}_{T+1},0)-\Pi(\mu_{T+1},0)]+\gamma^2[\Pi(\tilde{\mu}_{T+2},0)-\Pi(\mu_{T+2},0)]+...$$

and (23) implies

$$\gamma[\Pi(\mu_{T+2},0) - \Pi(\hat{\mu}_{T+2},0)] + ... \geq \Pi(\mu_{T+1},1) - \Pi(\mu_{T+1},0).$$

Since we know $\mu_{T+1} > \mu_T$, it then follows $\Pi(\mu_{T+1},1) - \Pi(\mu_{T+1},0) > \Pi(\mu_T,1) - \Pi(\mu_T,0)$. We then obtain:

$$\begin{aligned} & \gamma[\Pi(\mu_{T+2},0) - \Pi(\hat{\mu}_{T+2},0)] + ... \geq \Pi(\mu_{T+1},1) - \Pi(\mu_{T+1},0) \\ > \; & \Pi(\mu_T,1) - \Pi(\mu_T,0) \geq \gamma[\Pi(\tilde{\mu}_{T+1},0) - \Pi(\mu_{T+1},0)] + \gamma^2[\Pi(\tilde{\mu}_{T+2},0) - \Pi(\mu_{T+2},0)] + ... \end{aligned}$$

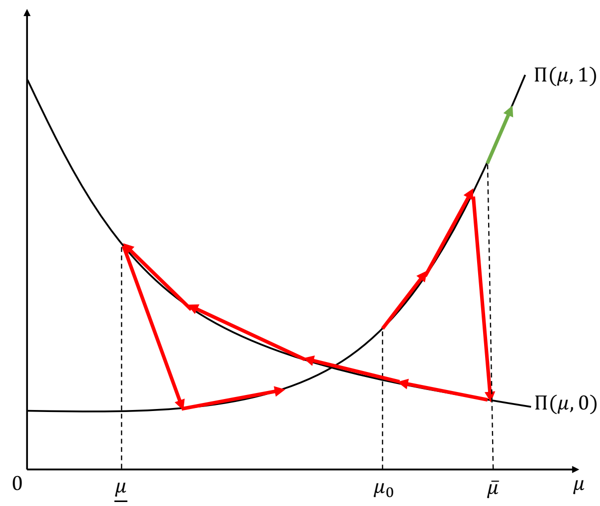

Figure 1: The "Shifting" Strategy of Mission Choice

To see this is a contradiction, note first that

$$\hat{\mu}_{T+2} > \mu_{T+1} > \mu_{T+2} > \tilde{\mu}_{T+1},$$

then by convexity of $\Pi(\mu, 0)$, we know that $\Pi(\hat{\mu}_{T+2}, 0)+\Pi(\tilde{\mu}_{T+1}) > \Pi(\mu_{T+1}, 0)+\Pi(\mu_{T+2}, 0)$, which then implies that

$$\Pi(\mu_{T+2}, 0) - \Pi(\hat{\mu}_{T+2}, 0) < \Pi(\tilde{\mu}_{T+1}, 0) - \Pi(\mu_{T+1}, 0).$$

Following the similar logic, we know

$$\gamma[\Pi(\mu_{T+2}, 0)-\Pi(\hat{\mu}_{T+2}, 0)]+... < \gamma[\Pi(\tilde{\mu}_{T+1}, 0)-\Pi(\mu_{T+1}, 0)]+\gamma^2[\Pi(\tilde{\mu}_{T+2}, 0)-\Pi(\mu_{T+2}, 0)]+....$$

Hence a contradiction.

