Incentives and Ethics: How Markets and Organizations Shape our Moral Behavior*

Mathias Dewatripont† Jean Tirole‡

December 20, 2019

Abstract: Do markets promote unethical behavior? This paper studies how the replacement logic impacts ethical behavior when suppliers are driven by both profit and ethical concerns. To this purpose, it proposes a unified model encompassing the three possible wedges between client and social demands: internalities, externalities, and shrouded attributes.

When supplier fees are constrained, a good approximation of many medical, apps and franchising environments, unethical behavior is more likely, the higher the fee and the more competitive the market. In contrast, with market-determined fees, supplier concentration has no impact on ethical behavior as less competition also means higher fees. More ethical firms are likely to command a lower (higher) market share under constrained (market-determined) fees.

The replacement logic also affects ethical behavior in organizations. Of particular interest is the design of managerial incentives by owners to align objectives despite differences in ethical concerns. Corporate choices are shown to be more ethical than owners would wish if and only if agents enjoy rents. The paper then concludes with a study of private, public, and industry (self-) regulations.

Keywords: Competition, ethics in markets and organizations, replacement logic, regulation.

JEL numbers: D21, D43, D6, I11, L21.

*This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 669217 - ERC MARKLIM). Jean Tirole acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). Financial support from the Chaire SCOR "Market risks and Value creation" is gratefully acknowledged. We gratefully acknowledge the comments of participants at a Bonn symposium in honor of Martin Hellwig, seminars at Université Libre de Bruxelles (ECARES and I3h), the TSE conference on "Markets, Morality and Social Responsibility", and of Aimé Bierdel, Paul-Henri Moisson and Charles Pébereau.

†Université libre de Bruxelles (Solvay Brussels School of Economics and Management, ECARES and I3h).

‡Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST).

1 Introduction

Do markets run contrary to the common good? While 18th century thinkers viewed markets as creating trust among otherwise unrelated individuals (Montesquieu's "doux commerce"), today's public opinion, many social scientists, politicians and religious leaders feel that markets promote unethical behavior. For instance, numerous prominent philosophers have lately expressed their distrust toward markets, with a variety of viewpoints from the necessity to ban repugnant markets to the stance that a market economy is an unlikely path to a harmonious society.[1]

Furthermore, recent experimental work demonstrates the power of the "replacement excuse", the fact that if a supplier refuses to engage in an immoral trade, "someone else will". This excuse has been used as an individual narrative by Nazis about-to-be convicted of crimes against humanity, by firms and countries selling weapons to dictatorships, by banks selling toxic products or providing short-term incentives to talents they want to attract, or by doctors overprescribing dangerous drugs. Newspapers and websites that violate the confidentiality of investigations and thereby inflict social sanctions on citizens who have not yet been proved guilty invoke the same excuse ("the news would have come out anyway"). So do we when we gossip about someone ("you will learn it from somebody else"). Other examples include the bribing by business executives of officials to win a contract, and professional athletes taking illicit drugs to defeat their competitors.

Indeed, the replacement excuse has impeccable logic to anyone with a consequentialist bent. Important theoretical work by Dufwenberg et al. (2011) and Sobel (2015) identified conditions under which an economy whose agents have other-regarding preferences delivers the same allocation as if these agents were perfectly selfish: Perfectly competitive markets completely destroy any velleity of doing good.

This paper studies how competition impacts ethical behavior when these conditions are not met. The framework, developed in Section 2, posits that suppliers are driven by both a profit motive and an ethical concern. They are consequentialist, in that they reflect on what would happen if they did not serve the client; their moral compass is activated by a disconnect between what is desired by the client and what is good for society. The first possible wedge may be traced to an *internality* (as is the case when a doctor overprescribes opioids, which is attractive to the client's "current self" but detrimental to her "long-term self"). Second, the disconnect may stem from an *externality* (as when doctors deliver fake medical certificates to allow their client not to be vaccinated or to take sick leave). A third wedge is associated with *shrouded attributes*[2] as when a supplier misrepresents the product or exploits the client's incorrect prior or inattention. We provide numerous examples of these three wedges between client and social demands. In all three cases, ethical concerns make suppliers care about the wedge between private and social values. Our model, which

[1] E.g. Anderson (1993), Sandel (2012), Satz (2010), and Walzer (2008).

[2] We borrow the terminology from Gabaix and Laibson (2006)'s work on imperfect competition subject to such attributes.

unifies the three wedges within a single framework, looks at supplier competition along two dimensions: the level of utility promised to customers/patients and, possibly, the level of fees (prices).

Section 3 studies competition among suppliers whose fees are constrained, a good approximation of many medical markets and apps or franchising environments. The fee may be set by a regulatory authority or by a private franchisor, platform or HMO. Alternatively, it may hit the "zero lower bound", as is so often the case in IT services. It first establishes that unethical behavior is more likely, the higher the stakes (the level of the fee), and the more competitive the market, vindicating the view that non-price competition may work against ethical behavior. The latter result's intuition is that ethical behavior implies a substantial loss in market share in unconcentrated markets. The results apply to both covered and imperfectly covered markets (in the latter case, the suppliers compete among themselves as well as with an outside option).

Our model has implications for advertising. It predicts that direct-to-consumer advertising of a product should be prohibited when the product generates internalities or negative externalities. In contrast, advertising is not problematic when the consumption of the product generates positive externalities or is neutral. We compare the predictions with the legal framework in force in several OECD countries.

Section 4 looks at competition among suppliers whose fees are market-determined. In contrast with the case of constrained fees, supplier concentration has no impact on ethical behavior. As earlier, intense competition gives more prominence to the possibility of replacement and thereby lowers the incentive to be virtuous; however, competition also reduces prices and therefore the stake in attracting clients through unethical behavior. The two effects, associated with price and non-price competition, exactly offset each other.

Section 5 then studies the equilibrium under heterogeneous ethical concerns, first with constrained fees and then with market-determined ones. Unsurprisingly, with constrained fees, the less ethical supplier commands a higher market share than the more ethical one. More interestingly, with market-determined fees, the more ethical supplier, while still exhibiting a more ethical behavior, also sets a lower price. Intuitively, the more moral type offers a more ethical, but less attractive service, and so must charge lower prices; furthermore, she is eager to gain market share as she is more virtuous than the other suppliers, inducing her to further lower the price. In the linear-demand duopoly version of the model, a rise in the ethical concern of the more ethical type raises her market share, meaning that her price cut more than offsets the lower pandering to the consumer (in contrast, we saw that the less ethical supplier has a larger market share when the fee is constrained). The same holds for a rise in the ethical concern of the less ethical type, provided the difference between the two types is not too large.

Section 6 analyzes the interaction between suppliers' own ethical concerns and "delegated philanthropy", in which customers and investors value ethical behavior (socially responsible consumption and investment). We provide a simple and explicit formula for the resulting ethics level.

We view our analysis as a cautionary statement regarding markets, not as a frontal attack on them; for, one cannot mount an attack against markets without considering their counterfactual. Section 7 studies the hierarchical interaction between owner's and manager's ethical concerns. It shows that when managers receive no rent, ethical behavior is entirely determined by the owners' social preferences. By contrast, managerial preferences are accounted for in policy when managers enjoy rents; and so ethical concerns in part add up. Section 7 also shows that the replacement logic applies to a wide range of institutions, and compares alternative methods for ensuring compliance when the managers' ethical concerns differ from the owners'. The owner prefers piece rates to exclusive territories, and tournaments/relative performance evaluation to piece rates. Finally, we show that the same logic applies when only "orders" are possible. A parallel can be established with moral hazard models that rely respectively on output-contingent and efficiency-wage models.

Section 8 studies several forms of regulations. It first compares the fees set by a social planner and a private entity (such as an HMO) to the laissez-faire fee, under a free entry condition. It then analyzes Pigovian taxation (when the latter is feasible). Finally, it asks whether self-regulation in the form of a club committing to good ethical practices can supplant formal regulation. Section 9 concludes with some alleys for future research.

Related literature

The paper takes the replacement logic as its starting point. This logic receives much support in experimental work (let alone anecdotal evidence). Falk and Szech (2017) show that (the perception of) pivotality is key to sustaining moral behavior. Organizations that aim at being ethical should accordingly attribute individual responsibility to their members. Bartling and Özdemir (2017) demonstrate that the replacement excuse is less prevalent when there is a strong social norm, though. On the theory front, Dufwenberg et al. (2011) provide conditions under which agents with other-regarding preferences behave like selfish agents in Walrasian equilibrium (in which agents take prices as given). Sobel (2015) studies strategic behavior in auction markets with full replacement or in large economies and comes to the same conclusion: markets make selfish, and the replacement logic implies that an observer cannot distinguish between selfish agents and agents with other-regarding preferences. We extend the Dufwenberg et al and Sobel contributions by assuming supplier differentiation; the replacement logic then operates only partially, which breaks the observational equivalence of economies populated with selfish and other-regarding agents and therefore allows us to study socially responsible behavior.

The paper also has a strong connection with the corporate social responsibility (CSR) literature. To situate this paper in the CSR literature, it may be useful to refer to the taxonomy in Bénabou-Tirole (2010). In that taxonomy, the first notion of CSR amounts to emphasizing sustainability. According to this approach (popular with sovereign wealth funds for instance), CSR is about avoiding the short-termism that has characterized many industries, most notably banking prior to the 2008 crisis (whether due to bonus compensation or career/glory concerns). One may wonder, though, how taking a long-term

perspective to maximizing (intertemporal) profit contributes to CSR. The answer lies in a correlation: protecting stakeholders (workers, environment) against managerial short-termism may also protect shareholders. This first view of CSR fits with much of the corporate governance literature, which emphasizes shareholder activism to promote long-termism, deferred compensation, etc. Related, although less conventional, is the view that shared control may help protect stakeholders.[3]

The second view of CSR equates it with "delegated philanthropy". The firm is a channel for the expression of citizen values; put differently, stakeholders have a demand for corporations to engage in philanthropy on their behalf: consumers will be willing to pay a bit more for their coffee if Starbucks embraces the fair-trade approach. Environmentally conscious investors will accept getting a smaller return from green funds. Workers will take a wage cut when employed by an NGO. As in the first approach, the consumer-product company or the investment institution maximize profit, as they pass through the higher cost or the lower return to the stakeholders. This second view is embraced in Aghion et al. (2019) Besley-Persson (2019), and Moisson (2019). Aghion et al. formalizes CSR behaviour as a quality parameter, shows that under some conditions competition induces greener behaviour and tests this hypothesis. Besley-Persson adds a political determination of the tax/subsidy system and allows values to adjust over time to the (endogenous) technological evolution. Moisson puts particular emphasis on image concerns and consequentialism; in the baseline model, consumers care about impact investment. The paper studies the evolution of the green premium paid by socially responsible investors as environmental consciousness grows.

The third approach is "insider-initiated corporate philanthropy", namely philanthropy that clashes with the profit-maximization hypothesis.[4] This is the approach taken in this paper as well as in Hart-Zingales (2017). In that paper, shareholders also compare their monetary gains with the ethical impact of their actions. This tradeoff has "bite" when they vote at the general assembly or board of directors, since both impacts are non-zero only if their vote is pivotal. By contrast, this leads them to focus solely only on monetary gains when they buy shares (there is no socially responsible investment), since they rationally expect to be pivotal and therefore affect the company's future actions only with a tiny probability.

Our paper is also related to several strands of the industrial organization literature. The result that a more intense competition may deliver poor non-price outcomes may be reminiscent of the work on the value of the banking franchise. In the industrial organization realm, Shapiro (1983) showed that the quality of experience goods may fall when intense competition reduces the payoff to maintaining a reputation for quality. Our frame-

[3]Jäger et al (2019) investigate the causal effects of shared governance building on a 1994 swift reform in Germany and argue that co-determination may act as a substitute for worker protection. They provide evidence that workers on the board, among other things, affect board composition (more women, fewer aristocrats) and -contrary to expectations- increase investment in long-term assets.

[4]Even leaving aside the agency literature, there is of course a long tradition of analyses of non-profit-maximization goals: Beckerian discrimination theory, labor-managed firms, etc.

work however links non-price choices to ethical concerns rather than reputation, and has distinct policy implications. More broadly, the IO literature on quality provision refers to "good quality", namely to quality that is demanded by both consumers and the social planner; the constraint on quality is then its cost. While Spence (1975) has shown that firms with market power may provide too little or too much quality, quality is not both demanded by consumers and frowned upon by the social planner ("bad quality"), unlike in this paper. Finally, and in relation with Section 7, Holmström-Milgrom (1991) and Laffont-Tirole (1991) stress that high-powered incentives by a principal may compromise the agent's provision of non-contractable quality;[5] we refine that insight by emphasizing its relationship with competition, both market competition (we identify the agent's cost of quality as an endogenous loss in market share), and relative performance evaluation.

