What's the Point of Efficiency? On Heath's Market Failures Approach

Richard Endörfer
Louis Larue
University of Gothenburg, Sweden

This article reviews and criticizes Joseph Heath's market failures approach (MFA) to business ethics. Our criticism is organized into three sections. First, we argue that, even under the ideal assumptions of perfect competition, when markets generate Pareto-efficient distributions, Heath's approach does not rule out significant harms. Second, we show that, under nonideal conditions, the MFA is either too demanding, if efficiency is to be attained, or not sufficiently demanding, if the goal of Pareto efficiency is abandoned. Finally, we argue that Heath's appeal to regulations and specific moral requirements as a remedy for market failures is unlikely to safeguard efficiency and exposes a number of general worries regarding the moral force of the MFA. We end this article with a constructive suggestion on how to adjust the MFA to avoid these problems while preserving its contractualist and Paretian spirit.

Key Words: Joseph Heath, market failures approach, efficiency, contractualism, market harm

Joseph Heath (2004, 2014) has proposed and defended one of the most prominent theories of business ethics in recent years. The market failures approach (MFA), as he called it, is gaining in influence (Norman 2011) and has recently been subject to some criticism (Cohen and Peterson 2019; Moriarty 2019; Steinberg 2017; Hsieh 2017; von Kriegstein 2016; Smith 2018). In this article, we aim to expand the literature on the MFA by discussing crucial objections that have, to the best of our knowledge, not yet been mentioned in the literature.[1]

Let us begin by summarizing the central tenets of the MFA. Similar to some of its rivals, such as stakeholder (Freeman 1984) and shareholder theory (Friedman 1970), the MFA is concerned primarily with the professional moral duties of managers in a market economy (Heath 2014, 69). Heath makes two central claims. First, he argues that the ultimate function of the market is to achieve Pareto-efficient distributions of goods and services. Hence Heath usually refers to the MFA as a "Paretian" theory of business ethics (5). Second, and consequently, he claims that the moral duty of managers (and possibly of other market participants

[1] McMahon (1981) has defended a similar theory of business ethics. Our focus here is on specific weaknesses of Heath's theory, which may not all apply to McMahon's.

Business Ethics Quarterly (2022), pp. 1–25. DOI:10.1017/beq.2022.21
Published by Cambridge University Press on behalf of the Society for Business Ethics.
© The Author(s), 2022.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

as well)[2] is to abide by a set of imperatives derived from what Heath calls the "Pareto conditions," that is, the conditions that must obtain if markets are to generate Pareto-efficient distributions (37). The common core of these imperatives can be summed up in a motto: "Do not exploit market failures."

Hence Heath's entire theory revolves around Pareto efficiency.[3] A distribution is Pareto-efficient if no one can be made better off (in terms of welfare, preference satisfaction, or else) without making someone else worse off. Closely related to the concept of Pareto efficiency is the concept of Pareto improvement (to which we will return in a later section of this article). Pareto improvements are shifts from one distribution to another, such that at least one person is made better off, while no one is made worse off. Heath subscribes to the view that an economic system that generates Pareto-efficient distributions is to the benefit of some (or most), while it is to the detriment of none. According to Heath, appeals to efficiency are inherently less controversial than appeals to, for example, equality. Whereas appealing to equality typically requires some form of redistribution, such that some gain while others lose, appealing to efficiency entails promoting distributions in which no one can be made better off without making anyone worse off. As Heath (2014, 2) puts it, "efficiency arguments [contrary to equality arguments] appeal to mutual benefit, or win–win transformations, and so do not necessarily create an oppositional constituency." The possibility of win–win transformations is hence of central importance in advocating in favor of the market over other economic arrangements, because they are an integral part of Heath's wider vision of a "minimally controversial contractualism" (155–56).

In economic theory, the so-called first fundamental welfare theorem (also referred to by Heath as the invisible hand theorem) demonstrates that under a number of idealizing assumptions, markets necessarily generate efficient distributions if their participants engage in utility-maximizing behavior (Arrow and Hahn 1971; see also Moriarty 2019).[4] Hence, as Heath makes repeatedly clear (e.g., Heath 2014, 28–29), profit-maximizing behavior is permitted in markets, because it enables markets to produce efficient distributions. Therefore, Heath argues, achieving efficiency requires that economic agents act according to an "adversarial ethic," which demands a weakening of, or, in some cases, elimination of, everyday moral obligations in markets (9–10).

[2] Heath (2014, 69) explicitly characterizes MFA as a normative theory addressing the professional moral duties of managers in a market economy. However, we believe that this characterization arbitrarily limits the potential and reach of the theory. Contrary to Heath's insistence, MFA clearly cannot merely be a theory about the professional duties of managers. Even if managers themselves were never to exploit any market failures, there would be no reason to believe that markets would tend to generate Pareto-efficient distributions if all other market participants exploited market failures. Hence MFA must also be a theory of market participant ethics. In what follows, we will discuss the implications of the MFA as an ethical theory addressing market participants in general, rather than merely managers.

[3] We use the terms *efficient* and *efficiency* to refer to Pareto efficiency, unless specified otherwise.

[4] To be more precise, consumers engage in utility-maximizing behavior, whereas firms engage in profit-maximizing behavior. Heath usually focuses on the case of firms; hence he typically uses the term *profit-seeking* behavior.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

However, real markets are not always perfectly competitive as required by the first fundamental welfare theorem. In Heath's (2014, 34) words, "when one or more of the Pareto conditions [i.e., the conditions necessary for the first fundamental welfare theorem to obtain] are not satisfied, the competitive equilibrium of a market economy will be less than Pareto-optimal. When a Pareto-inferior outcome is realized, this is referred to as a market failure." Market participants (including, but not limited to, managers) can hinder markets from producing efficient distributions by exploiting market failures.[5] For example, if information about the inferior quality of a product is available only to the seller of the product, the buyer might purchase it for the price she would have been willing to pay for the good only at a normal quality. Hence the transaction is not Pareto efficient. In many cases, market participants will have strong incentives to exploit market failures for their own benefit. However, by doing so, they undermine the market's tendency to generate efficient distributions. Heath argues that managers have a moral duty to refrain from taking advantage of or creating market failures. The upshot of his theory of business ethics is that market participants ought to follow heuristic imperatives, such as "minimize negative externalities" or "compete only through price and quality," in order not to disrupt the market's capacity to generate efficient distributions (37).

The criticisms against the MFA encompass five main strands. First, the existence of widespread market imperfections entails that the MFA is de facto not justified by appeals to Pareto efficiency (Moriarty 2019; Hsieh 2017). Second, although the MFA might apply under ideal circumstances, it has no real-world implications (Moriarty 2019; Steinberg 2017). Third, the MFA's imperatives apply to agents other than managers (von Kriegstein 2016). Fourth, there is nothing special about the adversarial character of the market: love relationships, legal disputes, and many other daily activities often have a competitive nature (Hsieh 2017). Fifth, the exclusive focus on efficiency is unsatisfactory for a theory in business ethics that is embedded in a larger theory of political philosophy (Singer 2018).

While we strongly agree with the vast majority of these criticisms, our arguments differ from earlier critiques of Heath's MFA in four substantive points. First, to the best of our knowledge, the current literature does not encompass any criticism of the MFA under the ideal conditions of perfect competition. We believe that to assess the MFA, we must first take seriously its vision for markets under ideal conditions. Second, whereas some commentators have pointed out that the MFA lacks moral force under nonideal circumstances (e.g., Moriarty 2019; Hsieh 2017; Steinberg 2017), our criticism will take seriously Heath's proposal to focus on Pareto-improving, "bottom-up" exchanges (Heath 2014, 40) under nonideal conditions. Third, whereas many commentators have noticed

[5] As we explain in section 2, there is an important distinction to make between two kinds of efficiency: constrained efficiency and social optima. Because this distinction is relevant only under nonideal conditions, we bracket it here and in section 1 (where all efficient distributions constitute social optima) for illustrative purposes.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

that actual markets do not fulfill the conditions of perfectly competitive markets (Singer 2018; Moriarty 2019; Hsieh 2017), we emphasize in various passages the crucial point that market participants in nonideal markets cannot live up to the implicit epistemic demands of the MFA. Fourth, another focal point in our critique that has been underappreciated in the literature thus far is the interrelation between state regulation and moral constraints. As we will point out, depending on the source of market failure, neither state regulation nor ethical constraints will help remedy the failure. In conclusion, we believe that our novel criticism of the MFA adds to the existing literature in a productive manner and hope that this article helps to improve on a promising, well-established theory in business ethics.