Step 3 From Step 2 we know that if in the future, a firm finds it optimal to stick to a mission choice, then the firm will also stick to such choice in all previous periods. We now examine the case where the firm may always shift between different mission choices, i.e. $\lim_{t\to\infty} \sigma_t$ does not exist. Denote the set of μ_t that is generated by such path of σ_t choice as $\mathbf{M}$. Then we define $\overline{\mu} = \sup \mathbf{M} \leq 1$ and $\underline{\mu} = \inf \mathbf{M} \geq 0$ (intuitively, these are the upper bound and lower bound of the μ_t path). Then it is easy to see that such strategy profile will be dominated by the following strategy: if $\Pi(\overline{\mu}, 1) > \Pi(\underline{\mu}, 0)$, then when $\overline{\mu}$ is reached (say, at time s)[21], the firm instead choose all the future σ to be 1, because all the future per period profit will be $\Pi(\mu_{s+1}, 1) > \Pi(\overline{\mu}, 1)$; similarly if $\Pi(\overline{\mu}, 1) \leq \Pi(\underline{\mu}, 0)$, then when $\underline{\mu}$ is reached, the firm instead choose all the future σ to be 0. Figure 1 captures the intuition of this step, where the "butterfly" pattern of always shifting between different mission choice (the red arrows) will be dominated by the path where when the highest possible μ is reached, the firm stick to $\sigma = 1$ instead for all the future periods (the green arrow). It is easy to see that such strategy profile should always give higher payoffs than the shifting strategy. Then from Step 2, we know that the shifting strategy should be dominated by the firm sticking to one mission choice from the beginning.

Therefore the dynamic optimization problem reduces to comparing whether always choosing $\sigma = 1$ or $\sigma = 0$ will yield higher discounted payoff in period 0. Suppose the initial condition for μ is μ_0, then define

$$\Pi^1(\mu_0) = \Pi(\mu_0, 1) + \gamma\Pi(\mu_1^1, 1) + \gamma^2\Pi(\mu_2^1, 1) + ...$$

and

$$\Pi^0(\mu_0) = \Pi(\mu_0, 0) + \gamma\Pi(\mu_1^0, 0) + \gamma^2\Pi(\mu_2^0, 0) + ...$$

We know $\Pi^1(\mu_0)$ is continuous and increasing in μ_0 and therefore $\frac{\Pi(0,1)}{1-\gamma} \leq \Pi^1(\mu_0) \leq \frac{\Pi(1,1)}{1-\gamma}$, and $\Pi^0(\mu_0)$ is continuous and decreasing in μ_0 and therefore $\frac{\Pi(1,0)}{1-\gamma} \leq \Pi^0(\mu_0) \leq \frac{\Pi(0,0)}{1-\gamma}$.[22] It easy to see $\Pi(1, 1) > \Pi(1, 0)$ by Assumption V, and $\Pi(0, 0) > \Pi(0, 1)$ (because $\hat{\Pi}(0, 1) = \hat{\Pi}(0, 0)$). Then there exists M such that always choosing $\sigma = 1$ will be the optimal path of firm mission choice if and only if $\mu_0 > M$; otherwise the firm should be always choosing $\sigma = 0$.

To see why Assumption V implies $\Pi(1, 1) > \Pi(1, 0)$, examine when $\sigma = 1$,

$$\hat{\Pi}(\mu, 1) = \max_{\beta \geq 0} \{[y - \beta] p [\mu F(p\beta + \theta) + (1 - \mu) F(p\beta)] - c\}$$

and with $\sigma = 0$, it is

$$\hat{\Pi}(\mu, 0) = \max_{\beta \geq 0} \{(y - \beta) p [\mu F(p\beta - \varepsilon) + (1 - \mu) F(p\beta)]\}.$$

[21] If $\sup M$ could never be reached, find s such that μ_s is close enough to $\sup \mathbf{M}$ such that $\mu_{s+1} = H(\mu_s, 1) > \overline{\mu}$.

[22] Notice that by the structure of $H(\mu_t)$, if μ_0 is either 0 or 1, then μ_t will always stay at such initial value.

If $\mu = 1$:

$$\begin{aligned}\hat{\Pi}(1,1) &= \left[F\left(p\hat{\beta}_1(1)+\theta\right)\right]p\left[y-\hat{\beta}_1(1)\right] \\ &\geq F\left(p\hat{\beta}_0(1)+\theta\right)p\left[y-\hat{\beta}_0(1)\right] \\ &> F\left(p\hat{\beta}_0(1)-\varepsilon\right)p\left[y-\hat{\beta}_0(1)\right]=\hat{\Pi}(1,0)\end{aligned}$$

for $\theta > 0$ by the fact that $F(.)$ is increasing where the first inequality holds since $\hat{\beta}(1)$ is the profit maximizing bonus. Hence there exists a range of $c \in [0, \overline{c}]$ where $\overline{c} > 0$ such that:

$$\hat{\Pi}(1,1) - c > \hat{\Pi}(1,0).$$

This completes the proof for Proposition 2.

