UNDERSTANDING MARKETS WITH SOCIALLY RESPONSIBLE CONSUMERS*

Marc Kaufmann
Peter Andre
Botond Kőszegi

Many consumers care about climate change and other externalities associated with their purchases. We analyze the behavior and market effects of such "socially responsible consumers" in three parts. First, we develop a flexible theoretical framework to study competitive equilibria with rational consequentialist consumers. In violation of price taking, equilibrium feedback nontrivially dampens the impact of a person's consumption on aggregate consumption, undermining the motive to mitigate. This leads to a new type of market failure, where even consumers who fully "internalize the externality" overconsume externality-generating goods. At the same time, socially responsible consumers change the relative effectiveness of taxes, caps, and other policies in lowering the externality. Second, since consumer beliefs about and preferences over their market impacts play a crucial role in our framework, we investigate them empirically via a tailored survey. Consistent with our model, consumers are often consequentialist, and many believe that they have a dampened impact on aggregate consumption. Inconsistent with our model, however, we also find many respondents who expect to have a one-to-one impact on aggregate consumption. Third, therefore, we analyze how such "naive" consumers modify our theoretical conclusions. They consume less than rational consumers in a single-good economy, but may consume

*Anik De, Khanh Hoang, and Wenjun Zheng provided fantastic research assistance. We are indebted to Stefanie Stantcheva and five anonymous referees for constructive comments, and to Youssef Benzarti, Simon Gächter, Hendrik Hakenes, Oliver Hart, Fabian Herweg, Mats Köster, Mechthild von Knobelsdorff, Andreas Lange, Stephan Lauermann, Tímea Molnár, Chris Roth, Klaus Schmidt, Fred Schroyen, Christoph Semken, Philipp Strack, James Stratton, and Paul Voss, as well as seminar and conference audiences for insightful discussions. Kőszegi thanks the European Research Council for financial support under grant no. 788918. Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A01) is gratefully acknowledged. Funded by the DFG under Germany's Excellence Strategy–EXC2126/1-390838866. We also gratefully acknowledge research support from the Leibniz Institute for Financial Research SAFE and Behavior and Inequality Research Institute (briq).

The Quarterly Journal of Economics (2024), 1989–2035. https://doi.org/10.1093/qje/qjae009.
Advance Access publication on March 18, 2024.

more in a multigood economy with cross-market spillovers. A mix of naive and rational consumers may yield the worst outcomes. *JEL codes:* D01, D11, D50, D62, D64, D91.

I. Introduction

When studying implications of externalities from market trade, economists overwhelmingly assume that individuals do not care about the externalities they cause. But many consumers do care: they are willing to reduce or modify their consumption to alleviate the associated climate change, adverse working conditions, animal suffering, or other externalities.[1] This article investigates equilibrium behavior, welfare, and the effectiveness of interventions in markets with such "socially responsible consumers."

We proceed in three steps. First, as a methodologically natural and flexible benchmark, we develop and study a competitive-equilibrium framework with rational consequentialist consumers. We show that equilibrium feedback nontrivially dampens the effect of a person's consumption on the externality, eroding the motive to mitigate. This implies both that price taking is violated and that even consumers who fully "internalize the externality" overconsume externality-generating goods. We compare the effectiveness of taxes and caps in lowering the externality, and identify policies that outperform both classical interventions. Second, since consumer beliefs about and preferences over their aggregate impact play a crucial role in our framework, we investigate them via a tailored survey. We find that consumers are often consequentialist, and many believe that they have a dampened aggregate impact. However, we also find a large subgroup of consumers who expect to have a one-to-one, undampened impact. Third, therefore, we return to our framework and analyze how the latter "naive" consumers modify our conclusions. We establish that they consume less than rational consumers in some situations, but more in others.

1. Studies abound. See Sundt and Rehdanz (2015), Jalil, Tasoff, and Bustamante (2020), and Schulze Tilling (2023) on climate change; Auger et al. (2003) and Hainmueller, Hiscox, and Sequeira (2015) on fair-trade products; Norwood and Lusk (2011) and Grethe (2017) on animal welfare; and Bartling, Weber, and Yao (2015) on externalities in a laboratory environment.

We begin the first part of our article in Section II by modeling the market behavior of a single rational consequentialist consumer. The consumer's utility equals $u(c) - pc - kg$, where $c > 0$ is her consumption, $k > 0$ is the strength of her social concern, $p > 0$ is the price, and g is an externality equal to total equilibrium output. To study a consumer who is small relative to the market, we posit I additional consumers who are identical to each other, as well as I identical suppliers. Whatever demand c the consumer submits, the price p is chosen to clear the entire market.