In this article, we roughly follow Heath's argumentative structure by which he establishes the MFA. In section 1, we discuss the MFA as an ideal theory, that is, under conditions of perfect competition. We argue that, even if conceived of as an ideal theory, the MFA allows for severely harmful transactions to take place. In section 2, we discuss the MFA as a nonideal theory and argue that Heath faces two options, neither of which is satisfactory. First, when market failures prevent the economy from reaching Pareto-efficient distributions, Heath may insist that market agents have a duty to refrain from exploiting market failures to the best of their ability by striking only Pareto-efficient bargains. In that case, we argue that the MFA is too demanding because it imposes a significant epistemic burden onto market participants, who need to act as if markets were competitive, but without perfect information. Such a duty would be overdemanding in principle and therefore implausible. Second, Heath might argue (as he explicitly does) that when the ideal of Pareto efficiency is too remote to be achievable, market participants must engage in mutually beneficial interactions instead. We argue that, in that case, the MFA is not demanding enough, because the focus on mutually beneficial transactions sets a much lower moral standard for market participants than Pareto efficiency. The third section focuses on Heath's justification for state regulations. Heath argues that the state's role is to correct or prevent market failures through establishing an appropriate legal framework. He then claims that, when state intervention fails to minimize market failures, market participants are required to abide by the MFA's specific ethical constraints. In response, we argue that Heath's recourse to regulations and ethics is unsatisfactory. We show that it is unclear how the MFA's heuristics can help correct market failures when regulations fail to do so and argue that this threatens their moral force. Finally, we offer a schematic, constructive suggestion on how the MFA could be adjusted to avoid its most central problem: the overdemandingness of a justification based on efficiency. We suggest that a less demanding yet morally more defensible version of the MFA requires market participants to engage in mutually beneficial transactions while respecting two moral side constraints: first, the requirement not to discriminate against other market participants on racist, sexist, ableist, ageist, and so on grounds, and second, the requirement not

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

to exploit those market failures that foreseeably generate significant morally relevant harm.

1. THE MFA AS AN IDEAL THEORY

The MFA derives much of its justificatory core from the market's capacity to generate efficient distributions. However, as Heath explains, actual markets are riddled with market failures and will often fail to achieve efficiency. While actual markets might coincidentally produce efficient distributions, only perfectly competitive markets necessarily do so. Nevertheless, it is useful to start our inquiry by assuming perfectly competitive markets and no market failures. Indeed, Heath (2014, 30) himself first sketches out the MFA in a perfectly competitive world. In this section, we argue contra Heath (2014, 174) that, even in this simplified scenario, the MFA's moral imperatives are explicitly not "pretty much all there is to business ethics, at least with respect to market transactions." Even assuming full compliance, the MFA permits voluntary transactions that might impose devastating harm on third parties.

The ideal scenario that forms the background to our discussion is the world of the first fundamental welfare theorem. As we have already stated, the theorem shows that perfectly competitive markets generate efficient distributions. To see how far removed this world is from ours, it pays to look into the concept of a perfectly competitive market. A market is perfectly competitive if it satisfies what Heath (2014, 34) refers to as the Pareto conditions. Hsieh (2017) and Moriarty (2019), among many others, have already noted that these conditions are excessively demanding. They include such items as complete and transitive preference rankings of market participants, complete markets, no transaction costs, no market power, and, most importantly for our purposes here, no (nonpecuniary) externalities.

Yet, abstracting from the complications of real markets permits us to discuss the simplest case of an application of the MFA. In a perfectly competitive market, no market failures emerge, and therefore there is no opportunity to exploit them. Hence the moral imperatives of the MFA are equivalent in their normative, action-guiding content to the dictates of rational utility maximization ("Buy x, sell y given your preferences—don't worry about market failures; there are none").[6]

A particularly important aspect of competitive markets is that they do not exhibit a specific kind of externality, so-called nonpecuniary externality. The textbook case describes a company engaging in pollution without compensating the victims of the pollution. What is special about nonpecuniary externalities is that they arise "outside" of the market, or as a side effect of the market process, when someone's utility or profit is decreased because a cost, such as the effects of pollution, has not been taken into account in the production process.

[6] Gauthier (1986) hence refers to competitive markets as "moral free zones." Notice, however, that even if rationality and MFA dictate the same actions, the normative reasons for acting according to MFA and rationality can differ even in a perfectly competitive market.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

Contrary to this, perfectly competitive markets necessarily exhibit pecuniary externalities. Pecuniary externalities arise "inside the market" when a market participant's utility or profit is diminished as a result of a change in prices. This should come as no surprise: if the price of good x falls in a perfectly competitive market, buyers of x are made better off, while sellers of x are made worse off. Nonetheless, if the market is perfectly competitive, the resulting distribution will be Pareto efficient. If the market is perfectly competitive, losses and gains will be balanced such that the social surplus is maximized, that is, efficient.[7]

As numerous authors have argued (Hausman 1992; Olsaretti 2004; Sen 1985), though they do not affect the efficiency of a distribution, pecuniary externalities are a cause for concern from a perspective of justice. This concern can be spelled out in two different ways. First, pecuniary externalities can cause severe inequalities: even if the end result of the market process is efficient, some people will be made better off or accumulate more resources than others (Sen 1985). Although we share this concern, we will bracket it here. Second, pecuniary externalities can amount to severe harm impositions (Hausman 1992; Endörfer 2021).

To see this, imagine a perfectly competitive market in which the price of red apples falls to the point of driving some producers out of the market, because red apples become unfashionable as a new species of blue apples is introduced. Consequently, the red apple producer Ann, who has no savings, loses her only source of income and starves. But her competitor, the blue apple producer Bob, who has accumulated enough savings to withstand a crisis, will benefit from her elimination from the market.[8] For illustration, consider Figure 1.

Both x and y show a specific distribution of indiscriminate units of utility between Ann and Bob. Initially, the market prices are such that they produce at point x. Then, owing to some change in the demand for red apples, the price of red apples decreases, and they move to point y. Remember that we assume perfectly competitive markets. Price changes are not under the control of Ann and Bob, who have no choice but to move from x to y. The shift from x to y thus marks the shift in the distribution of utility between Ann and Bob when the price of red apples decreases. For the sake of simplicity, we set the final price at 0, but our argument does not rely on this assumption. The difference between b_1 and b_2 shows Bob's increase in utility, while the difference between a_1 and zero shows Ann's loss. Both x and y are on the Pareto frontier, nonetheless. The graph illustrates that Ann suffers because of a shift from one Pareto-efficient distribution (x) to another (y).[9] This is important, because it demonstrates that Ann's starvation is explicitly not a result of a market failure. Ann is harmed because of her exposure to market competition.

[7] For a discussion of why pecuniary externalities do not affect efficiency, see Scitovsky (1954).

[8] To render the story consistent, we should add that blue apples are much more expensive than red apples, such that consumers are indifferent between buying one blue and ten red apples. But they simply choose only to buy blue apples.

[9] One might think that the first fundamental welfare theorem shows that a competitive market achieves a unique equilibrium, such that a shift from x to y is inconceivable. But for the first fundamental welfare theorem to prove a unique equilibrium, we have to make additional assumptions about consumer preferences. See Arrow and Hahn (1971).