2 A model of ethical concerns in markets

2.1 Framework

There are n suppliers, $i \in \{1, \ldots, n\}$ and a mass 1 of clients. Suppliers provide a service for clients. The cost of providing this service is γ per client. The outcome of their interaction is characterized by a perceived gross surplus $u \in [0, \bar{u}]$ for the client and an impact w on social welfare. The "Pareto frontier," described in Figure 1, obeys the following properties:

Assumption 1 *(Pareto frontier). The efficient frontier $w = W(u)$ is a smooth function on $[0, \bar{u}]$ and is single peaked, with a peak at $\hat{u} \in (0, \bar{u})$ (such that $\hat{w} \equiv W(\hat{u}) = \max W(u)$). Furthermore, W is concave on $[\hat{u}, \bar{u}]$ and $W'(\bar{u}) = -\infty$.*

Demand functions. Let $\{f_j\}_{j\in\{1,\ldots,n\}}$ and $\{u_j\}_{j\in\{1,\ldots,n\}}$ denote the fees charged by the suppliers and the gross consumer surpluses that they deliver. The demand $\tilde{D}_i$ for supplier i depends on net surpluses $\{u_j - f_j\}_{j=1,\ldots,n}$ and is symmetrical (it is invariant to permutations of net surpluses $\{u_j - f_j\}_{j\neq i}$). Let

$$q_i = \tilde{D}_i(f_1 - u_1, \ldots, f_i - u_i, \ldots, f_n - u_n).$$

Occasionally, we will specialize to a linear demand system (and indicate when we do so). In the relevant range:

$$\tilde{D}_i = \frac{1}{n} - (n-1)\sigma(f_i - u_i) + \sigma\left[\Sigma_{j\neq i}(f_j - u_j)\right]$$

for $\sigma > 0$. Letting $w_{-i} \equiv \frac{\Sigma_{j\neq i} w_j}{n-1}$, $u_{-i} \equiv \frac{\Sigma_{j\neq i} u_j}{n-1}$, $f_{-i} \equiv \frac{\Sigma_{j\neq i} f_j}{n-1}$ and $\hat{\sigma} \equiv (n-1)\sigma$, the demand function can be rewritten as

$$\tilde{D}_i = \frac{1}{n} - \hat{\sigma}[(f_i - u_i) - (f_{-i} - u_{-i})].$$

[5] Where quality here is viewed from the principal's standpoint.

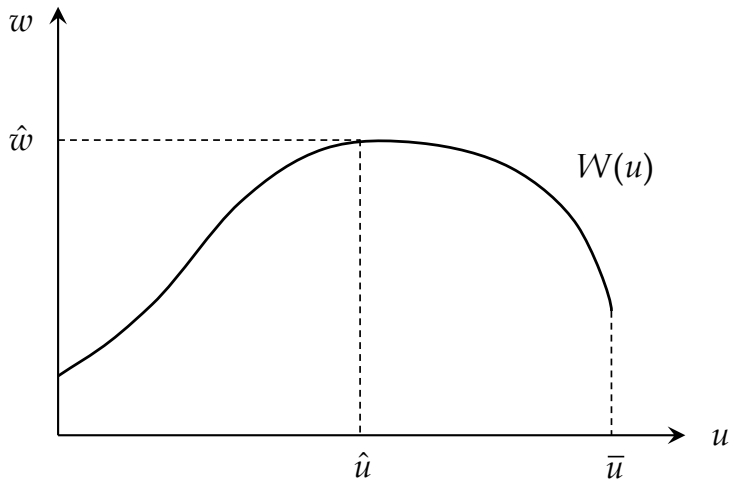

Figure 1: Pareto frontier.

We will focus on symmetric Nash equilibria $\{f^*, u^*\}$ and so we will rewrite demand functions as

$$q_i = D_i(f_i - u_i, f^* - u^*) \equiv \tilde{D}_i(f^* - u^*, \ldots, f_i - u_i, \ldots, f^* - u^*).$$

The suppliers are substitutes ($\partial D_i/\partial f_i < 0 < \partial D_i/\partial f^*$) and their marginal revenue is decreasing ($(f_i - \hat{\gamma})D_i(f_i - u_i, f^* - u^*)$ is concave in f_i for any $\hat{\gamma}$). We will let

$$\eta_i \equiv \frac{-\frac{\partial D_i}{\partial f_i}}{\frac{D_i}{f_i}}$$

denote the elasticity of demand for supplier i's services. Finally, we define supplier i's Lerner index:

$$L_i = \frac{f_i - \gamma}{f_i}.$$

Objective functions. Suppliers care about profit, but also have ethical concerns, reflected in their contribution to overall welfare. When "acquiring" a customer, they account for not only their markup over cost, but also how this makes a difference to social welfare.[6] Let $\alpha \geq 0$ denote the suppliers' relative weight on the ethical impact relative to profit. In the case of a symmetric equilibrium, with gross utilities u^*, associated welfare w^* and fees f^*, the supplier maximizes

$$V_i \equiv [f_i - \gamma + \alpha(w_i - w^*)]D_i(f_i - u_i, f^* - u^*).$$

We will first assume that, in the relevant range of parameters, the market is covered, that is $\Sigma_i D_i = 1$; for example everyone uses a doctor. We will later relax this assumption, which captures most starkly the replacement excuse. We will consider two cases:

[6]See the Appendix for alternative moral imperatives.

(i) *Constrained fees:* $f_i = f$ for all i. The fee f may be set by a regulator. The assumption of regulated fees is reasonable for medical markets in a number of countries. But it may also be relevant in economies in which fees are negotiated between, say, an HMO and an hospital or a group of doctors. Other examples of activities with regulated fees include taxis, apps whose prices are capped by platforms[7], and franchises whose tariffs are set by franchisors. Finally, we use the terminology "constrained fees" rather than "regulated fees", as an important subclass of fixed prices is associated with the zero lower bound on pricing, so ubiquitous in the tech industry (see Section 2.2).

(ii) *Competitive fees:* unconstrained suppliers set their fees non-cooperatively.

2.2 Illustrations: The three wedges

We provide three illustrations of the model. In the first two, being ethical means "being less nice" to the client (or at least her current self); in the third one, being ethical means "being nicer" to the client. In each case, there is a wedge between the quality perceived by the customer and that assessed by the social planner.

Example 1: Internalities (painkiller prescriptions). The supplier (a doctor) decides whether to prescribe an opioid to the client (the patient). The painkiller brings known benefit b, but has side effects with cost c. This cost is observed only by the doctor (who learns who is at risk) and is distributed according to distribution $G(c)$ and density $g(c)$. The doctor chooses a threshold c^* under which she prescribes the painkiller. Assume that the patient knows c^*; one may have in mind that patients know the doctor's reputation for being easy (pill mill doctor) or tough on prescriptions. Welfare is $b - c$, but clients have hyperbolic preferences with coefficient $\beta < 1$: They long for quick relief and value the prescription at $b - \beta c$. And so

$$u_i = \int_0^{c_i^*} (b - \beta c) g(c) dc \quad \text{while} \quad w_i = \int_0^{c_i^*} (b - c) g(c) dc.$$

The maximum gross surplus corresponds to $c_i^* = b/\beta$ and is equal to $\bar{u} \equiv \int_0^{b/\beta} (b - \beta c) g(c) dc$. The welfare optimum corresponds to $c_i^* = b$ and so

$$\hat{w} \equiv \int_0^{b} (b - c) g(c) dc \quad \text{and} \quad \hat{u} \equiv \int_0^{b} (b - \beta c) g(c) dc.$$

More generally, the function $w_i = W(u_i)$ is obtained by substituting c^*, and satisfies Assumption 1.[8]

[7] Examples include no-surcharge rules, and best-price guarantees.

[8] One has

$$W''(u) = \frac{d}{dc^*} \left(\frac{b - c^*}{b - \beta c^*} \right) \Big/ \frac{du}{dc^*} < 0.$$

Instances of overconsumption due to imperfect self-control or biases in predicting one's future behavior are many outside the health domain (excessive indebtedness, gambling, videogaming, impulsive clicking on privacy consent forms...).

Example 2: Externalities (vaccines, overprescription of branded drugs). This example replaces the internality of Example 1 by an externality. Patients have heterogenous probabilities x of getting sick in the absence of vaccination, in which case they suffer damage d and contaminate an expected number e of other people. Patients are selfish and value being vaccinated at $E[b-c]$, where $b=xd$ is the benefit and c is a cost of vaccination. The social planner attaches value $E[(1+e)b-c]$. It is easy to check that this externality example is mathematically akin to the internality example, Example 1.[9]

Underconsumption of vaccines may also be driven by a misperception of their side-effects. For example, the measles vaccine was falsely accused in *The Lancet* of causing autism, which led to a substantial drop in MMR vaccination. Such misperceptions way be captured as an underestimation of the true value of the vaccine, independently of contagion considerations.

Underconsumption, unlike overconsumption, raises the question of how the supplier can provide a quantity that exceeds the client's desired consumption: a doctor cannot physically vaccinate a patient who refuses to be inoculated. A first interpretation of the underconsumption model goes as follows: when the state mandates children to be vaccinated in order to be able to go to school or public sport facilities, parents may try to obtain a complacent (fake) vaccination certificate from the doctor (underconsumption of vaccines). Similarly, in some countries, occupational physicians may routinely deliver fake medical certificates allowing employees to take paid sick leave (underprovision of work). In both examples, unethical supplier behavior is associated with a fraudulent report to a third party. A second interpretation applies when no law or rule mandates a level of consumption in excess to that desired by the client. Ethical/unethical behavior then relates to the intensity with which the doctor puts pressure on the patient, say to be vaccinated; it may range from attempts at persuasion to outright refusal to keep seeing a patient who refuses the vaccination.

Overconsumption occurs in the case of antibiotics. Another case in point is the overconsumption of branded drugs when generics are available, imposing an externality on the social security system. A fraction of French patients has always viewed generics as inferior products. Until the mid-90s French doctors faced no cost of prescribing branded drugs instead of generics (and pharmacists' compensation was proportional to the price of the drug!). Lo and behold, doctors pandered to their patients and generics' market share was about 2%. A reform introduced incentives for doctors to prescribe generics, and also gave pharmacists the ability to replace a branded drug by an equivalent generic. The share of generics' prescriptions improved, especially with general practitioners (whose patients are more loyal than for specialists, in conformity with the theory developed below). But the

[9] One has $u \equiv \int_{b^*}^{\infty}(b-c)dF(b)$ and $W \equiv \int_{b^*}^{\infty}[(1+e)b-c]dF(b)$. A supplier with ethical concerns will choose b^* in $\left[\frac{c}{1+e}, c\right]$. Then $W'(u^*) = -\left[\frac{(1+e)b^*-c}{b^*-c}\right] < 0$ and $W''(u) = \frac{ec}{(b^*-c)^2}\frac{db^*}{du} < 0$.

low percentage (36%)[10] of generics among prescriptions reimbursed by the social security system suggests that there is still substantial pandering.

Example 3: Shrouded attributes. Firms typically emphasize positive attributes of their goods and services and rarely their flaws. To be certain, consumer protection agencies' and courts' mission is to combat inappropriate statements or frauds. But there is a thin line between outright misrepresentation and fraud on the one hand, and omission, opaque language and the exploitation of consumer inattention on the other hand. Unexpected obsolescence, vague recommended usage or the downplaying of side effects may not be illegal or else hard to regulate given their ubiquity and the limited means of the agencies. Socially-responsible investors may not be aware that some green-looking securities they purchase do not correspond to an actual impact on climate change. Like internalities and externalities, shrouded attributes leave scope for moral judgment.

One way to formalize this within our model goes as follows: Suppose that the good actually delivers gross surplus $\hat{u}$ to the consumers. Supplier i can inflate this surplus and claim it is $u_i \in [\hat{u}, \bar{u}]$. Consumers take the announcement at face value (see below for a more rational version) and plan around the announced value, leading to later inconvenience (complementary investments miscalibration, misleading claims made to downstream users...) cost $\Gamma(u_i - \hat{u})$, with $\Gamma(0) = 0$, $\Gamma'(0) = 0$, $\Gamma'(u_i - \hat{u}) > 0$ and $\Gamma''(u_i - \hat{u}) > 0$ for $u_i > \hat{u}$, and $\Gamma'(\bar{u} - \hat{u}) = +\infty$ for some maximum exaggeration level $\bar{u} - \hat{u}$. Then $W(u) \equiv \hat{u} - \Gamma(u - \hat{u})$ satisfies the general assumptions. Supplier i's objective function is $[f_i - \gamma + \alpha[W(u_i) - W(u^*)]]D_i(f_i - u_i, f^* - u^*)$.