Contradicting the typical view that small consumers are price takers, we demonstrate that our consumer keeps thinking about her impact on the price p even as $I \to \infty$. If she lowers her demand to reduce the externality, she lowers the price and thereby raises others' demand, dampening the reduction she brings about. Furthermore, while the price effect is proportional to $\frac{1}{I}$, there are I other consumers who react, so their dampening effect does not vanish. Understanding this in turn lowers the consumer's motivation to mitigate. But because she does have an impact on the aggregate externality, she still consumes less than a selfish consumer ($k = 0$).

The behavior described above lays the foundation for modeling markets with many socially responsible consumers. In this setting, we define a competitive equilibrium as a situation in which the market clears, each consumer maximizes her utility given the degree of dampening, and that degree is in turn consistent with consumer behavior. For the evaluation of welfare outcomes by a social planner, we denote by K the (exogenously given) marginal social cost of the externality.

Section III identifies fundamental properties of competitive equilibria. Adding a new insight to basic welfare economics, we show that the market fails not only when consumers do not care about the externalities they cause, but also when they do care. As an illustrative if extreme case, suppose that each consumer fully "internalizes the externality," assigning the same weight to it as the planner ($k = K$). This means that she and the planner are willing to pay the same amount out of her funds for reducing the externality. Still, while output is lower than with selfish consumers, dampening implies that any equilibrium features overconsumption, and multiple suboptimal equilibria may arise. Further, due to path dependence it is plausible that society converges to the highest-consumption, and hence worst, equilibrium. These are market failures: if the good is not traded in the market, then consumers with $k = K$ bring forth the socially optimal outcome.

In Section IV, we compare policies that aim to improve outcomes by introducing fee-based production permits. A broad principle emerges: policies that induce lower dampening are better at motivating consumers to cut consumption, and thus achieve better outcomes. As an instance of this principle, we confirm and generalize in our setting Herweg and Schmidt's (2022) insight that a unit tax is superior to a cap. We also show that the planner can do even better with a tax that increases when the consumer price decreases. Such a policy mutes other consumers' response to a price decrease, lowering dampening. We also identify conditions under which a cap is better than a tax when there is trade with a polluting partner who is not subject to domestic policy. With a cap, a consumer's effect on dirty foreign production is dampened less. This insight is also relevant for policy debates regarding the efficacy of abatement efforts under a cap.

In Section V, we generalize our model by allowing for two products that are perfectly interchangeable in consumption utility but generate different externalities. To understand the scope for mitigating the externality through product selection rather than quantity reduction, we impose as a start that consumers have unit demand. Then, there is always an equilibrium in which the two products sell at the same price, yet consumers are indifferent between them. In such an equilibrium, socially responsible consumers behave as if they were selfish, and the clean product enjoys no advantage in the market. Worse, this selfish equilibrium may be the sole equilibrium, for instance, if the clean product is not significantly cleaner than the dirty one. Intuitively, if—acting selfishly—everyone chooses the cheaper product, then the market always equilibrates at the uniform market-clearing price. This means that an individual consumer cannot affect the externality, that is, dampening is full. Consistent with equilibrium, therefore, everyone prefers the cheaper product.

When consumers also decide how much of the chosen product to purchase, a new cross-market effect arises. If a consumer raises her consumption of the clean product and thereby raises its price, she induces someone else to switch to the dirty product, raising production in that market. Unlike in a single-market setting, therefore, the externality generated from consuming a product can exceed the product's direct externality.

In Section VI, we turn to the second major part of our article: an empirical investigation of two crucial ingredients for the results above. First, due to rationality, consumers understand that reducing their own consumption has a dampened market-level effect. Instead, consumers could be “naive” in that they expect to have a one-to-one effect. Second, because of consequentialism, consumers care about their dampened effect on the externality. Instead, consumers could be “deontological” in that they care about their action or direct effect. To distinguish these possibilities, we survey 2,000 U.S. consumers. We measure beliefs about one’s impact by asking consumers how they think a change in their own consumption would affect aggregate consumption. Using eight specific externality-generating goods, we study reductions in consumption, including of fuel and meat, and reallocations of consumption, including from brown to green electricity and from new to secondhand clothing. We measure the nature of social concerns by comparing consumers’ values for a specific action, such as a reduction in one’s CO_2 emissions, when it does versus when it does not reduce the aggregate externality. In all cases, we gather qualitative text data in which consumers explain their responses, allowing us to shed light on their reasoning.

We document three main facts. First, many consumers (38%) believe that a change in their own consumption has a dampened effect on aggregate consumption. Some respondents explicitly argue that their effect would be offset by other consumers, while many simply observe that they are too small to have an aggregate impact. Second, however, even more consumers (54%) express the naive view that a change in their consumption affects aggregate consumption one to one. Third, consumers often care about the externalities they cause in a consequentialist way.