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

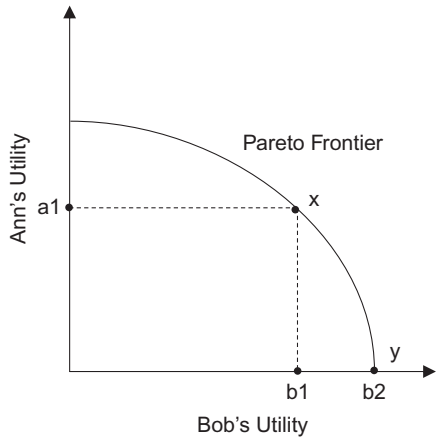

Figure 1: Ann and Bob's Pareto Frontier

Heath agrees that competition, within the market or otherwise, does indeed impose harms on the losers of competition. Nonetheless, he insists "that the wrongness of this harm is outweighed by the positive externalities generated by the competition as a whole" (Heath 2014, 103). He argues that we should permit market participants to expose each other to harm via changes in prices, because, in a perfectly competitive market, the price mechanism will lead to Pareto-efficient distributions of goods and services. But notice that in our example, there is no possibility to generate a Pareto improvement once *x* or *y* has been reached. By assumption, both x and y are efficient distributions. The move from *x* to *y* has therefore not generated any additional "positive externalities"—consumers do not gain any additional utility. Ann's starvation thus cannot be justified with reference to efficiency-related reasons. Ann starves because of a win–lose transformation consistent with efficiency.

If the MFA permits market participants to buy and sell as they please in a perfectly competitive market, they are permitted to do so even if they thereby significantly harm competitors like Ann. We thus conclude that even under ideal conditions, the MFA is not demanding enough, because it does not prevent market participants from imposing severe harms on third parties via pecuniary externalities. Contrary to this, we submit that if voluntary market transactions cause third parties severe harm, we have at least a strong pro tanto reason to consider these transactions morally impermissible. The MFA acknowledges no such reason when efficiency concerns remain unchanged.[10]

Heath (2014, 183) points out one way in which this conclusion can be resisted via the second fundamental welfare theorem (see also Singer 2018):

[10] Even if Heath does not discuss them, other, efficiency-independent defenses are available to argue that exposing others to harm via pecuniary externalities is permissible. For a full discussion of pecuniary externalities and market harms, see Endörfer (2021) and Wilkinson (2022).

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

> The second fundamental theorem of welfare economics showed that, with the proper allocation of initial assets, a competitive market could generate any one of the sets of possible Pareto-optimal distributions. This means that whatever distribution would be picked out by one's favored conception of equality, there is no obstacle in principle to the design of a market that would generate exactly that outcome.

Crucially, this implies that we could avoid a shift to a state of affairs in which Ann ends up severely harmed if assets were simply redistributed via a lump sum tax (Heath 2014, 183). This response is only partially satisfying on Heath's own line of reasoning.

Remember that the central reason for why Heath considers markets important is their capacity to generate win–win transformations. Heath's hope is that perfectly competitive markets are justified because they make the metaphorical cake as big as it possibly could be. But there are problems with how we are to understand the notion of win–win transformations in light of possible redistribution efforts a la the second fundamental welfare theorem. If lump-sum redistribution is compatible with Heath's minimally controversial contractualism, it is unclear whether the justification of markets is, all things considered, as uncontroversial as he claims. Any redistributive effort in the name of equality or aversion to harm necessarily generates win–lose transformations. Heath (2014, 2) admits as much himself: "equality arguments [i.e., arguments that do not appeal to efficiency] are essentially about how to resolve distributive conflict, and so always have a win–lose structure." So we might ask ourselves, even if the market itself exclusively generates win–win transformations, how much is achieved if these transformations are ultimately undone by redistributive efforts anyway? Whether markets make the cake as big as it could be matters little to Bob if he has to anticipate that whatever efficiency gains he enjoys from his participation (and perhaps even more) will ultimately be taxed away to preserve some form of justifiable equality. Heath's hope must be that (as he states in connection to his discussion of Cohen's camping trip) "the efficiency gains might be so great, in other words, that the campers are willing to surrender control over the distributive consequences" (185). But they might not be. As a result, Bob might have to give up more than he gained to prevent Ann's starvation. In other words, if some form of redistribution is part of Heath's minimally controversial contractualism (as one would presume necessary to preserve its plausibility), nothing necessarily precludes that the promise of win–win transformations remains but a shallow one.

This, in turn, brings us back to the MFA. To sum up, even under idealized conditions, either the MFA is a normative theory that stands on its own but considers severely harmful transactions morally permissible or the MFA is part of a wider vision of minimally controversial contractualism. In the latter case, the MFA is only plausible given lump-sum redistributive efforts that prevent severe harm via market transactions. But if the latter is the case, Heath has to demonstrate, rather than merely stipulate, that the efficiency gains are sufficiently large to constitute posttaxation win–win transformations while protecting market participants from severe harm. Therefore, at worst, the MFA as a stand-alone

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

normative theory is implausible. At best, as part of a wider contractualist theory, it is incomplete.

2. THE MFA AS NONIDEAL THEORY

Heath discusses the ideal scenario to demonstrate that for markets to reliably generate efficient distributions, market participants must refrain from exploiting market failures. Nonetheless, Heath agrees that actual markets are not perfectly competitive. In his own words, "if the benchmark [for the success of markets in correcting collective action problems] is Pareto-optimality, then markets always fail" (Heath 2011, 25). Despite this, Heath believes that the MFA can provide us with a number of reliable "heuristics" that should morally guide market participants' transaction decisions (41). Although Heath insists that these heuristic rules can be highly demanding, he also writes that "it cannot be argued that these demands are too onerous in principle, since the demands simply articulate the way that capitalist economies are supposed to function in the first place" (38).

In this section, we argue that the MFA as a nonideal ethical theory is either too demanding or not enough, depending on one's objectives. If Heath wants to safeguard Pareto efficiency, then the demands placed on market participants by the MFA are indeed "too onerous in principle," because they impose excessive epistemic burdens onto market participants. However, if Heath relinquishes Pareto efficiency and embraces instead a bottom-up perspective that merely requires transactions to constitute merely mutually beneficial transactions, we argue that the MFA is then not demanding enough.

To do so, we must first clarify an important conceptual point concerning the difference between two different kinds of Pareto-efficient distributions. We refer to these different kinds as *social optima* and *constrained efficient distributions.* Social optima are Pareto-efficient distributions that result under conditions of perfectly competitive markets. These are the distributions that are generated when the assumptions of the first fundamental welfare theorem hold. As Heath readily admits, these assumptions do not hold in real markets. But nonetheless, real markets also produce distributions that are "efficient," albeit in a much less inspiring way. We call these distributions constrained efficient distributions. To illustrate, imagine that I found your lost bike. The distribution we have now reached could easily be Pareto improved upon, because you value your bike more than I do. You could theoretically offer me fifty dollars for the bike, which I value at forty dollars, and we would both be better off. However, I do not know whose bike I found. This lack of information is a so-called transaction cost.[11] Because I do not know that you would be willing to pay me more for the bike than it is worth to me, we cannot engage in transaction. In other words, no Pareto improvement is available. The distribution resulting from me

[11] For further discussion of these issues, we refer the reader to the seminal articles by Coase (1960) and Dahlman (1979), as well as to a particularly illuminating article by Calabresi (1991) on the "pointlessness" of Pareto efficiency.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

finding your bike is hence "constrained efficient."[12] When firms maximize their profit, but transaction costs are rampant (such that they are not aware of potential Pareto-improving exchange opportunities), whatever distribution results from their actual exchanges will be Pareto efficient—but only in the constrained efficient sense. We therefore assume that when Heath refers to the failure of actual markets to achieve Pareto efficiency, he means that markets are unable to achieve a social optimum.