Finally, note that a more rational, asymmetric-information, version of the model would have consumers not know about the misreporting function. For example, with some probability they believe that misreporting is infeasible; the "irrational version" is just the limit of the "rational version" as this probability goes to 1.

The case of apps. The tech industry illustrates both the possibility of constrained fees and the three factors of wedge between consumer demand and social welfare. First, note that many applications (such as search, videos, GPS navigation software, social networks...) are available free of charge to the consumer. As emphasized in the industrial organization literature, such services really have negative opportunity costs (the profit from targeted advertising and data collection far outweighs their tiny physical cost), so that their optimal price is constrained by the zero lower bound linked to the threat of arbitrage by bots.

Apps also exhibit the "wedge trilogy". First, we all click impatiently, perhaps overemphasizing our immediate gratification over the long-term loss of privacy (internality). Second, our e-mails and postings on social networks reveal information about others, violating their privacy (externality).[11] Third, the false sense of privacy experienced by most consumers is an illustration of shrouded attributes; for instance most users of the "Do not track" privacy setting do not know that big tech companies do not abide by the spirit of

[10] Generics penetration in the US, UK and Germany exceeds 80%.

[11] See, e.g. Choi et al (2019).

the request[12]. More broadly, tech companies work on reassuring consumers more than on really protecting their privacy.

3 Equilibrium ethics under constrained fees

3.1 Equilibrium characterization

Suppose that the fee is constrained at some level f. We make the following assumptions:

Assumption 2 *:* $\frac{d(\eta L)}{df} > 0$.

Assumption 2 will ensure the monotonicity of equilibrium behavior with respect to f. It is reasonable. Note first that $\frac{dL}{df} > 0$. So for Assumption 2 to be satisfied, it suffices that the elasticity of demand does not decrease too fast with f. For example, the elasticity of demand is proportional to f (and so is increasing in f) when consumers' demand obeys a unit-demand, discrete choice model.[13] Another illustration relates to the free-entry outcome: intuitively, a higher fee leads to more entry, which exacerbates competition.[14]

Assumption 3 *: Let* $\eta(u^*) = -\frac{\partial D_i}{\partial f_i}(f-u^*, f-u^*)nf$ *denote the elasticity of demand in a symmetric equilibrium (*$u_i = u^*$ *for all* i*). Then* $\frac{\eta'(u^*)}{\eta(u^*)} \leq \frac{W''(u^*)}{W'(u^*)}$.

Assumption 3 will ensure uniqueness of a symmetric equilibrium. It is satisfied for example for CES and Hotelling-Lerner-Salop demand functions, as well as in the discrete choice model (for which $\eta'(u^*) = 0$). In the absence of Assumption 3, there may be strategic complementarities and multiple symmetric equilibria (see the Appendix).

We look for a symmetric equilibrium. Supplier i then solves:

$$\max_{\{u_i\}} [f - \gamma + \alpha[W(u_i) - w^*]]D_i(f - u_i, f - u^*),$$

yielding first-order condition in a symmetric equilibrium:

$$-\alpha W'(u^*) = \eta(u^*)L \tag{1}$$

where, recall, $L \equiv \frac{f-\gamma}{f}$ is the Lerner index, and $\eta(u^*)$ denotes the elasticity of individual demands when net prices are all equal to $f - u^*$.

[12]Web sites are legally entitled to decide what they think is right.

[13]A consumer h selects to buy from supplier i if $u_i + \varepsilon_{ih} - f > \max_{j \neq i}\{u_j + \varepsilon_{jh} - f\}$, where $\varepsilon_{\cdot h}$ are idiosyncratic taste shocks obeying a symmetric distribution. Then $D_i = \Pr(u_i - u^* \geq \max_{j \neq i}\{\varepsilon_{jh} - \varepsilon_{ih}\})$ in a symmetric equilibrium is independent of f and $\frac{-\frac{\partial D_i}{\partial f_i}}{\frac{D_i}{f}}$ is proportional to f.

[14]This is indeed the case for the Lerner-Salop model, for which both $\eta = \frac{f(f-\gamma)}{t\bar{V}}$ (where $\bar{V}$ is the reservation utility or fixed entry cost) and $L = \frac{f-\gamma}{f}$ are increasing in f.

We will assume that the second-order condition is satisfied; a sufficient condition for this is that the demand be linear or concave.[15]

Proposition 1 *(impact of competition on ethics under constrained fees and a covered market). There exists a symmetric equilibrium under constrained fees. The equilibrium level of ethics is unique under Assumption 3,*[16] *and is given by*

$$-\alpha W'(u^*) = \eta(u^*)L.$$

(i) It is decreasing (u^* is increasing) in the intensity of competition (η), from the socially desirable level ($u^* = \hat{u}$) when the clients are captive ($\eta = 0$) to the ethic-free outcome ($u^* = \bar{u}$) under perfect competition ($\eta = \infty$).

(ii) It is decreasing (u^* is increasing) in the constrained fee (f) under Assumption 2, from the socially desirable level for $f = \gamma$ to the ethics-free outcome as f tends to ∞.[17]

(iii) It is increasing (u^* is decreasing) with $\alpha \in [0, +\infty)$ from the ethic-free outcome to the socially desirable level.

Income effects. Condition (1) can be re-interpreted to shed some light on Shleifer (2004)'s insight that ethical behavior is a normal good, and therefore likely to improve as societies become richer. Suppose that a supplier's value of money is an increasing and concave function Φ. Then $V_i = \Phi((f-\gamma)D_i) + \alpha[W(u_i) - w^*]D_i$. The analysis is then the same, except that α is replaced in condition 1 by

$$\hat{\alpha} \equiv \frac{\alpha}{\Phi'\left(\frac{f-\gamma}{n}\right)}.$$

The denominator captures Shleifer's impoverishment effect: an exogenous downward shock on the marginal utility of money Φ' increases the suppliers' relative weight on money and ceteris paribus makes them behave less ethically. When the increased concern for money results from a decrease in the fee f however, a second effect operates in the opposite direction: the suppliers have a lower stake in behaving unethically (Proposition 1 (ii)). The net effect is a priori ambiguous. So it is not necessarily the case that, ceteris

[15] The second-order condition writes, using $W(u_i) = w^*$ at $u_i = u^*$:

$$\alpha\left[W''D_i - 2W'\frac{\partial D_i}{\partial f_i}\right] + [f-\gamma]\frac{\partial^2 D_i}{\partial f_i^2} < 0.$$

[16] Assumption 1 implies that the continuous function $-\alpha W'(u^*) - \eta(u^*)L$ is negative at $\hat{u}$ and positive at $\bar{u}$. Its derivative when it takes value 0 is $\alpha W'(u^*)\left[\frac{\eta'(u^*)}{\eta(u^*)} - \frac{W''(u^*)}{W'(u^*)}\right]$. So if $\eta'/\eta < W''/W'$, the function equals 0 at exactly one value of u^*.

[17] Of course, for very large f, the assumption that the market is covered becomes much less plausible. More on this shortly.

paribus, rich doctors will behave better than their poorer counterparts in other parts of the world.[18]

Patient imperfect information

Serious controversy arose regarding the aggressive advertising of OxyContin by Purdue. More generally, direct advertising to consumers is viewed as putting pressure on doctors to prescribe drugs. Suppose that a fraction of patients are not aware of the availability of the drug (and consequently do not know the doctor's reputation in the matter). Those patients go to the doctor nonetheless to be treated; under regulated fees, each doctor receives market share $1/n$ in this subgroup of patients. Other patients are as described above. If the doctors can discriminate between informed and uninformed, then $u^* = \hat{u}$ for the uninformed and $-\alpha W'(u^*) = \eta(u^*)L$ for the informed. If z is the fraction of informed clients, the lack of consumer information generates a social gain equal to $(1-z)[W(\hat{u}) - W(u^*)]$. In this model direct-to-consumer advertising should be banned as it augments the pressure on suppliers to please the client.[19]

This corollary of our previous analysis would be worth of empirical investigation. Casual empiricism, based on a few legislations (Table 1), indeed suggests that consumer advertising is much more tightly regulated for prescription drugs, for which overconsumption can be very costly, with some exceptions in the case of underconsumption (vaccines, tobacco addiction reduction). By contrast, regulations on advertising are much more lenient for non-prescription drugs.

	Canada	Australia	China	France	Germany	Sweden	USA
Advertising of non-prescription medicine	Yes, subject to FDA regulation on content	Yes, subject to prior approval	Yes	Yes if no risk for public health	Yes	Yes for adults	Yes
Advertising of prescription-only medicines	Yes but limited to name price and quantity	No	No	No except for vaccines and tobacco addiction reduction	No	No except for vaccines	Yes
Disease awareness campaign	Yes, can discuss treatment options	Yes, can mention treatment options	Yes but no reference to medicine	Yes if no reference to a medicine	Yes if no reference to a medicine	Yes but not focusing on treatments	Yes

Table 1: drug advertising

[Built using data from https://iclg.com/compare/pharmaceutical-advertising]

Do good doctors behave more ethically than bad ones?

Suppose now that suppliers differ in their talent. Let s_i denote supplier i's exogenous

[18]There is another caveat. A society's average other-regarding preference (α) may evolve to reflect changes in the benefit of prosocial behavior. Thus, while the effects analyzed in the text seem robust, they may not be the only relevant effects.

[19]Of course, the model abstracts from potential benefits of direct-to-consumer advertising, such as informing the patient that a treatment exists, inducing her to see a doctor (this effect is most likely for recently approved drugs).

talent. Assuming a duopoly and linear demand, supplier i's demand is $D_i(f-u_i-s_i, f-u_j-s_j) = \frac{1}{2} - \sigma[(f_i-u_i-s_i)-(f_j-u_j-s_j)]$). The first-order condition writes:

$$\sigma[f-\gamma+\alpha(w_i-w_j)] = \alpha[-W'(u_i)]D_i.$$

Suppose that the good doctor (call her i) is no more ethical than the bad one (call her j): $u_i \geq u_j$. Then $w_i \leq w_j$, $D_i > D_j$ and $-W'(u_i) \geq -W'(u_j)$, a contradiction. So $u_i < u_j$: *good doctors behave more ethically*, and (from the FOC and the properties that $W' < 0$ and $W'' < 0$) *command a larger market share* ($D_i > D_j$).

3.2 Non-covered market

In Section 3.1, clients had no choice but picking a supplier; in other words, outside options were never attractive. In contrast, let us introduce an outside option yielding utility u_0 and welfare w_0. This outside option might stand for not consulting a doctor, finding illegal drugs to replace OxyContin, etc. When performing comparative statics below, we will posit that w_0 covaries negatively with u_0 (although not necessarily according to W as the outside option might have a different nature).

We consider the same oligopoly model, but in which outside options are "co-located" with the products. This model offers the convenience of a smooth transition from perfect competition to pure monopoly.[20] In this nested discrete choice model, (1) the client chooses among the suppliers (so $\Sigma_{i=1,...,n} D_i = 1$); (2) the client chooses between her preferred within option (i, say) and an outside option co-located with i. The outside option offers utility u_0 and welfare w_0; it involves an idiosyncratic cost $\kappa \in \mathbb{R}$ of using the outside option, distributed according to cumulative distribution $X(\kappa)$. Without loss of generality, we assume that the price of the outside option is equal to that of the regulated-fee products ($f = f_0$); if this is not the case, one can renormalize the distribution of κ to obtain the expressions below. We assume that $X'/(1-X)$ is increasing (monotone hazard rate property). The demand for good i is then $[1-X(u_0-u_i)]D_i(f-u_i, f-u^*)$.