Proof of Corollary 1 To see how M changes with γ, notice first that M is defined as

$$\Pi^1(M) = \Pi^0(M) \tag{24}$$

where $\Pi^1(\mu_0) = \sum_{t=0}^{\infty} \gamma^t \Pi(\mu_t, 1)$ is the discounted present value of future profits from always choosing $\sigma = 1$ with initial condition μ_0, and $\Pi^0(\mu_0) = \sum_{t=0}^{\infty} \gamma^t \Pi(\mu_t, 0)$ is that from always choosing $\sigma = 0$. Then it follows that higher θ and/or ϵ, as well as lower c will shift up the left hand side, which leads to lower M.

D Proof of Proposition 3

Then given the values of $\hat{\Pi}(\mu,1)$ and $\hat{\Pi}(\mu,0)$ at $\mu = 0$ and $\mu = 1$, as well as the continuity and monotonicity of $\hat{\Pi}(\mu,1) - \hat{\Pi}(\mu,0)$, there exists $\hat{\mu}$ such that $\hat{\Pi}(\mu,1) - c \geq \hat{\Pi}(\mu,0)$ if and only if $\mu \geq \hat{\mu}$.

Differentiate both side of (24) with respect to γ, we have

$$\left[\frac{\partial \Pi^1(M)}{\partial \mu_0} - \frac{\partial \Pi^0(M)}{\partial \mu_0}\right]\frac{\partial M}{\partial \gamma} = \frac{\partial}{\partial \gamma}\left[\Pi^0(M) - \Pi^1(M)\right].$$

Note first it is easy to see that $\frac{\partial \Pi^1(M)}{\partial \mu_0} - \frac{\partial \Pi^0(M)}{\partial \mu_0} > 0$. Hence the sign of $\frac{\partial M}{\partial \gamma}$ simply depends on how $\Pi^0(M) - \Pi^1(M)$ changes with γ. Define T_1 to be the time period when $\mu = 1$ is reached by constantly choosing $\sigma = 1$, and T_0 to be the time period when $\mu = 0$ is reached by constantly choosing $\sigma = 0$. Let $T = \max\{T_1, T_0\}$, then

$$\begin{aligned}\Pi^0(M) - \Pi^1(M) &= \Pi(M,0) - \Pi(M,1) + \gamma\left[\Pi(\mu_1^0,0) - \Pi(\mu_1^1,1)\right] + \gamma^2\left[\Pi(\mu_2^0,0) - \Pi(\mu_2^1,1)\right] \\ &\quad +... + \frac{\gamma^T}{1-\gamma}\left[\Pi(0,0) - \Pi(1,1)\right],\end{aligned}$$

where $\mu_1^0, \mu_2^0, ...\mu_{T-1}^0$ denotes the path of μ as firm continues to choose $\sigma = 0$, and $\mu_1^1, \mu_2^1, ...\mu_{T-1}^1$ is the path of μ as $\sigma = 1$ is chosen for all future periods. Therefore

$$\begin{aligned}\frac{\partial}{\partial\gamma}\left[\Pi^0(M) - \Pi^1(M)\right] &= \left[\Pi(\mu_1^0,0) - \Pi(\mu_1^1,1)\right] + 2\gamma\left[\Pi(\mu_2^0,0) - \Pi(\mu_2^1,1)\right] \\ &\quad +... + \frac{\gamma^T\left[1 + T\frac{1-\gamma}{\gamma}\right]}{(1-\gamma)^2}\left[\Pi(0,0) - \Pi(1,1)\right].\end{aligned}$$

This says that the sign of $\frac{\partial M}{\partial \gamma}$ not only depends on $\Pi(0,0) - \Pi(1,1)$, i.e. a larger gap of $\Pi(0,0) - \Pi(1,1)$ makes $\frac{\partial M}{\partial \gamma} > 0$ more likely to be true, but also depends on the entire paths of $\mu_1^0, \mu_2^0, ...\mu_{T-1}^0$ and $\mu_1^1, \mu_2^1, ...\mu_{T-1}^1$.