With this caveat out of the way, let us begin with the overdemandingness of the MFA. Heath himself presents a first reason for which the MFA may be overly demanding. In many cases, market participants will forgo opportunities for profit if they do not exploit market failures, which can be severely costly to them (Heath 2014, 37). However, the MFA is also overdemanding in another sense. Even in a "deontically perfect world" (Heath 2014, 36), in which market participants in imperfectly competitive markets fully withstand the incentive to exploit market failures, they ought to recognize a market failure when they see one and know how to avoid it to collectively generate a social optimum, rather than a constrained efficient distribution. This is not as easy as it might seem at first glance, and Heath's appeal to heuristics might not help much in this regard. To be fair, all other things equal, some heuristics might be simple to follow, such as "do not seek tariffs or other protectionist measures" (Heath 2014, 37). But others, such as "minimize negative [nonpecuniary] externalities" (Heath 2014, 37), are overdemanding, because they place an excessive epistemic burden onto market participants. Consider a typical pollution case in which a firm is able to reduce its production costs for a particular good simply by dumping toxic by-products into a nearby river. Thereby, it is affecting the health of its customers, who live downstream of the river. Prima facie, it seems as if the social optimum (i.e., the efficient distribution) would amount to the absence of pollution. But this is not true. The social optimum could just as well be some nonzero level of pollution that enables the polluting firm to provide the good at a lower cost to its polluted customers. In order not to threaten the social optimum, the firm (qua agent, in this example) would have to know the precise degree to which the river must be polluted and the exact price at which pollution pays off for its customers. This obviously implies that the firm must have some form of perfect information about the pollution preferences of its customers, which (assuming, for example, positive transaction costs and incomplete preference sets of the customers) is a tall order, to say the least.

The core problem, however, extends beyond the issue of oversimplified heuristics. Another way to frame the problem of the pollution case has to do with market-clearing prices. Generally speaking, the price of some good x is market clearing when it is such that the demand for x precisely equals its supply. If prices are not market clearing, Pareto improvements are possible. Pareto improvements are in such

[12] Dahlman (1979) conversely refers to externalities that permit for the emergence of what we call "constrained efficient" distributions as "not Pareto relevant." He prefers this terminology to demonstrate that the idea of a social optimum is normatively infused, but for our purposes here, it is better to cling to the idea of the social optimum.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

cases possible, because some excess of supply could have been sold if the price were lower, making both seller and buyer better off. Alternatively to providing its products at a cheaper price, the firm could simply offer to bribe those who suffer from its pollution, that is, propose a price for its pollution.[13] The social optimum might be a mixture between the bribe and some nonzero level of pollution, such that the "market for pollution" clears. But, depending on which conditions of perfect competition are violated, it is unclear whether the market-clearing price can even be determined. Transaction costs of various kinds might be in the way of efficient bargaining, property rights over the river might not be clearly defined, and so on. The end of the story is that the bribe will likely be too high or too low, and even if the resulting distribution will qualify as a constrained efficient, it will not be a social optimum.[14] Hence, if the MFA permits market participants in an imperfectly competitive market to engage only in transactions that promote a social optimum, it requires them to *know* and exclusively exchange at market-clearing prices. But this is not possible, because only a perfectly competitive market ultimately reveals these prices. The MFA hence demands from market participants to know and act on what they cannot know in an imperfectly competitive market.[15]

This brings us to our next point: the lack of demandingness of only Pareto-improving transactions. In a Hayekian turn,[16] Heath (2014, 40) proposes to consider the functioning of the market from a bottom-up perspective:

> In day-to-day life, this optimum [i.e., a social optimum] is irrelevant. Every voluntary exchange generates a Pareto improvement. It is through these tangible, incremental efficiency gains that the private market system has established its merit.[17]

[13] What we discuss here is one implication of the so-called Coase theorem. For the original article, see Coase (1960).

[14] The end result is that government intervention might bring about the social optimum, but we will say more on that in the next section.

[15] Heath (2014, 181) indicates a possible solution to the problem of overdemandingness, the so-called n-best framework. Heath states that the simple MFA heuristics might need to be adapted to account for potential "significant compliance problems" (180), which presumably include the kinds of epistemic constraints we have pointed out in this section. In this way, normative requirements of the MFA framework can be weakened to account for complications under nonideal conditions. Heath's move from social optima–promoting to bottom-up Pareto-improving exchanges (which we discuss in the next paragraph) could be understood as an application of this weakening of the normative framework. The problem with the suggested response is that Heath promotes the weakening of the normative framework of the MFA as an "iterative" process (180): instead of second- and third-best frameworks, the MFA operates with n-best frameworks. However, Heath tells us very little about which "compliance problems" legitimately justify a shift from a j- to a k-best framework. If the weakening is iterated sufficiently often, the danger is hence that the MFA becomes an uninformative normative framework. The worry is that an i-best framework undermines the entire point of the MFA by simply permitting market participants to exploit market failures and conduct business as usual, because they are unwilling to comply with any other normative framework. Because Heath is, to the best of our knowledge, unspecific about under which conditions he believes a weakening of the normative MFA framework to be justified, we will not pursue this counterargument any further here.

[16] See von Hayek (1945) for his vision of how prices emerge in free markets.

[17] Quite confusingly, in another passage, Heath (2014, 198) states, "The market is designed to promote Pareto efficiency as a byproduct of competitive behavior on the part of firms, *no single instance of which will be Pareto-improving*" (italics added). This is correct. Because of the presence of pecuniary externalities,

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

In other words, Heath argues ultimately that prices need not be market clearing and that the social optimum can be ignored. We can achieve something close enough to a social optimum by relying on the fact that even imperfectly rational market participants will engage only in mutually beneficial exchanges.

This leaves us with the following rationale. The MFA as a nonideal theory dictates that market participants face only two moral obligations when transacting in markets: first, the obligation not to exploit market failures (insofar as they can know about them) and, second, the obligation to engage only in exchanges that are mutually beneficial (Heath 2014, 11).[18] Let us call markets in which full compliance with these two obligations prevails *MFA-constrained markets.*

The important point to notice is that MFA-constrained markets are not perfectly competitive markets and hence do not generate social optima. The well-intentioned market participants nonetheless face transaction costs and are subject to bounded rationality, nonobserved externalities, and so on. The distributions they collectively generate might constitute Pareto improvements but are not social optima. We think that MFA-constrained markets are problematic for two reasons.

First, market participants can exploit market failures and nonetheless be mutually beneficial. A monopolist can sell her good x at a price that is higher than the market-clearing price and thereby not saturate demand; that is, the monopolist could sell above the marginal cost of production of x. Nonetheless, the customers who buy the product are made better off than they would be if they could not purchase x. This is a mutually beneficial transaction, even though a market failure has been exploited. Such examples show that if all that ultimately matters are mutually beneficial transactions that culminate in "incremental efficiency gains," a prohibition on the exploitation of market failures seems unmotivated on the grounds of the MFA. But Heath needs to provide arguments that demonstrate that MFA-constrained markets are Pareto superior to possible alternatives, in which market failures are exploited. In light of such arguments, a prohibition on the exploitation of market failures could be justified based on the MFA. But as Moriarty (2019) argues, we may doubt that such arguments are available according to our best economic theories.[19]

Second, even if Heath can justify the prohibition of failure-exploiting behavior, any market distribution that falls short of a social optimum fails to adhere to what Heath considers the most central reason to prefer MFA-constrained markets over their alternatives: that markets generate win–win transformations.

almost no exchange in the market will be a Pareto improvement (unless compensated via offsetting price changes). For discussion, see Holcombe and Sobel (2001). Later in this section, we make a similar argument by stating that Heath is implicitly promoting, not win–win, but win–lose transformations. To preserve precision, we hence state that in MFA-constrained markets, each market participant engages in mutually beneficial transactions, rather than in strictly Pareto-improving transactions.

[18] More precisely, Heath (2014, 187) does not expect the actual distribution to be even Pareto efficient but merely "by-and-large Pareto improving."

[19] This is so because of the "general theory of the second best" (Lipsey and Lancaster 1956). We bracket the discussion, as it has already been addressed at length by Heath (cf. chapter 6 in Heath 2014; Heath 2019) and his critics (Moriarty 2019).