Supplier i then solves (assuming a symmetric equilibrium, so we can rewrite the demand function as $D_i(f-u_i, f-u^*)$)

$$\begin{aligned}\max_{\{u_i\}} V_i = {} & X(u_0-u_i)D_i(f-u_i, f-u^*)[\alpha w_0] + [1-X(u_0-u_i)]D_i(f-u_i, f-u^*)[f-\gamma+\alpha W(u_i)] \\ & + X(u_0-u^*)[1-D_i(f-u_i, f-u^*)][\alpha w_0] + [1-X(u_0-u^*)][1-D_i(f-u_i, f-u^*)][\alpha w^*].\end{aligned}$$

Let $\eta \equiv -\frac{\frac{\partial D_i}{\partial f_i}}{\frac{D_i}{f_i}}$ denote the "within elasticity" and $\varepsilon \equiv \frac{-X_i'}{\frac{1-X_i}{f_i}}$ denote the "across elasticity". The total elasticity is $\hat{\eta} \equiv \eta + \varepsilon$. Similarly the Lerner indices in equilibrium are equal to:

$$L \equiv \frac{f-\gamma}{f} \qquad \text{and} \qquad K \equiv \frac{f-\gamma+\alpha(w^*-w_0)}{f}.$$

[20] Bénabou-Tirole (2016). This formulation, unlike Hotelling-style specifications, allows a distinction between the within-industry elasticity and the elasticity with respect to the outside good, as seen below.

The following proposition shows that unethical and attractive outside options exacerbate the need to be attractive to capture demand and reduce suppliers' ethics:

Proposition 2 *(Outside options). In the presence of co-located outside options, the ethical choices are given in a symmetric equilibrium by*

$$-\alpha W'(u^*) = \eta L + \varepsilon K = \hat{\eta} L + \frac{\alpha(w^* - w_0)}{f}\varepsilon.$$

Assume linear demands and exponential cumulative distribution $X(\kappa)$. Then, suppliers are less ethical, the less ethical the outside option (the lower w_0 is).[21]

Application: There is detailed empirical evidence that, when the FDA forced Purdue pharma to make OxyContin "abuse-resistant",[22] patients massively shifted towards heroin. As we suggest, a wider availability of illegal drugs (e.g. when fentanyl made illegally in China becomes increasingly available in the US and Europe) has an additional effect: Doctors become less ethical (not in preferences, but in behavior as they increase u^*) because they feel that keeping patients addicted to prescription opioids limits a (worse) addiction to illegal opioids.[23]

4 Equilibrium ethics under market-determined fees

Suppose now that suppliers choose their fees as well as the gross surplus they offer to their clients.

First-order conditions. Price setting adds the following first-order condition:

$$\eta L = 1. \tag{2}$$

Combining (2) with (1), one obtains:

$$-\alpha W'(u^*) = 1. \tag{3}$$

Condition (3) states that the marginal cost of being unethical ($-\alpha W'(u^*)$) is equal to the marginal benefit: Indeed, an increase in u_i can be offset by an equal increase in f_i, leaving demand D_i constant. The Appendix shows that the global second-order condition is satisfied.

[21] In a symmetric equilibrium, the linear demand for supplier i is: $D_i = 1/n - (n-1)\sigma[(f-u_i)-(f-u*)]$, while the exponential cumulative distribution is: $X(\kappa) = 1 - e^{-\lambda\kappa}$. Given these assumptions, one can show that the first-order condition is: $-\alpha W'(u^*) = [n(n-1)\sigma + \lambda](f-\gamma) + \alpha\lambda(w^* - w_0)$, which implies that $0 < dw^*/dw_0 < 1$.

[22] The abuse-deterrent version of OxyContin, approved by the FDA in 2010, makes the pill difficult to crush or dissolve, making it less dangerous. Alpert et al (2018) show that the OxyContin reformulation significantly reduced OxyContin misuse, but also led to a large increase in heroin deaths.

[23] It would be worth thinking more about the general implications for the legal treatment of various drugs. Another case in point is cannabis, which is viewed both as a drug and a painkiller which could help reduce the opioid epidemic. In our model, if doctors could prescribe it, it could prevent OxyContin addiction for some patients, while now it is "just an outside option", which is moreover less attractive to dealers than heroin or fentanyl.

Proposition 3 *(impact of competition on ethics under market-determined fees). The level of ethics under market-determined fees is given by*

$$-\alpha W'(u^*) = 1.$$

The intuition goes as follows: An increase in competition implies that keeping customers becomes more of a concern, which induces suppliers to behave less ethically. However, competition also lowers fees, reducing the gain from behaving unethically. With demand depending on net fees (fees minus gross surplus), these two opposite forces cancel out.[24]

Remark (good and bad doctors). We saw that under constrained fees, duopoly and linear demand, talented doctors behave more ethically and have a larger market share. Do similar results obtain under market-determined fees? Simple computations show that both select the same ethical behavior (as given by (6): $-\alpha W'(u_i) = -\alpha W'(u_j) = 1$), that the more talented doctor selects a higher fee ($f_i - s_i = f_j - s_j$) and they have the same market share ($D_i = D_j = 1/2$). Again one obtains different predictions under constrained and market-determined fees.

5 Heterogeneous ethical concerns: does the market drive out ethical types?

Suppose now that suppliers differ in their ethical concerns; without loss of generality, assume that $0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$.[25] Supplier i solves:

$$\max \{(f_i - \gamma)D_i + \alpha_i [w_i D_i + \Sigma_{j \neq i} w_j D_j]\}$$

We specialize the analysis by assuming a linear demand system in this section.

(a) Constrained fees

In the constrained-fee context, $f_i \equiv f$ for all i, and the only decision variable is u_i. The maximization of supplier i's objective function with respect to u_i yields first-order condition

$$\hat{\sigma}[(f_i - \gamma) + \alpha_i(w_i - w_{-i})] + \alpha_i W'(u_i) D_i = 0. \tag{4}$$

The intuition behind condition (4) goes as follows: A unit increase in u_i attracts $\hat{\sigma}$ new clients, bringing markup $(f_i - \gamma)$ on each of them. This increase in market share

[24] Another way to think about (3) is through its comparison with Spence (1975). Supplier i can be viewed as offering a product of quality u_i (note that marginal and inframarginal consumers have the same valuation for quality) at cost $\gamma - \alpha[W(u_i) - w^*]$. When fees are endogenous, the supplier equates the marginal cost of quality, $-\alpha W'(u_i)$, and the marginal benefit, 1.

[25] We assume that these social preferences are common knowledge. Otherwise, assuming that suppliers are reputation-conscious, the objective function below has to be augmented with an image term as in, e.g., Bénabou-Tirole (2006).

further improves welfare by $w_i - w_{-i}$ (decreases it if $w_i < w_{-i}$). Finally, the decrease in the welfare corresponding to the D_i clients of supplier i has an ethical cost for supplier i.

The following proposition is proved in the Appendix.

Proposition 4 *(heterogenous ethical concerns under constrained fees) Suppose a linear demand system. Under duopoly, the less ethical firm commands a higher market share.*

(b) Market-determined fees

The first-order condition with respect to f_i is

$$-\hat{\sigma}[(f_i - \gamma) + \alpha_i(w_i - w_{-i})] + D_i = 0. \tag{5}$$

The intuition for (5) is similar to that underlying condition (4). Furthermore, combining (4) and (5) yields

$$-\alpha_i W'(u_i) = 1. \tag{6}$$

Thus the characterization of ethical choices obtained in the symmetric environment extends to heterogenous ethical concerns. While (5) relies on a linear demand system, (6) actually holds for general demand.

We next study whether firms with higher ethical concerns have a low market share. Condition (6) implies that their ethical choices will make them unattractive to clients. But this is not the end of the story. Their ethical concerns also make them eager to capture market shares away from less scrupulous suppliers, who conversely do not want to gain market share for that specific reason. We study the horserace between these two forces in the duopoly case.

Duopoly case. Suppose $n = 2$. Substracting conditions (5) for firms i, j $(i \neq j)$ and using the expressions $D_i = \frac{1}{2} - \hat{\sigma}[(f_i - u_i) - (f_j - u_j)]$, one obtains

$$3[(f_j - u_j) - (f_i - u_i)] - [(u_i - u_j) + (\alpha_i + \alpha_j)(w_i - w_j)] = 0.$$

Suppose that $\alpha_i > \alpha_j$ and so from (6) $u_i < u_j \Leftrightarrow w_i > w_j$.

The strict concavity of W, together with (6) implies that $u_i - u_j + \alpha_i(w_i - w_j) > 0$. And so

$$f_j - u_j > f_i - u_i \Leftrightarrow D_i > D_j.$$

The more ethical rival has a larger market share. Furthermore, one can show that this market share increases when α_i increases, while the impact of α_j is ambiguous (it decreases with α_j for α_j close to α_i, though).[26]

Proposition 5 *(heterogenous ethical concerns under market-determined fees)*

[26] $\frac{\partial D_i}{\partial \alpha_i} \propto \frac{\partial}{\partial \alpha_i}[u_i - u_j + (\alpha_i + \alpha_j)(w_i - w_j)]$, which, using (6), yields $\frac{\partial D_i}{\partial \alpha_i} \propto (w_i - w_j) + \alpha_j W'(u_i)\frac{du_i}{d\alpha_i} > 0$.

(i) For any n, ethical choices are governed by

$$-\alpha_i W'(u_i) = 1.$$

And so more ethical suppliers make more ethical choices.

(ii) Suppose a linear demand system. Under duopoly, the more ethical firm commands a larger market share despite its more ethical choice. This market share increases with its own level of ethics; by contrast, the less ethical rival's ethics has an ambiguous impact on market shares.

6 Combining investor/consumer social responsibility with supplier ethical concerns

We have so far focused on supplier ethical concerns. Let us add similar concerns for the clients. For instance, ethical behavior is incentivized both through socially responsible investment by individual investors and by consumers' engaging in fair trade, green or more generally ethical consumption. So, let us allow the client to care about welfare, a concern much more relevant in the case of externalities than in the other two illustrations of the model.[27] Thus, the net fee, as perceived by the client is $f_i - [u_i - \mathfrak{a}[W(\hat{u}) - W(u_i)]] \equiv f_i - \mathring{u}_i$. The demand function is then $D_i(f_i - \mathring{u}_i, f^* - \mathring{u}^*)$. Under unregulated fees, say, supplier i solves

$$\max_{\{f_i, u_i\}} (f_i - \gamma) D_i(f_i - \mathring{u}_i, f^* - \mathring{u}^*) + \alpha[W(u_i) D_i(f_i - \mathring{u}_i, f^* - \mathring{u}^*) + W(u^*)[1 - D_i(f_i - \mathring{u}_i, f^* - \mathring{u}^*)]].$$

Simple derivations yield:

Proposition 6 *(customer ethical concerns) When both the suppliers and the clients partially internalize their impact on welfare (in proportions α and $\mathfrak{a}$ respectively), ethical choices are given by*

$$\begin{aligned} &-(\alpha + \mathfrak{a}\eta L) W'(u^*) = \eta L \quad \textit{under constrained fees, and} \\ &-(\alpha + \mathfrak{a}) W'(u^*) = 1 \quad \textit{under market-determined fees.} \end{aligned}$$

This proposition shows how the "delegated philanthropy" view of CSR interacts with the "corporate philanthropy" view of CSR. Naturally, the two forms of internalization combine to generate a more ethical behavior.

[27] In case of an internality, the current incarnation cares about a distorted view of her own welfare. Under shrouded attributes, the client is imperfectly informed or unaware of a misrepresentation by the supplier.

7 Is this about markets? Ethics in hierarchies

An organization's ethical behavior is often set by subordinates facing implicit or explicit expectations from their superiors. The opioid misuse was promoted by middlemen and doctors incentivized by Purdue Pharma. War crimes are almost never perpetuated by the top brass; similarly many Nazis in Nuremberg and in other court trials emphasized that "they were only following orders". And of course there are plenty of everyday instances of less dramatic failures to pursue social welfare.

To be clear, we are not claiming that every organization wants to achieve "evil" ($\bar{u}$ in our model). For instance, in a market economy, such a policy may in the long run backfire because of lawsuits or mere consumer disaffection, and so may not be privileged by long-term-oriented management or shareholders; furthermore, the latter may themselves have ethical concerns. In this section, we will assume that the organization's objective diverges from the maximization of social welfare and will look at two mechanisms that define incentives for agents with other-regarding concerns: formal incentives and orders.

7.1 Ethical concerns in hierarchies

Let us investigate the interaction of ethical concerns along the supply chain. There are n owners, each with a manager. The owner's and the manager's ethical concerns are indexed by $\mathfrak{a}$ and α, respectively. We distinguish two cases, depending on whether rents have to be left to the manager.

(i) No managerial rent. Suppose that the "management" has sufficient career concerns or else stands for a well-capitalized subcontractor. The owner of firm i can then require any level u_i he wants by subjecting the manager to a sufficient capital loss if the demand D_i ends up differing from its set target, say $D_i(f_i - u_i, f^* - u^*)$, where $\{f^*, u^*\}$ is the equilibrium choice made by other firms. Under a constrained fee, the owner solves

$$\max_{\{u_i\}} \{(f-\gamma)D_i(f-u_i, f-u^*) + \mathfrak{a}[[W(u_i)-w^*]D_i(f-u_i, f-u^*)+w^*]\}$$

yielding as usual

$$-\mathfrak{a}W'(u^*) = \eta L.$$

Similarly, under market-determined fees, the outcome is given by $-\mathfrak{a}W'(u^*) = 1$.