These findings imply that an empirically realistic model must include consequentialist consumers who expect to have a dampened effect on markets. Hence, our framework above modeling such consumers is a natural theoretical benchmark, especially in light of the general-equilibrium and price-theory literature’s traditional focus on rational consumers. However, the evidence also highlights that it is important to understand the effect of naive consumers, particularly when they coexist with rational consumers in the market.

Motivated by this last observation, the third part of our article in Section VII begins analyzing how naive consumers affect equilibrium outcomes. Since deontological consumers behave identically to naive consumers, our analysis also applies to this group. In our single-good model, naive consumers generate higher welfare than rational consumers. Intuitively, a naive consumer ignores dampening, so she overestimates her impact on the market, which encourages her to reduce consumption. But in multimarket settings, naive consumers may underestimate their impact and hence consume more than rational consumers. As an illustration, consider our two-good model with both product and quantity choice when the clean product is in fully inelastic supply and the dirty product is in fully elastic supply. If the marginal consumer is rational, then the presence of naive consumers leaves dampening unchanged, so prices are uniform. A naive consumer therefore chooses the clean product, believing that this does good at no cost to herself. Furthermore, not understanding the cross-market effect—that consumption of the clean product raises dirty output—she consumes more than a rational consumer. With such a mixed population, therefore, equilibrium welfare is lower than with a rational population.

We conclude by emphasizing two points. First, if read as advice to consumers wondering about their market impacts, our theory has an unambiguous message. It is: yes, you do have an impact, and if you care about it, you should modify your behavior (being aware of cross-market effects). This is true even though our market-failure results make clear that dealing with externalities requires systemic changes too. Second, our equilibrium framework can be adapted in a formulaic way to other situations with externalities and is easily adjusted to other kinds of social concerns. Hence, it opens the possibility for studying the behavior of socially conscious consumers in myriad other settings, some of which we mention.

The Appendix and the Online Appendix provide background on our framework and survey, and contain all our proofs.

I.A. Related Literature

Our article relates to several literatures. No previous work, however, recognizes the violation of price taking for socially responsible consumers or incorporates the resulting dampening into a model of behavior and competitive equilibrium in a standard product market.

There is a large literature on how to think about equilibria in markets with many consumers. A long-standing axiom is that small consumers take the market price as fixed

(Debreu 1959; Arrow and Hahn 1971), and extensive research establishes conditions under which this is approximately optimal (e.g., Mas-Colell 1980). Because market behavior depends only on prices, an intuitive implication of taking prices as fixed is that a person takes others' behavior as fixed as well. Existing research on the market consequences of social preferences has overwhelmingly—and sometimes implicitly—adopted such an interpretation (Sobel 2007; Dufwenberg et al. 2011; Hakenes and Schliephake 2021; Pástor, Stambaugh, and Taylor 2021; Piccolo, Schneemeier, and Bisceglia 2022; Aghion et al. 2023; Arnold 2023; Dewatripont and Tirole forthcoming). We show that our socially responsible consumers violate price taking, and we define a variant of competitive equilibrium that appropriately accounts for such consumers' incentives.

Starting from Heinkel, Kraus, and Zechner (2001), a few researchers have pointed out that price changes induced by unselfish behavior can alter the benefits of that behavior. Our dampening effect stems from the same mechanism—but applied to vanishingly small consumers. Although they do not frame it as a violation of price taking and show dampening in a different form, this starting point parallels two previous analyses.[2] Norwood and Lusk (2011, chapter 8) heuristically derive the dampening effect in the context of advice to consumers who care about animal welfare. Broccardo, Hart, and Zingales (2022) study the behavior of small investors and consumers who take their price-mediated impacts into account. In their paper, the impacts act through managers' investment decisions and are not crucial for the main results. In contrast, our dampening acts through the product market, and its implications are our primary interest, leading to different insights.

Other research exploring the above types of price effects focuses on actors with market power. The literature on carbon leakage (e.g., Felder and Rutherford 1993; Babiker 2005; Burniaux and Oliveira Martins 2012; Perino 2015) argues that CO_2 reductions in one region may lower prices and thus raise emissions in other regions. This research typically assumes selfish decision makers. In some recent models of socially responsible investment, investors—worrying about the "additionality" of their

2. See also Trammell (2023), who understands that consumers with ethical concerns violate price taking but does not develop a framework like ours or discuss economic implications.

efforts—are cognizant of their price impacts and modify their behavior accordingly (e.g., Moisson 2020; Green and Roth 2021; Hakenes and Schliephake 2021; Oehmke and Opp 2022; Krahnen, Rocholl, and Thum 2023). Because we study competitive equilibria in product markets with small socially responsible consumers, our results are completely different from those in either literature.