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

At first glance, this may seem counterintuitive: how is it possible that some market participants suffer losses if all transactions in an MFA-constrained market are mutually beneficial? The reason is that even a transaction that is *in isolation* mutually beneficial to buyer and seller might create adverse third-party effects (also referred to as *pecuniary externalities*). For example, if the price of x is sufficiently responsive to the demand of x, then, all things equal, the mutually beneficial transaction of many units of x between two parties will increase the price of x. Unfortunately, an increase in the price of x makes prospective future buyers worse off, because they will have to find an adequate substitute for x or accept a higher price for x. We have argued in the previous section of this article that such third-party effects can theoretically, even in a perfectly competitive market, amount to severe harm.[20] Clearly, transactions that generate such harm are not win–win transformations but win–win–lose. Furthermore, nothing precludes the possibility that such adverse third-party effects emerge in imperfectly competitive markets, including MFA-constrained markets. Notice that such third-party effects do not necessarily emerge because of ill will or otherwise flawed moral motivation of market participants.[21] They can simply be the result of profit (or utility) maximization under nonideal conditions.

To sum up, contrary to what Heath argues, under nonideal conditions, full compliance with the MFA will effectively generate win–lose transformations, either because market agents cannot know market-clearing prices or because they engage only in mutually beneficial transactions, which may have deleterious effects on third parties. But if MFA-constrained markets, not unlike real markets, by assumption generate win–lose transformations, it is unclear what could convince us to promote the MFA.

We thus conclude that, under nonideal conditions, the MFA is either overly demanding or not demanding enough. It is overly demanding if invoked as a way to salvage social optima despite market imperfections. For, in that case, it requires market participants to know which prices are market clearing and which actions correspond to the exploitation of market failures. If, however, the goal of achieving social optima is effectively abandoned, the MFA is not demanding enough. For market participants are then permitted (or even required) to engage in transactions that only in isolation constitute win–win transformations but could imply catastrophic third-party effects. In short, there is nothing inherent to the MFA that guarantees that nonideal markets generate win–win transformations.

3. ETHICAL CONSTRAINTS OR STATE INTERVENTION?

Even if it is true that nonideal markets on their own do not deliver on the promise of win–win transformations, there is still the hope that the MFA, complemented

[20] See Hausman (1992, 99) for a related discussion on third-party effects in mutually beneficial transactions.

[21] For examples of similar results in economics, see Greenwald and Stiglitz (1986) and Loong and Zeckhauser (1982).

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

with adequate state interventions, can correct their shortcomings. Heath (2014, 11) believes that state interventions in combination with adequate ethical constraints might correct the "movement of prices in the direction that will clear the market." In this section, we argue that Heath's appeal to this twin solution of regulation and moral constraints creates severe problems for the relevance of the MFA.

As we have seen, Heath acknowledges the failure of actual markets to reach efficiency. If markets fail to produce efficient distributions, he suggests a twofold solution. First, he appeals to state interventions aimed at correcting market failures. Second, whenever the state is not successful in remedying market failures, market participants ought to follow the MFA's heuristics (or ethical constraints) and refrain from exploiting existing market failures. In summary, Heath's conclusion is to defend a mixed economy in which market participants strive to achieve Pareto improvements while the state's role is to minimize market failures when they occur. The heuristic rules resulting from the MFA are hence required as a complement to state regulations when the latter fail to achieve efficiency, and vice versa. Let us examine Heath's arguments more closely.

First, consider Heath's arguments in favor of state interventions into markets. He writes that "the primary function of the legal regulation of the market is to prevent market failures" so as to achieve "perfect markets" (Heath 2014, 34; see also Heath 2011). According to Heath (2014, 11), one of the primary mechanisms by which the state can minimize market failures is through the enforcement of property rights. He shares the view of neoinstitutional economists, such as Coase (1960), that an adequate system of property rights in combination with a low-transaction-cost environment can, in theory at least, allow private actors to internalize externalities. Other existing legal ways to reduce market failures are the tort system; the provision of public goods, such as roads; and regulations (broadly conceived by Heath as comprising all sorts of restrictions on pollution, waste, etc.) (Heath 2014, 11, 88–89).

But Heath (2014) acknowledges that "perfect" regulation is as much a fiction as "perfectly competitive" markets. He concedes that "the law is a somewhat blunt instrument" (89). Very often, even the state lacks the necessary information to implement optimal measures. Moreover, the administrative costs of regulating the economy may be too high and impose an excessive deadweight loss on the economy, which more than outweighs the potential efficiency gains of reducing market failures. In these various senses, then, state intervention often fails to correct market failures and secure efficiency.

But when regulation fails to bring about efficiency, the "ethical constraints" on profit maximization set by the MFA enter the picture (Heath 2014, 89). More specifically, Heath argues that firms ought to abide by the heuristic rules established by the MFA, which delineate the appropriate rules of competition: "minimize negative externalities," "reduce information asymmetries between firm and customers," and so on (for additional examples, see Heath 2014, 37, 111–13). If the state cannot enforce the rules of competition and minimize the incidence of market failures, because it would be too costly or too complex to do so, market

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

participants "should behave as though market conditions were perfectly competitive, even though they may not in fact be" (Heath 2014, 37; see also 111).

Thus Heath (2014) requires market participants to behave as profit maximizers, which are encouraged to take advantage of free-riding strategies (9), while at the same time acting in accordance with the MFA's heuristics, or, as Heath puts it, the rules of "sportsmanship" (103) as established by the ideal of the perfectly competitive markets. The combined result is that they ought to abide by several "moral obligations that extend beyond their legal obligations" (Heath 2014, 12; see also Heath 2014, 89–90).

Heath acknowledges that the demandingness of the MFA stands in conflict with profit maximization. As he states, "in the real world, any firm that began to unilaterally respect these [moral] constraints would be quickly eliminated from the marketplace" (Heath 2014, 37, 112; 2018). In short, firms are caught in a conflict between the morality of competition (as conceived by Heath) and the harsh reality of competition. Heath (2014, 37) nevertheless argues that "the fact that other people are not going to respect their moral obligations does not undo the obligation for everyone else." For Heath (2018), failing to abide by these moral norms may be excusable under specific circumstances, but it is still a serious moral wrong. He concludes his argument by making clear that "the imperatives outlined above are extremely demanding, so much so that competitive pressures would probably prevent any corporation from respecting all of them in the near term" (Heath 2014, 113). Even when it is not in the firm's interest to abide by the MFA's heuristics, it is nevertheless a firm's moral duty to do so to promote efficient distributions.

We believe that Heath may be overly optimistic about the level of successful compliance with the MFA. Even taking into account Heath's own considerations, the obvious epistemic problem remains: as we have already argued, in incomplete markets, it is simply not clear in many cases why firms themselves should be particularly well placed to assess whether they are exploiting market failures. As stated in the previous section, one particularly poignant problem is that the relevant information to assess whether a market failure is present is partially created via the price mechanism itself, which, in an incomplete market, is severely distorted. Our main point of contention, though, is that Heath's strategy to minimize market failures through a twin appeal to regulation and ethical constraints reveals a much more severe problem of the MFA.

Recall that a market failure is a situation in which "the competitive market fails to produce a Pareto-efficient outcome" (Heath 2014, 87). According to Heath, regulation is needed to correct market failures and enable social optima. However, Heath does not say much on why market failures may arise. Understanding their origin, however, is of crucial importance. Schematically, we can say that they come into existence for three kinds of reasons.

First, market failures can arise when market participants are unwilling to abide by the rules of competition. This includes cases in which market participants are fully aware that they create or exploit market failures. Call the failures that arise out of a lack of appropriate motivation to abide by the MFA's heuristics *moral failures.* If moral failures arise, optimal regulations may incentivize market participants to

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

behave in accordance with the MFA's heuristics and thereby safeguard a social optimum.

Second, market failures can arise even if market participants are fully compliant with the MFA's heuristics. This includes cases in which market participants are willing to compete fairly but are unable to acquire adequate information on the market in order to do so. For example, if property rights are indeterminate or information unevenly distributed among market participants, they cannot know the market-clearing prices and thus cannot collectively generate an efficient distribution. Call these market failures *epistemic failures.* If epistemic failures arise, regulations may ideally enable market participants to act in accordance with the rules of competition, for example, by correcting the distorted price mechanism via optimal taxation, and thereby help market participants to collectively secure efficiency.