Either way, the manager's ethical concerns are irrelevant.

(ii) Managerial rents. Suppose next that the managers have no cash and are protected by limited liability.[28] Owner i's objective function is optimized by setting some policy u_i and some compensation $T_i \geq 0$ for the manager, together with the threat of withdrawing

[28] The case in which the managers have some cash is intermediate between the two polar cases we focus on.

T_i if D_i differs from $D(f_i - u_i, f^* - u^*)$. Under constrained fees, owner i solves

$$\max_{\{u_i, T_i\}} \{(f-\gamma)D_i(f-u_i, f-u^*) + \mathfrak{a}[[W(u_i)-w^*]D_i(f-u_i, f-u^*)+w^*] - T_i\}$$

subject to the incentive constraint:

$$T_i + \alpha[[W(u_i)-w^*]D_i(f-u_i, f-u^*)+w^*] \geq \max_{\{\tilde{u}_i\}} \alpha[[W(\tilde{u}_i)-w^*]D_i(f-\tilde{u}_i, f-u^*)+w^*] \equiv \max_{\{\tilde{u}_i\}} \Delta(\tilde{u}_i, u^*) \qquad \text{(IC)}$$

when $\Delta(\tilde{u}_i, u^*)$ is the deviation payoff. This deviation payoff satisfies

$$\frac{\partial \Delta}{\partial \tilde{u}_i} < 0 \quad \text{for} \quad \tilde{u}_i \geq u^* \quad \text{(ethical deviations must be downward deviations)}$$

and

$$\frac{\partial \Delta}{\partial \tilde{u}_i}(\hat{u}, u^*) > 0 \quad \text{(at } \tilde{u}_i = \hat{u}\text{, a marginal improvement in ethics implies a second-order effect on welfare, but a first-order loss in demand).}$$

Constraint (IC) is binding (otherwise $T_i = 0$ and (IC) is not satisfied), defining a function $T(u_i, u^*)$, with

$$\frac{\partial T}{\partial u_i} = -\alpha \frac{\partial}{\partial u_i}[[W(u_i)-w^*]D_i(f-u_i, f-u^*)] > 0.$$

Managerial rents lead to a reduction in u^*, the more so the higher the managers' ethical concerns (the higher α is).

The analysis is similar for *market-determined fees*. One feature of interest, though, is that $\eta L > 1$: The owners raise their price beyond that given by the inverse-elasticity rule so as to make it less tempting for their managers to behave ethically.[29]

Proposition 7 *When managers receive no rent, ethical behavior is entirely determined by the owners' preferences. By contrast, managerial preferences are accounted for in policy when managers enjoy rents; and so ethical concerns in part add up along the hierarchy.*

In either case, the owners' ethical concerns have an important role on the outcome. For example, the analysis suggests that NGOs will, ceteris paribus, offer low-powered incentives to their staff.

7.2 Explicit and implicit incentives that overcome agents' ethical objections

Proposition 7 shows that agents' ethical objections have "bite" if and only if they receive rents. The Appendix refines this analysis by looking at specific incentive schemes: exclusive territories, piece rates, tournaments or simple "orders". The common thread of these

[29] The left-hand side of (IC) is (locally) invariant to f_i, while the derivative of the right-hand side is equal to $\alpha[W(\tilde{u}_i)-w^*]\frac{\partial D_i}{\partial f_i} < 0$, where $\tilde{u}_i$ is the arg max of $\Delta(\tilde{u}_i, u^*)$.

specifications is that, when overcoming agents' ethical objections involves a monetary cost for the owners, these objections lead to more ethics in equilibrium than what the owners would prefer. This is reminiscent of moral hazard and adverse selection models where the principal induces less effort than in the first-best outcome in order to leave agents with fewer rents. The normative implication is different here however: society benefits from more ethical outcomes.

Beyond this insight, the Appendix derives predictions about ethics which have a natural parallel with moral hazard results:

- The owner prefers incentive schemes that minimize the amount of rents for the agents. In our context this means inducing competition between them. Therefore, exclusive territories, which suppress competition, are dominated for the owner by piece rates, and these are in turn dominated by tournaments (relative performance evaluation).
- When output incentive schemes are not available and the owner needs to rely on input monitoring (aka efficiency wages), equilibrium ethics rise when the agent has more bargaining power (because, say, she is essential and so does not fear much being fired).

8 Regulation

8.1 Optimal fee

(a) Socially optimal fee under free entry.

In the absence of constraints, setting the fee equal to marginal cost ($f = \gamma$) would kill two birds with one stone: The clients would pay a low price and the suppliers, having no stake would pick $u_i = \hat{u}$. Assume therefore that the suppliers have reservation utility $\bar{V} > 0$, and look for a free entry equilibrium for an arbitrary administered fee f . Adding an entry decision implies that a supplier can impact the overall outcome both through its entry decision and the ethical decision once in the market. We assume that the replacement effect operates fully at the entry stage; this amounts to assuming that if a supplier does not enter, another supplier will. Put differently, if in equilibrium n suppliers find it individually rational to enter, were one of them not to enter, another would replace him. Thus, the individual rationality constraint for the entrants is $\frac{f-\gamma}{n} + \alpha[W(u^*(n)) + \mathcal{D}(n)] \geq \bar{V} + \alpha[W(u^*(n)) + \mathcal{D}(n)]$ where $\mathcal{D}(n)$ measures the diversity benefits, with $\mathcal{D}' >$

0.[30] So let us rewrite the individual rationality constraint as:

$$V_i = \frac{f-\gamma}{n} \geq \bar{V}.$$

This constraint, which is binding, defines a function $n(f) \equiv (f-\gamma)/\bar{V}$.[31]

Letting $u^*(f,n)$ denote the equilibrium level of u, the social planner's objective function reflects the clients' welfare (suppliers have no surplus under free entry):

$$\mathcal{W}(f) \equiv W(u^*(f,n(f))) - f + \mathcal{D}(n(f)).$$

For example, in the Lerner-Salop generalization of the Hotelling model (n firms are symmetrically located on a circle of length 1, and consumers are uniformly distributed along the circle and have unit transportation cost t), $\mathcal{D}(n) = -t/4n$. We assume that $\mathcal{W}(f)$ is quasi-concave (a sufficient condition for this is that the diversity benefit $\mathcal{D}(n)$ be concave, as is the case in the Lerner-Salop model).

The social planner maximizes $\mathcal{W}$ subject to

$$-\alpha W'(u^*(f,n(f))) = \eta(f,n(f))L(f).$$

Let $f^R(\alpha)$ denote the optimal regulated fee. Absent ethical concerns ($\alpha = 0$), the optimal regulated fee $f^R(0)$ would satisfy

$$\mathcal{D}'(n(f^R(0)) = \bar{V}.$$

With ethical concerns ($\alpha > 0$), the optimal regulated fee $f^R(\alpha)$ further embodies its ethical impact. To see how, letting $\frac{du^*}{df} \equiv \frac{\partial u^*}{\partial f} + \frac{\partial u^*}{\partial n}\frac{dn}{df}$ and similarly for $d(\eta L)/df$, one has:

$$-\alpha W''(u^*)\frac{du^*}{df} = \frac{d}{df}(\eta L).$$

Next, from Assumption 2

$$\frac{d\mathcal{W}}{df} = \frac{[-W']}{\alpha W''}\frac{d}{df}(\eta L) - 1 + \frac{\mathcal{D}'(n(f^R(\alpha))}{\bar{V}} = 0$$

and so

$$f^R(\alpha) < f^R(0).$$

(b) HMO

[30] In theory it is possible that each of the n suppliers would prefer an outcome in which he refrains from entering and nobody replaces him. This would yield a partial replacement effect. Our assumption, which simplifies expressions, can be justified as a perfect equilibrium of a game in which firms take turn making an entry decision and do enter as long as the number of firms that have entered yet is smaller than the level, n, given by the free entry condition.

[31] As usual, we treat n as a continuous variable rather than as an integer, solely for conciseness purposes.

Consider now a health maintenance organization, which sells coverage to clients and contracts with health-care suppliers. So the price for coverage if $u^* + \mathcal{D}(n)$. Let us assume that the HMO faces the same participation constraint as the social planner. The HMO has no ethical concern and maximizes its profit

$$\pi \equiv u^* - f + \mathcal{D}(n)$$

subject to

$$\frac{f-\gamma}{n} \geq \bar{V}$$

and

$$-\alpha W'(u^*(f,n)) = \eta(f,n(f))L(f).$$

The same reasoning shows that, under Assumption 2, the fee set by the HMO satisfies

$$f^{\text{HMO}} > f^R(0).$$

Proposition 8 *The fees set by a social planner under ethical concerns ($f^R(\alpha)$) and no ethical concerns ($f^R(0)$), and by an HMO under ethical concerns (f^{HMO}) satisfy*

$$f^R(\alpha) < f^R(0) < f^{HMO}.$$

In particular, a social planner sets a low fee and reduces the intensity of competition so as to enlist the suppliers' ethical concerns, while an HMO panders to the clients and to do so raises the fee and the intensity of competition.

8.2 Pigovian policies

a) Design. Divergences between social welfare and individual goals call for Pigovian corrections when feasible. At an abstract level, and assuming that u_i is verifiable, the government might want to levy a tax $t(u_i - \hat{u})$ on "excessive attractiveness" $(u_i - \hat{u})$. For example, under regulated fees, supplier i would solve

$$\max_{\{u_i\}}[f - \gamma - t(u_i - \hat{u}) + \alpha[W(u_i) - W(u^*)]]D_i.$$

Simple computations show that to achieve the most ethical level $u^* = \hat{u}$, the tax should be set at level

$$t = \eta L$$

where $L \equiv (f-\gamma)/f$.

Example. Consider the overconsumption model (Example 1). While we expressed the tax t in terms of utils, it is more natural to express it in terms of quantities (number of prescriptions). Let τ denote the unit tax on "excess supply" $G(c_i^*) - G(b)$. Simple computations show that $\tau = [(1-\beta)b]t$. Supplier i maximizes

$$\Big[f - \gamma - \tau[G(c_i^*) - G(b)] + \alpha[W(c_i^*) - w^*]\Big]D(f - u(c_i^*), f - u(c^*))$$

with $W'(c_i^*) = (b - c_i^*)g(c_i^*)$ and $u'(c_i^*) = (b - \beta c_i^*)g(c_i^*)$. In the case of a *constrained fee*, the tax that delivers the optimal consumption $c_i^* = b$ is given by:

$$\tau = \eta L(1 - \beta)b. \tag{7}$$

It is instructive to see how this differs from the classical Pigovian prescription. We know from DellaVigna-Malmendier (2004) that the optimal Pigovian tax when the supply side is competitive and profit-maximizing exactly offsets the internality on the future self:

$$\tau = (1 - \beta)b.$$

For a *market-determined fee*, the first-order condition with respect to f_i is given by

$$\eta L = 1.$$

And so, unlike in the constrained-fee case, the optimal tax is the same as when the market is competitive and suppliers are selfish, i.e. the DellaVigna-Malmendier level:

$$\tau = (1 - \beta)b. \tag{8}$$

b) Limits to Pigovian policies. Pigovian policies require fine information about the parameters of the environment: Taxing a doctor for her "excessive prescription of opioids" requires a good knowledge of the doctor's baseline prescription level given her clientele, which is highly heterogeneous across doctors.[32] Pigovian policies also require keeping track of opioid supplies and preventing both a secondary market and an illegal market.[33] By contrast, the ability to correct various distortions through a consumption tax does not depend on the suppliers having all the same ethics. This can be seen either from (7) and (8), which do not depend on α, or from the suppliers' objective function ($W(c_i^*)$ is maximized at the consumption induced by the Pigovian, ethics-free tax).

Another reason for the absence of Pigovian corrections of the wedge between private and social incentives is that the government may not be pursuing social welfare and its Pigovian corollary, either because it does not want to (a case in point is given by the low carbon prices almost everywhere in the world) or because it cannot (it cannot regulate ethical choices abroad).

A third reason why governments may fail to apply Pigovian taxation arises when either the perpetrators or the victims of unethical choices (corruption, fiscal optimization, pollution) reside abroad.