It is easy to see that for sufficiently large γ, i.e. $\gamma \to 1$,

$$\frac{\gamma^T\left[1 + T\frac{1-\gamma}{\gamma}\right]}{(1-\gamma)^2} \to \infty.$$

Then the right hand side will be dominated by the last term, i.e. $\frac{\partial M}{\partial \gamma} > 0$ if and only if $\Pi(0,0) > \Pi(1,1)$. This says when the firm is sufficiently patient (i.e. the firm cares enough about future profits), if $\Pi(0,0) > \Pi(1,1)$, then the initial threshold M will be higher if γ is higher, i.e. the firm is willing to tolerate a higher threshold level of μ_0 to choose $\sigma = 0$. On the other hand, if $\Pi(0,0) < \Pi(1,1)$, then the initial threshold M will be lower if γ is higher, i.e. the firm will tolerate a lower μ_1 to choose $\sigma = 1$.

On the other hand, if $\gamma \to 0$, i.e. the forward looking maximization problem reduces to the problem of mission choice that maximizes per period profits, i.e. it is the threshold $\hat{\mu}$ that matters now. But let us still look at the local behavior of M when γ is very low:

$$\lim_{\gamma\to 0}\frac{\partial}{\partial\gamma}\left[\Pi^0(M) - \Pi^1(M)\right] = \Pi(\mu_1^0,0) - \Pi(\mu_1^1,1).$$

This says that how the initial threshold level M changes with γ is simply determined by the behavior of $\Pi(\mu_1^0, 0) - \Pi(\mu_1^1, 1)$, i.e. how the two paths yield different profits at period 1.

To see the limiting argument as $\gamma \to 1$, note from the expression of $\Pi^0(\mu_0) - \Pi^1(\mu_0)$ that the last term, i.e. $\frac{\gamma^T}{1-\gamma}[\Pi(0,0) - \Pi(1,1)]$, will be dominating all the sum of previous finite terms because $\frac{\gamma^T}{1-\gamma} \to \infty$. Hence as $\gamma \to 1$, $\Pi^0(\mu_0) - \Pi^1(\mu_0) > 0$ for any μ_0 (which implies $M = 1$) if and only if $\Pi(0,0) - \Pi(1,1) > 0$.

E Proof of Proposition 5

We can write down welfare at the steady state $\mu = 1$:

$$W_1 = \hat{\Pi}(1,1) - c + \int_0^{p\hat{\beta}_1(1)+\theta} F(\psi) d\psi$$

Similarly, one can write down the welfare at steady state $\mu = 0$:

$$W_0 = \hat{\Pi}(0,0) + \int_0^{p\hat{\beta}_0(0)} F(\psi) d\psi$$

Hence we have

$$W_1 - W_0 = \hat{\Pi}(1,1) - \hat{\Pi}(0,0) - c + \int_{p\hat{\beta}_0(0)}^{p\hat{\beta}_1(1)+\theta} F(\psi) d\psi. \qquad (25)$$

Since $\hat{\Pi}(0,1) = \hat{\Pi}(0,0)$, and the fact that $\hat{\Pi}(\mu, 1)$ is increasing in μ, we have $\hat{\Pi}(1,1) > \hat{\Pi}(0,0)$. We can also show that $p\hat{\beta}_1(1) + \theta > p\hat{\beta}_0(0)$. To see this, note first that when $\theta = 0$, $\hat{\beta}_1(1) = \hat{\beta}_0(0)$. Now we show $p\hat{\beta}_1(1) + \theta$ is increasing in θ: from firm's first order condition:

$$\frac{\partial[p\hat{\beta}_1(1) + \theta]}{\partial\theta} = \frac{pF'(p\hat{\beta}_1(1) + \theta)^2 \frac{\partial\hat{\beta}_1(1)}{\partial\theta}}{F''(p\hat{\beta}_1(1) + \theta)F(p\hat{\beta}_1(1) + \theta) - F'(p\hat{\beta}_1(1) + \theta)^2},$$

which is positive from Proposition 1 and Assumption R that F is log concave. Hence the right hand side is positive. It is also easy to see that it is increasing in θ.