Finally, market failures can arise for a third reason, which we shall call *moral ignorance.* Suppose that economic agents have perfect information, that they are motivated to act morally, and that regulations are fine-tuned to the MFA such that if general compliance with the MFA is achieved, social optima will result. There is thus no epistemic or moral failure, at least not in the sense that market participants are not sufficiently motivated to act morally. It is merely the case that they do not know about the MFA heuristics and thus fail to act accordingly. If they knew the heuristics, they would immediately act in accordance, and no market failures would arise.

Regulations might, at first glance, help in the former two cases. First, regulation can incentivize market participants to act in accordance with the MFA's heuristics if they lack motivation to do so. Second, regulation can provide information when market participants suffer from epistemic shortcomings. In mere moral ignorance cases, providing information about the MFA on its own is sufficient to generate Pareto-efficient distributions, even if regulation is otherwise deficient. We return to these cases later in this section.[22] Heath's hope is that whenever regulation fails, the MFA's heuristics will safeguard social optima, and vice versa. However, we shall argue that Heath's twin solution of regulation and ethical constraints is unhelpful in a number of relevant cases, even if we assume that nonideal markets could in principle generate efficient distributions.

Let us begin with a case of epistemic market failure. Consider an MFA-constrained market, that is, a market in which everyone fully complies with the MFA's heuristics but which nonetheless does not generate efficient distributions because market-clearing prices are not known. Regulatory efforts will probably not succeed in revealing these prices (in part because the market itself is supposed to reveal market-clearing prices). In such a situation, as has been argued in the previous section, it is unclear how full compliance with the MFA's heuristics could be of any help. Market participants do not even know whether and where inefficiencies are present. We have already assumed full compliance with the MFA's heuristics.

[22] We assume here that it is not of any relevance whether information about the MFA in mere moral ignorance cases is distributed via the state or other parties.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

Therefore, if regulations fail, a social optimum will not be attained. Nonideal markets of this kind are damned to stay inefficient.

But what about moral failures cases? Consider a situation in which, for some reason, market-clearing prices are common knowledge and if each market participant bought and sold at these prices, a social optimum would result. However, some monopolists are unwilling to sell at these prices. They are simply in principle not motivated by any appeals to do what is morally right according to the MFA. By assumption, regulations fail to rein in these monopolists. Even knowing full well that a social optimum could be obtained if they sold at market-clearing prices, they will not act in accordance with the MFA's heuristics, and regulations will not incentivize them to do so. Markets with such unresponsive monopolists are similarly doomed to inefficiency as markets with fully compliant participants that face epistemic market failures. In conclusion, the twin solution of regulation and MFA's heuristics will not safeguard efficiency, at least in these cases.

This brings us to moral ignorance cases. In these cases, efficiency is by assumption achieved if market participants merely acquire knowledge about the MFA heuristics. These are the cases in which MFA delivers on its promise of bringing about a social optimum. But if the MFA heuristics address merely moral ignorance cases, we should seriously question how helpful these heuristics are. Actual markets suffer from far more shortcomings than the mere fact that market participants unknowingly exploit market failures and that actual regulations are not fine-tuned to the MFA; we have already elaborated on many of these shortcomings earlier. If the MFA heuristics only have moral force in idealized moral ignorance cases, it remains unclear why they ought to be followed by actual market participants. As long as additional moral and epistemic failures plague markets, social optima cannot be safeguarded.

At first glance, the fact that, in some cases, the twin solution of MFA heuristics and optimal regulation will not reliably generate social optima seems to be a minor problem for Heath. Irrespective of whether market participants abide by the MFA's heuristics in the absence of perfect regulation, the heuristics' moral force remains intact. Not exploiting market failures is simply the right thing to do. An analogy from utilitarianism might be helpful to understand the response: even if nobody strives for bringing about the best consequences (and hence the best consequences are not achieved), acting such as to bring about the best consequences is still morally obligatory.

On closer inspection, this response may not be available to Heath. First, in the case of moral market failures at least, this reply appears to rely on a rather peculiar conception of the moral psychology of market participants. On one hand, it is assumed that, under nonideal circumstances, market failures will arise because market participants refuse to abide by the MFA's heuristics. On the other hand, Heath argues that when the state fails to correct these moral failures, market participants should then abide by the MFA's heuristics and strive for efficiency. If it is assumed that market failures exist because of moral failures, one cannot appeal to ethical constraints for safeguarding efficiency, for the existence of these market

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

failures rests on the assumption that market participants will not abide by the MFA's heuristics in the first place.

Second, and most importantly, recall that the moral force of the MFA's heuristics is conditional on whether markets can, at least in principle, generate social optima, because only social optima correspond to win–win transformations. Pareto efficiency in the sense of a social optimum is the very reason why the MFA's heuristics strive for perfectly competitive markets, instead of settling for actual markets that generate constrained efficient distributions in the presence of unavoidable market failures. But our examples from above show that in most cases, even with Heath's twin solution, the achievement of social optima is ruled out. If circumstances that prevent markets from generating win–win transformations prevail, the MFA's heuristics lose their moral force. Again, it might be helpful to consider the analogy from utilitarianism: we know that, theoretically, φ-ing would bring about the best consequences in context *c* if events *x*, *y*, and *z* do not occur. We also know that *x*, *y*, and *z* do occur. Therefore, in circumstances *c*, φ-ing does not bring about the best consequences, and the utilitarian will not acknowledge the obligation to φ in *c*. Similarly, there can be circumstances in which it is de facto impossible for market participants to bring about efficient distributions even if they fully comply with the MFA's heuristics. It follows that, in cases like the preceding examples, the heuristics carry no moral force. To be quite clear, if Heath is indeed a "Paretian," as he claims, he is an odd kind of Paretian, because even if fully complied with, the MFA's heuristics will not bring about social optima in the absence of (impossible) optimal regulation. It is hence questionable why a Paretian would insist on their moral force.

To sum up, the problem with Heath's twin solution is that, without optimal regulation in epistemic failure cases, a social optimum will not be reached. But in moral failure cases, optimal regulation renders the MFA's heuristics superfluous. Even under highly charitable assumptions, the MFA's heuristics are in principle only apt to address moral failures when optimal regulation is not possible. And even in these cases, by Heath's own admission, the "benchmark" of a social optimum will not be reached for various reasons, which should make us doubt the MFA's heuristics' moral force and indeed, on appropriate occasions, go against them.

4. SHOULD THE POINT OF MARKETS BE EFFICIENCY?

Thus far, we have argued that the MFA faces significant problems qua ideal as well as nonideal theory of business ethics. In this section, we propose to adapt the MFA in a manner that not only preserves its Paretian motivation but also improves on its applicability as a nonideal theory. Our main suggestion will be to focus on a procedural, rather than an outcome-oriented, Paretian theory of business ethics and to take into account the epistemic constraints faced by market participants. However, owing to constraints of space, what we offer in terms of constructive criticism in this article can at best be understood as a schematic attempt at improving the MFA. As a result, we will not be able to defend some assumptions that are central to our proposal in this article.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

At its heart, the MFA is a contractualist theory, that is, a theory that roughly states that if interactions between agents generate mutual advantage, then such conduct is justifiable to each party and thus permissible (Heath 2014, chapter 7).[23] What renders the MFA a theory of business ethics is that it is applied to the context of a market economy: participants in a market economy (perhaps even first and foremost managers) ought to behave in a way that enables mutual advantage.

We share with Heath a commitment to contractualism and believe, like him, that what makes an act permissible is that it is justifiable to all parties concerned. We also agree that, if market transactions are mutually advantageous, they are (pro tanto) justifiable and thus (pro tanto) permissible. The problem with the MFA's vision, as we have explained in this article, is that its justifiability relies excessively on the notion of Pareto efficiency or, more explicitly, social optima. As we have seen in sections 1 and 2, social optima are guaranteed only in the far-removed world of the first fundamental welfare theorem (Arrow and Hahn 1971; Sen 1985). Despite their best efforts, market participants will not be able to bring about social optima.