[32] And taxing patients may conflict with health insurance.

[33] Replacing a tax on excess prescriptions by a uniform tax on opioids raises other problems. It reduces the doctors' profit, and so to keep the same number of doctors, one need to add a lump-sum subsidy and entry must be monitored so as to prevent fake or part-time doctors from pocketing the subsidy.

8.3 Self-regulation: Socially responsible clubs

An alternative to regulation is self-regulation. Here we have in mind the voluntary adoption of ESG criteria, emphasizing the "G" (for "governance") in the acronym. Prominent along the governance dimension since the 2008 financial crisis is the pledge to incentivize managers and traders in a more responsible way (fewer bonuses, introduction of clawbacks...).

We look at the following situation: There are n firms. All suppliers in the economy have the same ethical concerns parameter α. So suppliers' payoffs are $\frac{f-\gamma}{n} + \alpha W(u^*)$. Suppliers collectively would prefer to behave more ethically, but they individually cannot help misbehaving to gain market share. Could self-regulation lead to more ethical outcomes? Suppose that they can choose between two values $\hat{u}$ and $\bar{u}$ (with associated welfares $\hat{w}$ and $\bar{w}$), that the fee f is regulated and that suppliers can commit to $\hat{u}$ if they so desire. A mechanism for commitment might be to put the manager in charge on a fixed wage. Prices are regulated at some level f.

To be able to solve for equilibrium, we assume a linear demand system and a covered market. Suppose that m suppliers join the socially responsible club. Let $V_1(m)$ and $V_0(m)$ denote the payoff of owners who have (resp. have not) signed the responsible pledge. An equilibrium of the first stage requires at least that:

$$V_1(m) - V_0(m-1) \geq 0 \geq V_1(m+1) - V_0(m)$$

Let $Z \equiv \widehat{\sigma}(\bar{u} - \widehat{u})$. Simple computations show that demands facing members (D_1) and non-members (D_0) are:[34]

$$D_0(m) = \frac{1}{n} + Z\frac{m}{n-1} > D_1(m) = \frac{1}{n} - Z\frac{n-m}{n-1}$$

and that

$$\begin{aligned} V_1(m) - V_0(m-1) &= (f-\gamma)D_1(m) + \alpha[mD_1(m)\hat{w} + (n-m)D_0(m)\bar{w}] \\ &\quad - [(f-\gamma)D_0(m-1) + \alpha[(m-1)D_1(m-1)\hat{w} + (n-m+1)D_0(m-1)\bar{w}]] \\ &= -(f-\gamma)Z + \alpha(\widehat{w} - \bar{w})[\frac{1}{n} + Z\frac{2m-n-1}{n-1}] \end{aligned}$$

The incentive to join the club increases with the number of its members: the membership decisions are *strategic complements*. There may be multiple equilibria (everybody joins, no-one joins). The virtuous equilibrium ($m = n$) can exist only if financial stakes $f - \gamma$ are not too large:

$$f - \gamma \leq \alpha(\widehat{w} - \bar{w})[1 + \frac{1}{nZ}] \tag{9}$$

[34] Assume that $n \leq 1/Z$, so as to ensure that D_1 is always non-negative ($D_1(1) = \frac{1}{n} - Z$). Being the only virtuous firm in a low-concentration industry results in zero demand.

Conversely, there is an equilibrium in which no-one joins the club ($m = 0$) if and only if:

$$f - \gamma \geq \alpha(\widehat{w} - \bar{w})[\frac{1}{nZ} - 1] \tag{10}$$

The decision of whether to join the club depends on the size of the club in the following way:

- The loss in market share associated with better behavior is smaller if others are socially responsible as well. However, the market share gain of being unethical is larger if others are socially responsible. For linear demand these two effects cancel (the term $(f - \gamma)Z$ does not depend on m).
- From an ethical perspective, it is more important not to lose market share when others are non-ethical; this is the driver of strategic complementarity in our model.

Proposition 9 *(socially responsible clubs). In the binary-choice, covered-market, linear-demand specification, membership decisions are strategic complements. The fully inclusive equilibrium exists if (9) is satisfied, and a no-member equilibrium exists if (10) holds.*

9 Conclusion

Whether in markets or in organizations, our moral judgment is affected by our incentives. As opponents of market economies have long emphasized, our institutional context frames our ethical choices and is not neutral; and while economics rightly emphasizes that competition creates incentives to please the consumer, competition may also reduce ethical concerns. Our analysis shows that

1) Critics of the market are correct in arguing that competition among suppliers weakens their ethical concerns. The replacement excuse is more potent, the higher the number of competitors and the less differentiated they are.

2) However, competition, to the extent that it reduces prices, also lowers the suppliers' benefit from attracting consumers in ethically-dubious ways. Thus high administered prices induce less ethical behavior than competitively determined ones, not more.

3) When suppliers differ in their ethical concerns, constrained fees work to the advantage of low-ethics suppliers. By contrast, price competition has interesting compositional benefits: More ethical suppliers, who have lower market shares under constrained prices, regain the upper hand, as they charge lower prices and more than make up for their non-price handicap.

Market advocates can "fight back" in two other ways. First, regulation (e.g. carbon pricing) may restore some alignment between private and social incentives; regulation is a complement to the market. Rephrasing the same point, the shareholder value paradigm is less appealing in industries where non-price competition raise ethical issues (externalities, internalities and shrouded attributes); and there is a complementarity between shareholder value and public regulation.

Second, attacks against the market economy must envision its counterfactuals. In that respect, our study of ethics in markets and hierarchies only begins to analyse the general issue of the morality of markets, raised in the 18$^{\text{th}}$ century and very prominent in the political discourse since the financial crisis and the backlash against globalization. More broadly, the unifying theme of our analysis is the relationship between ethical behavior and the power of incentives. The market is only a specific system of incentives. Within organizations, different ways of incentivizing managers and workers lead to different ethics. And who runs the organizations matter; we for instance derived the theoretical prediction that NGOs will offer lower-powered incentives to their agents.

Let us conclude with a few of the many issues of interest left open by our paper:

Self-selection into professions and organizations. Agents with different ethical concerns are likely to select different professions.[35] The ethical composition will depend on the rents attached to the professional choice: The "caring profession" may no longer be so caring with a rise in doctors' variable compensation; similarly, NGOs may no longer behave as nicely once people understand that a stint at an NGO may be a springboard for their career. Would one expect the more ethical types to opt for professions with high levels of externalities/internalities/deception opportunities? Furthermore, moral wiggle room may affect choices: While it may be the duty of a high-ethics agent to enter a low-ethics industry to make it more ethical, he may find the required sacrifice unappealing. Finally, would regulatory constraints on high-powered incentives, such as the ones imposed for bank managers after the 2008 financial crisis, lead owners to pick less ethical managers in an attempt to restore profitability?

Social norms. There is a fair amount of evidence showing that moral behavior is driven not only by intrinsic and extrinsic incentives, but also by self- and social-image concerns. Because social preferences have been assumed to be common knowledge, there was no loss of generality in ruling out such image concerns. One could by contrast allow for privately-known social preferences and add a term, $\mu E[\alpha_i \mid u_i]$,that would reflect both an intensity of image concerns, μ, and inferences about the individual's empathy, α_i, as a function of his action u_i. While the introduction of image concerns would not affect the main insights of this paper, it would open the door to social norms based on "what is deemed acceptable in society"[36].

[35] As emphasized by Prendergast (2007). See also Brekke-Nyborg (2008) and Barigozzi and Buranib (2019). People with different ethical concerns may also self-select when firms in the same industry offer differentially structured compensation schemes. Kosfeld and von Siemens (2011) study the screening of employees with unobserved other-regarding preferences when teamwork matters.

[36] As in, say, Bénabou-Tirole (2006, 2013).

Leading by example. Suppliers may have information on the gravity of unethical behavior, i.e. on the magnitude of the externality. By moving first, they may signal to other suppliers that unethical behavior is just unacceptable (for example by withdrawing from the market altogether); and even if the other suppliers are also informed about the size of the externality, society may not be, and the revelation that "there are other ways to run the business" may put social pressure on these other suppliers[37].

Customers' ethical concerns. Section 6 showed how the "delegated philanthropy" view of CSR interacts with the "corporate philanthropy" view of CSR. Introducing heterogeneous ethical concerns on the suppliers' and customers' sides would provide a richer picture. One would expect some "assortative matching" with more ethical suppliers enjoying a higher market share with ethical clients. The analysis of the impact of this partial specialization under heterogeneous clienteles on overall ethical behavior would be of interest.

References

Aghion, P., Bénabou, R., Martin, R., and A. Roulet (2019), "Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?," mimeo.

Alpert, A., Powell, D. and R. Liccardo Pacul (2018), "Supply-Side Drug Policy in the Presence of Substitutes: Evidence from the Introduction of Abuse-Deterrent Opioids", *American Economic Journal: Economic Policy*, 10(4): 1–35.

Anderson, E. (1993), *Value in Ethics and Economics*, Harvard University Press.

Bala, R., and P. Bhardwaj (2010), "Detailing vs. Direct-to-Consumer Advertising," *Management Science* 56: 148–160.

Barigozzi, F., and N. Buranib (2019), "Competition for Talent when Firms' Mission Matters," *Games and Economic Behavior*, 116: 128–151.

Bartling, B., and Y. Özdemir (2017), "The Limits to Moral Erosion in Markets: Social Norms and the Replacement Excuse," Mimeo.

Bénabou, R. Falk, A. and J. Tirole (2019), "Narratives, Imperatives and Moral Persuasion", mimeo.

Bénabou, R. and J. Tirole (2006), "Incentives and Prosocial Behavior," *American Economic Review*, 96(5): 1652-1678.

Bénabou, R., and J. Tirole (2010), "Individual and Corporate Social Responsibility," Coase lecture, *Economica*, 77: 1–19.

Bénabou, R., and J. Tirole (2013), "Laws and Norms," mimeo.

Bénabou, R., and J. Tirole (2016), "Bonus Culture: Competitive Pay, Screening and Multitasking," *Journal of Political Economy*, 124(2): 305–370.

[37]This is reminiscent of both Hermalin (1998)'s theory of leading by example and the informed principal approach in Bénabou-Tirole (2013).

Besley, T., and T. Persson (2019), "Escaping the Climate Trap: Values, Technology and Politics,", mimeo.

Bonnefon, J.F., Landier, A., Sastry, P. and D. Thesmar (2019), "Do Inverstors care about Corporate Externalities? Experimental Evidence", mimeo.

Bowles, S. (1998), "Endogenous Preferences: The Cultural Consequences of Markets and Other Institutions," *Journal of Economic Literature*, 36: 75–111.

Brekke, K., and M. Kuhn (2006), "Direct to Consumer Advertising in Pharmaceutical Markets," *Journal of Health Economics*, 25: 102–30.

Brekke, K.A., and K. Nyborg (2008), "Attracting Responsible Employees: Green Production as Labor Market Screening," *Resource and Energy Economics*, 30: 509–526.

Brodback, D., Guenster, N. and S. Pouget (2019), "The Valuation of Corporate Social Responsability: An Experiment," mimeo.

Choi, J.P., Jeon, D.S., and B.C. Kim (2019), "Privacy and personal data collection with information externalities," *Journal of Public Economics*, 173: 113–124.

Dana, J., Weber, R., and J. Kuang (2007) "Exploiting Moral Wriggle Room: Experiments Demonstrating an Illusory Preference for Fairness," *Economic Theory*, 33: 67-80.

DellaVigna, S. and U. Malmendier (2004), "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, 119(2): 353–402.

Diamond P.A. (1971), "A Model of Price Adjustment," *Journal of Economic Theory*, 3(2): 156–168.

Dufwenberg, M., Heidhues, P., Kirchsteiger, G., Riedel, F. and J. Sobel (2011), "Other-Regarding Preferences in General Equilibrium," *Review of Economic Studies*, 78: 613–639.

Falk, A., and N. Szech (2013), "Morals and Markets," *Science*, 340(6133): 707–711.

Falk, A., and N. Szech (2017), "Diffusion of Being Pivotal and Immoral Outcomes," under revision at *Review of Economic Studies*.

Gabaix, X., and D. Laibson (2006), "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 121(2): 505–540.

Hart, O., and L. Zingales (2017), "Companies Should Maximize Shareholder Welfare Not Market Value," *Journal of Law, Finance, and Accounting,* 2: 247–274.

Hermalin, B. (1998), "Toward an Economic Theory of Leadership: Leading by Example", *American Economic Review*, 88: 1188-1206.

Holmström, B., and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 7: 24–52.