However, even if the ideal of Pareto efficiency is probably unreachable, we contend that it is a minimal requirement of any theory of business ethics to demand that market participants engage in transactions that are mutually beneficial in the sense of being Pareto improving in isolation. Yet, by relying on a contractualist framework, we will also argue that market participants face at least two additional moral constraints.

As Heath argues, market economies excel at generating mutually advantageous transactions. If market participants are free to exchange goods with each other as they like and attempt to maximize profit (or utility), it follows that only mutually advantageous transactions will take place. As we have argued earlier, if considered in isolation, each of these exchanges will constitute a Pareto improvement precisely because the exchanges are mutually advantageous.

There is justificatory value in requiring market participants to engage only in transactions that they, from their epistemic standpoint, assess as mutually beneficial —not to do so would clearly generate deadweight losses. If you are willing to consistently trade with me at a loss to yourself, we are de facto intentionally generating a distributional pattern that is worse for you but better for me. The resulting distribution is quite straightforwardly less desirable than a distribution in which we are both made better off. These are deadweight losses that can be avoided even in real economies.

However, once third-party effects enter the picture, it becomes clear why economists so often insist that proper Pareto improvements are few and far between in real markets. A normative theory that would request of market participants to engage only in Pareto improving, rather than exchanges that are merely mutually beneficial in isolation, would be overdemanding. It would require market participants to keep track of how their transactions affect the overall achievement of a social optimum.

[23] Much of what follows is inspired by the Scanlonian understanding of contractualism, as developed in Scanlon (1998). Most centrally, we assume here that justifiability is owed not to aggregates of persons but to individual persons.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

In sum, requiring market participants to engage in transactions that are Pareto improving in isolation is an important element of any theory of business ethics—not making this demand would generate unnecessary losses. Yet, as we have argued at length in the article, this demand does not require much from market participants. Therefore, in the remainder of this section, we argue that whatever transactions they engage in ought to be restricted by two additional moral constraints on how participants in a social contract can be treated. These constraints constitute the main proposed improvements on the MFA, which will help preserve its contractualist core.

In an article that business ethicists are all too familiar with, Friedman (1970) states that firms ought to maximize their profits, meaning searching for mutually advantageous opportunities of exchange, within "the rules of the game." For Friedman, these rules were restricted to avoiding deception and fraud in the marketplace. For a contractualist, the rules of the game must ensure that all who participate can consent to their participation in the marketplace.[24]

How should the rules of the game hence be determined? A contractualist would insist on at least two kinds of moral constraint that determine the rules of the game. The first relates straightforwardly to the inclusion of participants in the marketplace. If the point of markets is justified on contractualist grounds, it follows that participation in markets must be justifiable to all who de facto are or will be participating in the market. Furthermore, if the point of markets is justified on Paretian grounds, it follows that the intended addressees of this justification are individual persons, rather than aggregates of persons.

For Heath, this justification is built on the claim that markets have the potential to generate social optima. But much more straightforwardly, justification demands first that no potential participant be arbitrarily excluded from the market. This means that racist, sexist, ableist, ageist, and so on forms of discrimination are straightforwardly unjustifiable in the marketplace on any plausible contractualist theory of business ethics. At least in principle, every market participant ought to have access to the full range of (in isolation) mutually beneficial exchanges available in the marketplace and be restricted only by her own epistemic, rather than principled, constraints to participation. This is the first type of moral constraint.

The second requirement relates to the avoidance of adverse third-party effects of market exchanges. In short, we argue that the reason why contractualist theories in general, including contractualist theories of business ethics, ought to condemn negative externalities is simply because they make third parties worse off without compensation and without their consent.

Insofar as mutually beneficial exchanges generate significant benefits for market participants, some market failures can be tolerated. Others, however, clearly cannot. Which market failures can and cannot be tolerated is a matter of justifiability and

[24] It is debatable whether ex ante or ex post consent is relevant in this context. Von Hayek (1966, 611) famously argued that only the former matters morally in the sphere of the market: "once we have agreed to play the game and profited from its results, it is a moral obligation on us to abide by the results, even if they turn against us." We bracket this discussion here owing to space constraints.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

therefore a deeply moral concern. For example, some technical (or nonpecuniary) externalities might be justifiable. It might be justifiable for workers to tolerate a nonthreatening level of air pollution if they have access to good employment opportunities at the local steel mill. Whether this marks a social optimum is beside the point of justifiability. But if their range of opportunities to engage in the labor market is limited to jobs that pay a below-subsistence wage (even if the wage is set at the socially optimal level!) and thus gradually diminish their labor power, then they cannot rationally consent to participating in the market.

The important point here is that the plausible level of justification is not determined by unattainable and unknown social optima or by the constrained efficient status quo. It is determined by the balance of *morally relevant* burdens and benefits to which the market economy exposes market participants. When market participants know or should have known that they are generating or contributing directly or indirectly to market failures that in turn generate morally unjustifiable burdens, they are exploiting market failures that are not only "almost always bad" (Heath 2019) but *necessarily* always bad from a contractualist standpoint. This is the second type of moral constraint.

To conclude, we schematically propose to improve the MFA by committing to its contractualist and Paretian character, while taking seriously the challenge that social optima are unattainable. The result is the sketch of a theory of business ethics that requires market participants to engage in those transactions they consider mutually beneficial, while avoiding exclusionary practices and the exploitation of foreseeable market failures that undermine the contractualist nature of justification.[25] The central advantage that this version of the MFA has over Heath's original vision is that it stands on a much firmer moral ground of justification: market participants are not required to avoid all market failures—they are required only to avoid those that are morally unacceptable to the contractualist spirit. Regulations can help achieve this objective either by minimizing epistemic constraints, such that a wider range of opportunities for mutually beneficial transactions becomes available to market participants, or by determining which market failures can justifiably be tolerated in the market and which cannot.

5. CONCLUSION

Heath has proposed a powerful theory of business ethics, one that promises to take economics seriously, while imposing strong moral obligations on managers. The main claim of his theory is that the market's main function is to achieve Pareto-efficient distributions, that is, social optima. Because competitive markets are, under ideal conditions, particularly apt in generating such distributions, market participants are required to act competitively, and profit-seeking behavior is encouraged. As we argued in section 1, though, even under such ideal conditions, Heath's

[25] In Heath's (2014, 163) terms, we have ultimately defended a version of "microcontractualism" as a defensible theory of business ethics. We reject the notion that a particular pattern of distributions serves as grounds for justification and instead embrace the idea of procedural fairness as grounds for justification.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

approach would allow for significant harm. Perfectly competitive markets do not exhibit market failures, but they exhibit pecuniary externalities, that is, welfare losses incurred by market participants on the losing side of competition. We argue that, even if the resulting distribution is a social optimum, these pecuniary externalities are effectively harm impositions that should be a concern from a perspective of justice, especially in consideration of Heath's wider vision of a minimally controversial contractualism.

As Heath acknowledges, the conditions of perfect competition never apply in real markets. Actual markets are subject to numerous market failures that prevent them from reaching social optima. Heath is acutely aware of these issues, and he significantly modifies his own original theory to account for them. As we have seen in section 2, he faces a choice. He can safeguard the ideal of social optima, but only at a tremendous cost. For the MFA then requires market participants to reconstruct the ideal market mechanism so as to know which prices are market clearing. This renders the MFA overdemanding. Alternatively, he can move the normative focus of his theory from Pareto efficiency toward merely mutually beneficial exchanges. In response, we argue that this constrained version of the MFA is not demanding enough because it does not safeguard win–win transformations, which form the backbone of Heath's justification of the market and the MFA.