Inderst, R., and M. Ottaviani (2012), "Competition through Commissions and Kickbacks," *American Economic Review*, 102(2): 780–809.

Kosfeld, M. and F. Von Siemens (2011), "Competition, Cooperation and Corporate Culture," *Rand Journal of Economics*, 42: 23-43.

Jäger, S., Schoefer, B., and J. Heining (2019), "Workers in the Board Room: The Causal Effects of Shared Corporate Governance," mimeo.

Laffont, J.J., and J. Tirole (1991), "Provision of Quality and Power of Incentive Schemes in Regulated Industries," in *Equilibrium Theory and Applications (Proceedings of the Sixth International Symposium in Economic Theory and Econometrics*, J. Gabszewicz and A. Mas-Colell (eds.), Cambridge: Cambridge University Press, pp. 161-196.

Moisson, P.-H. (2019), "Socially Responsible Investment, Environmental Consciousness and the Green Premium," mimeo.

O'Donoghue, T., and M. Rabin (2006), "Optimal Sin Taxes", *Journal of Public Economics* 90: 1825–1849.

Prendergast, C. (2007), "The Motivation and Bias of Bureaucrats," *American Economic Review*, 97: 180–196.

Sandel, M. J. (2012), *What Money Can't Buy: The Moral Limits of Markets*, Macmillan.

Satz, D. (2010), *Why Some Things Should Not Be For Sale: The Moral Limits of Markets*, Oxford University Press.

Shapiro, C. (1983), "Premiums for High Quality Products as Returns to Reputation," *Quarterly Journal of Economics*, 98(4): 659–680.

Shleifer, A. (2004), "Does Competition Destroy Ethical Behavior?," *American Economic Review, Papers & Proceedings*, 94: 414–418.

Sobel, J. (2015), "Do Markets Make People Selfish?" Mimeo.

Spence, A.M. (1975), "Monopoly, Quality, and Regulation," *The Bell Journal of Economics*, 6(2): 417–429.

Tirole, J. (2001), "Corporate Governance," *Econometrica*, 69(1): 1–35.

Tirole, J., (2019), "Assumptions in Economics," mimeo.

Walzer, M. (2008), *The Free Market and Morality* on YouTube.

Weyl, G. and M. Fabinger (2013), "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, 121(3): 528–583.

Appendix

Multiplicity of equilibria when Assumption 3 is violated

Equation (1) is guaranteed to have a unique solution if $\frac{\eta'}{\eta} < \frac{W''}{W'}$, which is for instance satisfied for CES and Hotelling-Lerner-Salop demand functions. But there may exist situations in which a supplier is keener to behave ethically when other suppliers also do and multiple equilibria co-exist (in which case the "most ethical equilibrium" Pareto dominates the others[38]).

To illustrate the potential multiplicity, assume that the market is covered and that a uniform change in fees leaves all demands invariant (which is a necessary implication of full coverage in discrete choice demand models, since a uniform change in fee does not alter the relative attractiveness of products). Finally, suppose that there are *two levels of utility*, say $\bar{u}$ and $\widehat{u}$. For these two levels to be both equilibria, it must be the case that

$$\frac{f-\gamma}{n} \geq [f-\gamma-\alpha(\widehat{w}-\bar{w})]D(f-\bar{u}, f-\widehat{u}) \quad \text{and} \quad \frac{f-\gamma}{n} \geq [f-\gamma+\alpha(\widehat{w}-\bar{w})]D(f-\widehat{u}, f-\bar{u}).$$

We now show the following:

Suppose that $n = 2$ or/and demand is linear. Then, in the binary utility case, as f increases from γ to $+\infty$,

- *first, the only equilibrium is the ethical one ($u^* = \widehat{u}$);*
- *then both levels of utility are sustained in symmetric equilibrium;*
- *finally, for large fees, the only equilibrium is the unethical one ($u^* = \bar{u}$).*

Proof. As we noted, demands are independent of f. For conciseness let $D^+ \equiv D(f - \bar{u}, f - \widehat{u}) > \frac{1}{n} > D^- \equiv D(f - \widehat{u}, f - \bar{u})$. The two conditions can be rewritten as:

$$1 \geq [1 - \frac{\alpha(\widehat{w} - \bar{w})}{f - \gamma}]nD^+$$

and

$$1 \geq [1 + \frac{\alpha(\widehat{w} - \bar{w})}{f - \gamma}]nD^-$$

Note that the RHS of the first inequality increases with $f \in (\gamma, \infty)$ from $-\infty$ to $nD^+ > 1$, while the RHS of the second inequality decreases from $+\infty$ to $nD^- < 1$.

[38] To define Pareto dominance, one must look at the suppliers' gross utilities, and not the net utilities that are used to derive optimal strategies; indeed the latter are equal to $(f - \gamma)/n$ in all symmetric equilibria.

To conclude the proof it suffices to show that when, say, the first inequality is satisfied with equality, the second inequality is strictly satisfied. Suppose, first, that $n = 2$. Then $D^+ + D^- = 1$ (as the market is covered). The result then follows from the property that $D^+ > 1/2$. When n is arbitrary, but demand is linear ($\tilde{D}_i = \frac{1}{n} - \hat{\sigma}\left[(f-u_i) - (\frac{\sum_{j\neq i}(f-u_j)}{n-1})\right]$), the reasoning goes as follows: define p as the "average net price" $p \equiv f - \frac{\bar{u}+\widehat{u}}{2}$ and rewrite $D^+ = \frac{1}{n} - \hat{\sigma}(p-\varepsilon) + \hat{\sigma}(p+\varepsilon)$ and $D^- = \frac{1}{n} - \hat{\sigma}(p+\varepsilon) + \hat{\sigma}(p-\varepsilon)$. Using the facts that the two inequalities are satisfied as equality for exactly the same f when $\varepsilon = 0$, that $\frac{\partial D^+}{\partial \varepsilon} = 2\hat{\sigma} = -\frac{\partial D^-}{\partial \varepsilon}$ and that $D^+ > D^-$, we obtain the result announced in the text. ■

Global second-order condition under market-determined fees

Let us check that the tentative equilibrium is a global optimum for each supplier. Equilibrium behavior requires that there be no (f_i, u_i) such that

$$(f^*-\gamma)D_i(f^*-u^*, f^*-u^*) < [f_i - \gamma + \alpha[W(u_i) - W(u^*)]]\, D_i(f_i-u_i, f^*-u^*) \equiv \mathcal{V}(f_i, u_i).$$

The concavity of W and condition (3) imply that

$$\alpha[W(u_i) - W(u^*)] \leq -[u_i - u^*].$$

So

$$\mathcal{V}(f_i, u_i) \leq (f_i - \gamma - u_i + u^*)D_i(f_i - u_i, f^* - u^*).$$

We also know that for all $\hat{f}_i$, $(f^*-\gamma)D_i(f^*-u^*, f^*-u^*) \geq (\hat{f}_i-\gamma)D_i(\hat{f}_i-u^*, f^*-u^*)$. Applying this to $\hat{f}_i = f_i - u_i + u^*$ yields

$$(f^* - \gamma)D_i(f^* - u^*, f^* - u^*) \geq (f_i - \gamma - u_i + u^*)D_i(f_i - u_i, f^* - u^*),$$

a contradiction. ■

Proof of Proposition 4

We consider the two first-order conditions (4) for suppliers i and j, and we substract one from the other. We obtain (using $f_i = f_j = f$):

$$H(\alpha_i, \alpha_j) \equiv \hat{\sigma}(\alpha_i + \alpha_j)(w_i - w_j) + \alpha_i W'(u_i)D_i - \alpha_j W'(u_j)D_j = 0.$$

We know that $H(\alpha_i, \alpha_i) = 0$ (as $w_i = w_j$, $u_i = u_j$, $D_i = D_j$).

Let us keep α_i constant and increase α_j above α_i:[39]

$$\begin{aligned} dH &= \hat{\sigma}(\alpha_i + \alpha_j)W'(du_i - du_j) + \alpha_i W'\hat{\sigma}(du_i - du_j) \\ &\quad -\alpha_j W'\hat{\sigma}(du_j - du_i) - W'Dd\alpha_j - \alpha_j DW''du_j + \alpha_i DW''du_i \\ &= 0 \end{aligned}$$

[39] We omit the index (i, j) when the variable has the same value for both firms.

$$\Leftrightarrow\ [4\hat{\sigma}\alpha W' + \alpha D W'''][du_i - du_j] = W' D d\alpha_j$$

and so for $d\alpha_j > 0$, $du_i > du_j$: The less ethical supplier has a higher market share (as $u_i - f > u_j - f$).

More generally, at any point $\alpha_j > \alpha_i$ such that u_j were equal to u_i (and so $D_i = D_j$, $w_i = w_j$, etc.),

$$\frac{dH}{d\alpha_j} = 0 \quad \Leftrightarrow \quad 2\hat{\sigma}(\alpha_i + \alpha_j)W'(du_i - du_j) + DW''(\alpha_i du_i - \alpha_j du_j) = W' D d\alpha_j < 0.$$

Imagine that $du_j > du_i$ (which is required for u_j to overcome u_i from below). Then the left-hand side of this equality is positive (implying a contradiction), unless $\alpha_i du_i > \alpha_j du_j$, which requires $du_i < 0$.

But differentiate (4):

$$\hat{\sigma}\alpha_i W'(du_i - du_j) + \alpha_i W'' D du_i + \alpha_i W' \hat{\sigma}(du_i - du_j) = 0$$

or

$$[-2\hat{\sigma}\alpha_i W'][du_j - du_i] = [-\alpha_i W'' D] du_i,$$

which is impossible as the LHS is positive and the RHS negative. ■

Alternative moral imperatives

We have adopted a consequentialist, impact-based approach: Suppliers care about the actual impact of their decisions. There are of course alternative moral attitudes. A celebrated one is rule consequentialism, the imperative to act only in accordance with the maxim through which one can at the same time will that it become a universal law.[40] Supplier i imagines that her behavior u_i is mimicked by other suppliers ($f_j = f_i, u_j = u_i$). Let $D^+(f-u) \equiv nD(f-u, f-u)$. Supplier i solves

$$(f_i - \gamma)\frac{D^+(f_i - u_i)}{n} + \alpha[W(u_i)D^+(f_i - u_i) + w_0[1 - D^+(f_i - u_i)]].$$

If the market is covered ($\Sigma_i D_i = 1$), then (letting u^K denote, by an abuse of terminology, the "Kantian level")

$$u^K = \hat{u}.$$

When the market is not covered ($\Sigma_i D_i < 1$), and the fee is regulated ($f_i = f$) then u^* is given by

$$\left[\frac{f-\gamma}{n} + \alpha[W(u^*) - w_0]\right]\left(\frac{-\partial D^+}{\partial f}\right) + \alpha W'(u^*)D^+ = 0.$$

[40] See Brekke-Nyborg (2008) for one of the early formal treatments of Kantian approaches for ethical behavior.

Assuming that a uniform price increase decreases individual demands,[41] one obtains

$$\hat{u} < u^K < u^*.$$

We thus see that rule consequentialism disposes of the replacement excuse and so nullifies the impact of competition intensity that is so prominent under purely consequentialist preferences. We also showed that there is still unethical behavior if the market is not covered. The latter result hinges on the clients' willingness to take an unethical outside option. Rule consequentialism reasoning is sensitive to the contours of the set of agents to whom it is applied.

This calls to a broader discussion of consequentialism. This question is age-old and a proper treatment lies outside the limited scope of this paper. Let us note, though, that it has two dimensions: positive and normative. On the positive front, consequentialist preferences receive strong support[42]; for example individuals are more prone to behave morally when externalities are large; and - more to the point for this paper- the highly-consequentialist replacement excuse is a popular one. But there are violations of consequentialism. A common "violation" is associated with the now-well-developed notion of "moral wiggle room"[43] . Often individuals are pivotal in decisions with moral overtones, but manage to distort reality so as not to feel responsible for otherwise-unethical behavior (as if they were not pivotal). The violation of consequentialism is an observed violation, rather than a conceptual one: It is precisely because individuals assess moral behavior through the lens of the consequences of their actions that they distort reality. Another apparent violation relates to the role of intentions: suppose that I meant to hurt you, but by clumsiness failed to do so. There is then a signaling aspect to my choice, that hinges on the consequences that would have prevailed, had I been more skillful[44].