Third, the last section analyzed Heath's claim that both the state and specific moral requirements must play a role in helping to achieve efficient distributions. He argues that the state has a role to play when markets fail to generate Pareto-efficient distributions. Its main task is to enforce regulations and laws that prevent or correct market failures. But when regulations alone are unable to achieve Pareto efficiency, Heath resorts to ethical constraints on market participants' behavior. He argues that market participants have a moral duty to act as if they participated in a perfectly competitive market and thus to abide by specific moral requirements not to exploit market failures. In response, we argued that Heath's appeal to regulation and ethical constraints is deeply problematic. We showed that market failures do not arise out of thin air but because economic market participants are either unable or unwilling to avoid taking advantage of information asymmetries, incomplete property rights, or externalities. The MFA will be relevant in cases where market participant are willing to act morally but are ignorant about the right moral actions. In other cases, we argued that it is unclear how the MFA's heuristics could help correcting market failures that were not addressed by regulation. As we demonstrated, this threatens their moral force.

In sum, Heath's theory appears far less morally attractive than he claims. Under ideal conditions, it is unclear how severe harms due to pecuniary externalities ought to be mitigated. Under nonideal conditions, what market participants are required to do merely amounts to securing mutually beneficial exchanges, which undermines efficiency as the justification for markets. The requirement for market participants to act morally when regulations fail will not reliably bring about social optima either.

Finally, we provide some constructive suggestions on how to improve the MFA. Our main change relates to the contractualist justification of the MFA. Rather than aiming to capture the contractualist spirit of the MFA via an outcome-oriented

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

standard of justification that focuses on efficiency, we suggest that this contractualist spirit is better preserved by focusing on a procedural standard of justification. We argue that market participants have an obligation to respect two kinds of moral constraints: first, they ought to avoid straightforwardly discriminatory practices in the marketplace, and second, they ought to avoid market failures that will foreseeably generate morally relevant, severe burdens on third parties.

In conclusion, despite its weaknesses, the MFA marks an important step toward a defensible theory of business ethics. The acknowledgment that business ethics is but part of a larger project in political philosophy is an essential insight that is preserved in the contractualist spirit of the MFA. Given that our more general theory of justice shares this contractualist spirit, a modified version of the MFA constitutes a fitting piece of the puzzle.

Acknowledgments

The authors thank Alexander Andersson, Jasmine C. Elliott, Francesco Montana-Rodriguez, Nina van Heeswijk, Abraham Singer, and Adrian Walsh, as well as the editor and three anonymous reviewers for their insightful comments and suggestions.

REFERENCES

Arrow, Kenneth J., and Frank H. Hahn. 1971. *General Competitive Analysis*. San Francisco: Holden-Day.

Calabresi, Guido. 1991. "The Pointlessness of Pareto: Carrying Coase Further." *Yale Law Journal* 100 (5): 1211–37.

Coase, Ronald H. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3: 1–44.

Cohen, Marc A., and Dean Peterson. 2019. "The Implicit Morality of the Market and Joseph Heath's Market Failures Approach to Business Ethics." *Journal of Business Ethics* 159 (1): 75–88.

Dahlman, Carl J. 1979. "The Problem of Externality." *Journal of Law and Economics* 22 (1): 141–62.

Endörfer, Richard. 2021. "Should Market Harms Be an Exception to the Harm Principle?" *Economics and Philosophy*. DOI: 10.1017/S0266267121000092.

Freeman, R. Edward. 1984. *Strategic Management: A Stakeholder Approach*. Boston: Pitman.

Friedman, Milton. 1970. "The Social Responsibility of Business Is to Increase Its Profits." *New York Times Magazine*, September 13.

Gauthier, David. 1986. *Morals by Agreement*. Oxford: Clarendon Press.

Greenwald, Bruce C., and Joseph E. Stiglitz. 1986. "Externalities in Economies with Incomplete Information and Incomplete Markets." *Quarterly Journal of Economics* 101: 229–64.

Hausman, Daniel M. 1992. "When Jack and Jill Make a Deal." *Social Philosophy and Policy* 9 (1): 95–113.

Heath, Joseph. 2004. "A Market Failures Approach to Business Ethics." In *The Invisible Hand and the Common Good*, edited by Bernard Hodgson, 69–89. Berlin: Springer.

Heath, Joseph. 2011. "Three Normative Models of the Welfare State." *Public Reason* 3 (2): 13–43.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

Heath, Joseph. 2014. *Morality, Competition, and the Firm: The Market Failures Approach to Business Ethics*. Oxford: Oxford University Press.

Heath, Joseph. 2018. "'But Everyone Else Is Doing It': Competition and Business Self-Regulation." *Journal of Social Philosophy* 49 (4): 516–35.

Heath, Joseph. 2019. "Is the 'Point' of the Market Pareto or Kaldor–Hicks Efficiency?" *Business Ethics Journal Review* 7 (4): 21–26.

Holcombe, Randall G., and Russell S. Sobel. 2001. "Public Policy towards Pecuniary Externalities." *Public Finance Review* 9 (4): 304–25.

Hsieh, Nien-hê. 2017. "The Responsibilities and Role of Business in Relation to Society: Back to Basics?" *Business Ethics Quarterly* 27 (2): 293–314.

Lipsey, Richard G., and Kelvin Lancaster. 1956. "The General Theory of Second Best." *Review of Economic Studies* 24 (1): 11–32.

Loong, Lee H., and R. Zeckhauser. 1982. "Externalities Do Matter When Contingent Claims Markets Are Incomplete." *Quarterly Journal of Economics* 97: 171–79.

McMahon, Christopher. 1981. "Morality and the Invisible Hand." *Philosophy and Public Affairs* 10 (3): 247–77.

Moriarty, Jeffrey. 2019. "On the Origin, Content, and Relevance of the Market Failures Approach." *Journal of Business Ethics* 165 (1): 113–24.

Norman, Wayne. 2011. "Business Ethics as Self-Regulation: Why Principles That Ground Regulations Should Be Used to Ground Beyond-Compliance Norms as Well." *Journal of Business Ethics* 102 (1): 43–57.

Olsaretti, Serena. 2004. *Liberty, Desert and the Market: A Philosophical Study*. Cambridge: Cambridge University Press.

Scanlon, Thomas M. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press.

Scitovsky, Tibor. 1954. "Two Concepts of External Economies." *Journal of Political Economy* 62 (2): 143–51.

Sen, Amartya K. 1985. "The Moral Standing of the Market." In *Ethics and Economics*, edited by Ellen Frankel Paul, Fred D. Miller, and Jeffrey Paul, 1–19. Oxford: Blackwell.

Singer, Abraham. 2018. "Justice Failure: Efficiency and Equality in Business Ethics." *Journal of Business Ethics* 149 (1): 97–115.

Smith, Jeffery. 2018. "Efficiency and Ethically Responsible Management." *Journal of Business Ethics* 150 (3): 603–18.

Steinberg, Etye. 2017. "The Inapplicability of the Market-Failures Approach in a Non-ideal World." *Business Ethics Journal Review* 5 (5): 28–34.

von Hayek, Friedrich A. 1945. "The Use of Knowledge in Society." *American Economic Review* 35 (4): 519–30.

von Hayek, Friedrich A. 1966. "The Principles of a Liberal Social Order." *Politico* 31: 601–17.

von Kriegstein, Hasko. 2016. "Professionalism, Agency, and Market Failures." *Business Ethics Quarterly* 26 (4): 445–64.

Wilkinson, Hayden. 2022. "Market Harms and Market Benefits." *Philosophy and Public Affairs* 50 (2): 202–38.

. . .

RICHARD ENDÖRFER is a doctoral candidate and member of the Financial Ethics Research Group at the University of Gothenburg, Sweden. His dissertation combines insights from both philosophy and economics to investigate financial crises qua collective risk impositions. His areas of expertise are risk ethics, collective action problems, and externalities.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press

LOUIS LARUE (louis.larue.2@gu.se, corresponding author) is a researcher in practical philosophy at the University of Gothenburg, Sweden. His research focuses on issues of justice in finance and money and on several topics in the philosophy of economics.

This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is included and the original work is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use.

https://doi.org/10.1017/beq.2022.21 Published online by Cambridge University Press