On the normative front, economists' and some other scientists (such as philosopher Peter Singer)'s utilitarian perspective calls for embodying the consequences of one's acts into one's moral judgment. This approach leads to concepts such as "impact behavior", according to which we should pick philanthropic actions (as investors, consumers or workers) that have a strong impact onto others (rather than, say, posturing actions that have a strong impact on one's social image). But advocates of a consequentialist approach still must confront the following paradox: in the context of the replacement excuse, what they deem an intrinsically moral stance (caring about the size of externalities) delivers an immoral outcome!

[41] Which is necessarily the case when consumers choose between an outside option and consuming one of the n products.

[42] See e.g. references in Bénabou et al (2019).

[43] See the large literature that followed Dana et al (2007).

[44] See e.g. Tirole (2019) for further discussion.

Hierarchies : explicit incentives that overcome agents' ethical objections

There have been numerous cases of companies incentivizing their agents to ignore ethics and raise profit. In the Wells Fargo account fraud scandal, the pressure and sales incentives from higher-level management to open as many accounts as possible through cross-selling pushed branch workers and managers to engage in ethically dubious behavior. Similarly, Goldman Sachs incentivized its sales people to sell products it knew to be toxic, and Dexia did the same with municipalities. These banking examples relate to the "shrouded attribute" justification of the wedge.[45] Another case in point is the opioid crisis. It is reported that Purdue paid $40 million in sales incentive bonuses to promote its sales.[46] We therefore study how suppliers can overcome a potential resistance of their agents to behave unethically.[47]

Consider a hierarchy composed of a principal (a firm, a platform) and agents (employees or sales representatives). The principal employs exactly n agents (which we label $i \in \{1, \ldots, n\}$). The potential agents, in number at least $n+1$, have initial wealth A each, are risk-neutral and have the same ethical-concerns parameter α. Each of the n agents chooses a level u_i. The principal provides the agents with symmetric incentives (more on this shortly), and all agents choose the same u^* in equilibrium. An agent joins the firm only if his compensation is at least equal to A. In particular, agents do not internalize "ex ante" ethical concerns, as the replacement logic plays fully at that stage: the welfare outcome will be $w^* = W(u^*)$ regardless of whether the agent joins the firm or not. In contrast, once in the firm, the replacement logic is only partial, which is why the gross surplus u^* may exceed $\hat{u}$.

Consumers pay a fee f and perceive net utility $u^* - f$; let $\mathcal{D}(n)$ denote the diversity benefit/expected distance to one's preferred agent. So, consumers are willing to join the firm/platform if and only if $u^* - f + \mathcal{D}(n) \geq u_0$. This modeling implicitly assumes that consumers do not know their "location" in the product space when choosing between the firm/platform and the outside option[48]. So the timing is: 1) The platform sets fee f as well as the agents' incentive scheme. 2) Consumers decide whether to join the platform. 3) Consumers who have joined the firm/platform learn their idiosyncratic preferences and choose the agent who will supply them.

[45] In the case of Dexia, the more finance-savvy municipalities probably wanted to benefit from low-interest conditions in the short term or to gamble with municipality money. The justification is then more that of an externality on the residents.

[46] The non–cancer-related pain market constituted 86% of the total opioid market in 1999.

[47] Both the suppliers' eagerness to overcome agent's ethical resistance and the form of the other-regarding preferences are reminiscent of Inderst-Ottaviani (2012). In their paper, multiple suppliers set commissions for an agent in order to induce the latter to steer the consumer toward their own product. The focus of Inderst-Ottaviani is on the level of commissions (which decrease with α, in our notation) and on the impact of disclosure. It is there rather different from the current one.

[48] The analysis can easily be generalized to consumers knowing their location, giving rise to a downward sloping demand.

To simplify the exposition, we specialize the model to a linear demand system: In the relevant range ($D_i \in [0,1]$),

$$D_i = \frac{1}{n} - (n-1)\sigma(f_i - u_i) + \sigma\left[\sum_{j\neq i}(f_j - u_j)\right].$$

The principal is a monopolist[49] without ethical concern (the analysis can be extended along the lines of Section 7.1 to accommodate two-tier ethical concerns). Given proposed utility u^*, the principal can set fee $u^* - f + \mathcal{D}(n) = u_0$, yielding gross profit

$$\pi(u^*) \equiv u^* - \gamma + \mathcal{D}(n) - u_0 - nT,$$

where T is the net compensation (transfer) paid to an agent. Her unconstrained optimum is therefore $\{u = \bar{u}, f = \bar{u} - u_0 + \mathcal{D}(n)\}$. However the choices of levels u_i of gross utility are delegated to the agents, whose actual behavior hinges on the incentive scheme they face. The monopolist maximizes $\pi(u^*)$.

(a) *Exclusive territories.* Suppose, first, that the principal allocates $1/n$ of the demand to each hired agent and pays a fixed wage. The diversity benefit under exclusive territories (ET) satisfies:

$$\mathcal{D}^{\mathrm{ET}}(n) \leq \mathcal{D}(n).$$

In the Hotelling-Lerner-Salop model, $\mathcal{D}^{\mathrm{ET}}(n) = \mathcal{D}(n)$ if the territories are allocated according to distance to supplier. But in general incomplete information about consumer location in the taste space implies imperfect matching between supplier and consumer under exclusive territories.

Second, note that the absence of competition for market share implies that $u_i = \hat{u}$. Finally, $T = 0$: Each agent knows that, were she to reject the offer made by the principal, the agent who would replace her would pick $\hat{u}$ as well, and so $\alpha(w_i - w^*) = \alpha[W(\hat{u}) - W(\hat{u})] = 0$. Thus the principal's profit is

$$\pi^{\mathrm{ET}} = \hat{u} - \gamma + \mathcal{D}^{\mathrm{ET}}(n) - u_0.$$

(b) *Piece-rate incentives.* Suppose that each agent receives an amount $r \in \mathbb{R}^+$ per unit sold, plus a lump-sum payment R (without loss of generality, the agent's wealth A is paid to the firm). The agent's participation constraint is then $\frac{r}{n} + R \geq A$.

Agent i's behavior is dictated by:

$$\max_{\{u_i\}}[r + \alpha[W(u_i) - w^*]]D_i(f - u_i, f - u^*).$$

First, let us show that in equilibrium $R = 0$. If not, agents have no rent ($\frac{r}{n} + R = A$) and the firm solves $\max_{\{u*\}} u^* - \gamma - u_0$, yielding $u^* = \bar{u}$. But implementing $\bar{u}$ requires $r = +\infty$, a contradiction. And so $T = \frac{r}{n} - A$.

[49]The same qualitative insights would hold if the principal were competing in a product market or if she had (limited) ethical concerns herself.

For a linear demand system and in a symmetric equilibrium, the first-order condition is:

$$-\alpha W'(u^*) = n(n-1)\sigma r.$$

For a given f, u^* is an increasing function of r, $u^*(r)$, with $u^*(0) = \hat{u}$ and $u^*(\infty) = \bar{u}$. Let $u^\dagger$ be defined by $\frac{-\alpha W'(u^\dagger)}{n^2(n-1)\sigma} = A$. Either $u^* \geq u^\dagger$, and then the firm selects $r^* = \frac{-\alpha W'(u^*)}{n(n-1)\sigma}$; or $u^* < u^\dagger$ and the firm selects $r^\dagger = \frac{-\alpha W'(u^\dagger)}{n(n-1)\sigma}$.

The monopolist's profit is then

$$\max_{\{r \in \mathbb{R}^+\}} u^*(r) - \gamma - u_0 - r - nA,$$

yielding

$$\frac{du^*}{dr} = 1 \Leftrightarrow -\alpha W''(u^*) = n(n-1)\sigma.$$

The optimal piece rate is $r^* = \frac{-\alpha W'(u^*)}{n(n-1)\sigma}$.

In words, the monopolist must set a low piece rate so as to avoid leaving rents to the agents. But low piece rates in the presence of ethical agents induce ethical behavior and low profits. Conversely, a high piece rate overcomes the agents' ethical concerns but are expensive for the principal.

Finally, note that the profit under piece-rate incentives, π^{PR}, satisfies:

$$\pi^{\text{PR}} \geq u^*(nA) - \gamma + \mathcal{D}(n) - u_0 > \pi^{\text{ET}}.$$

Thus piece-rate incentives are always preferred to exclusive territories.

(c) *Relative performance evaluation.* Suppose now that agent i is paid according to the difference between her performance and the average performance of her colleagues:

$$T_i = R + r\left[D_i - \frac{\sum_{j \neq i} D_j}{n-1}\right].$$

Note that in symmetric equilibrium, the variable compensation is equal to zero.

Agent i solves

$$\max_{\{u_i\}} r\left[D_i - \frac{\sum_{j \neq i} D_j}{n-1}\right] + \alpha\left[w_i D_i + \sum_{j \neq i} w_j D_j\right].$$

The first-order condition is:

$$-\alpha W'(u^*) = n^2 \sigma r.$$

Now the monopoly can raise the incentive rate r to promote arbitrarily high unethical behavior ($u^* = \bar{u}$) without leaving rents to the agents. The one deviation from equilibrium

behavior we have not yet checked is that associated with limited liability. Given the very-high-powered incentive scheme, deviations from $\bar{u}$ imply the loss of A. At the same time, the deviating agent's market share is at most $1/n$. Hence, a *sufficient* condition for a deviation exploiting limited liability not to be optimal is:

$$A \geq \frac{\alpha(\hat{w} - \bar{w})}{n}.$$

Proposition 10 *(managing ethical agents).*

(i) Under exclusive territories, the principal induces ethical, but low-profit behavior ($u^* = \hat{u}$) *and may further lose from restricting consumer choice.*

(ii) Under piece-rate incentives, the firm cannot reach the upper bound on its profit $\bar{u} - \gamma - u_0$*. It is confronted with a trade-off: A high piece rate overcomes the agents' ethical concerns but are expensive for the principal. The optimal piece rate* r *is* $\max\{r^\dagger, r^*\}$.

(iii) Under relative performance evaluation, a sufficient condition for the firm to reach the upper bound on its profit is that $A \geq \frac{\alpha(\hat{w}-\bar{w})}{n}$.

The section's running theme is that abandoning ethical behavior is akin to an "effort". There is a psychological cost of behaving badly, with the caveat that the ethical cost is proportional to demand. The tradeoff between effort and rent extraction for piece-rate schemes can be interpreted in this light, and so is the value of tournaments to overcome this tradeoff.

Hierarchies : implicit incentives (following orders)

Should a subordinate do what is expected from her even when this conflicts with her moral judgment? How responsible is the subordinate when harm results from her behavior? To address such questions, we now formalize a notion of "power" as the capacity for the firm to overcome subordinates' ethical concerns and exact obedience from them. Suppose that monetary incentives are ruled out and that the principal sets some expectation u^* for subordinates' behavior. Subordinates then select their policy under the threat of dismissal. Dismissal involves a known cost $c_P > 0$ for the principal, and a cost c_A for the agent, where c_A is drawn in an i.i.d manner from cumulative distribution $F(c_A - \theta)$, where θ shifts the distribution toward higher costs (the cost for the agent of a court martial differs substantially from that associated with returning to a fluid labor market). The distribution F satisfies the monotone hazard rate property ($(1-F)/f$ is decreasing).

The timing thus goes as follows: 1) The principal sets fee f and selects some expectation u^*. 2) Each agent learns her own cost of dismissal and decides whether to comply. 3) Agents who did not comply are dismissed.

When deciding not to follow the order, an agent optimally chooses $u_i = \hat{u}$.

Assume a single agent for simplicity. The agent does not follow orders if $c_A < \alpha[\hat{w} - W(u^*)]$. Using the same model as in Section 9 to depict the consumers' reservation utility, profit maximization amounts to:

$$\max_{\{u*\}} \{-\gamma - u_0 + (\hat{u} - c_P)F(\alpha[\hat{w} - W(u^*)] - \theta) + u^*[1 - F(\alpha[\hat{w} - W(u^*)] - \theta)]\}.$$

The first-order condition is:

$$[-\alpha W'(u^*)](u^* - \hat{u} + c_P) = \frac{1 - F(\alpha[\hat{w} - W(u^*) - \theta])}{f(\alpha[\hat{w} - W(u^*)] - \theta)}.$$

This defines a unique u^*.

Proposition 11 *(following orders). Orders u^* are less ethical, the more costly being fired for the agent (the higher θ is), the lower the agent's ethical concern (α), and the less essential the agent is (the lower the cost of separation c_P for the principal).*

As in Proposition 10, the parallel with "effort" is relevant here, except that we consider input monitoring in this subsection, like in efficiency-wage models, rather than output monitoring. Intuitively, the outcomes become closer to pure profit maximization if the agent is less ethical or if she is more easily 'directed' towards profit maximization by monitoring of her effort.