Following Weitzman (1974), many researchers have investigated the optimal regulation of externalities. We contribute to this literature by studying policies in markets with socially responsible consumers. Most related, Herweg and Schmidt (2022) show that a tax dominates a cap in a closed economy with such consumers. They capture consumers' motive to mitigate through an exogenous parameter, β^R, that can be different under different policies. Then, recognizing that under a cap rational consumers anticipate (what we call) full dampening, they impose $\beta^R = 0$ for this case. In our model, the motive to mitigate depends explicitly on the degree of dampening, which we derive endogenously from the economic situation. This allows us to consider other questions, such as identifying better policies or comparing a cap and a tax under trade, without making additional exogenous assumptions.

Our theoretical analysis raises the novel empirical question of whether consumers anticipate and care about dampening. Previous experimental work on small-scale strategic settings has found that people follow the "replacement logic," choosing less-morally if someone else would otherwise bring about the same outcome anyway (Falk, Neuber, and Szech 2020; Ziegler, Romagnoli, and Offerman forthcoming). We show that many consumers think in similar terms about their impact in large markets. We thus identify a new channel that affects individuals' willingness to mitigate externalities in markets, relevant for the large literature on socially responsible consumer behavior.

II. Large Markets with Socially Responsible Consumers

We develop our benchmark framework in two steps. First, we derive a condition for the optimal consumption choice of a single socially responsible consumer. Second, we show how to incorporate this condition into the definition of competitive equilibrium for a market with many socially responsible consumers.

Following the traditions of utility theory, general-equilibrium analysis, and price theory, our benchmark framework assumes that consumers are rational and consequentialist. We reconsider these assumptions in Sections VI and VII. The insights we uncover in the rational consequentialist model will remain pertinent there too.

II.A. A Single Consumer's Perspective

1. *Setup.* We consider a single-good market and study a single consumer, such as a household, municipal government, or small organization, who is a tiny part of the market. To do so, we use a "replicator economy" (Shubik 1973; Roberts and Postlewaite 1976): we introduce identical copies of the other participants, and let the number of copies approach infinity.

There are, then, I other consumers and I suppliers. The other consumers all have the same demand curve $D(p)$, and the suppliers all have the same supply curve $S(p)$. Both curves are continuously differentiable, with $D'(p) < 0$ and $S'(p) > 0$ everywhere. There is a price $p^* > 0$ for which $S(p^*) = D(p^*)$, and $\lim_{p\to\infty} S(p) - D(p) = \lim_{p\to 0} D(p) - S(p) = \infty$.

The market mechanism, consistent with the notion of a Walrasian auctioneer, is the following. First, the consumer submits her demand $c \in \mathbb{R}$. Then, the price $p(c) > 0$ is chosen to clear the market, that is, to satisfy $c + ID(p(c)) = IS(p(c))$; $p(c)$ clearly exists and is unique. Finally, the equilibrium quantity $q(c) = IS(p(c))$ is produced and consumed, generating an aggregate externality $g(c) = q(c)$.[3]

The consumer correctly predicts the above outcomes, and maximizes

$$u(c) - p(c)c - k(g(c) - g(0)), \tag{1}$$

where $u(\cdot)$ is a thrice differentiable strictly concave function satisfying $\lim_{c\to-\infty} u'(c) = \infty$ and $\lim_{c\to\infty} u'(c) = 0$, and $k > 0$ is a constant. The term $u(c)$ is consumption utility, $u(c) - p(c)c$ is private utility, and $-k(g(c) - g(0))$ represents social concerns. Capturing our notion of social responsibility—being willing to modify one's consumption to mitigate the associated

3. In this model, the notation g is redundant. Below, we extend our model to markets with multiple products that generate different externalities. By assuming that the consumer cares about g, we can introduce such modifications without changing the consumer's utility function.

externality—the consumer derives disutility from her effect on the aggregate externality relative to when she consumes nothing. Because from the consumer's perspective $g(0)$ is a constant (it equals $IS(p^*)$), we drop it when analyzing consumer behavior, and work with the utility function $u(c) - p(c)c - kg(c)$. We did not start with this formulation because assuming that the consumer internalizes the aggregate externality generated by everyone seems psychologically less realistic.

In the Appendix, we show how to derive the objective (1), as well as the social-welfare function (5), from a more fully specified utility model. In that model, an externality arises because average consumption affects consumers' private utilities, for instance, through the effects of global warming on health and well-being. A socially responsible consumer internalizes part of her externality effect on others. While she cares about the externality also because it affects her own private utility, this motive vanishes as $I \to \infty$, so expression (1) excludes it to start with.

2. *Analysis.* Differentiating the market-clearing condition with respect to c and rearranging gives

$$p'(c) = \frac{1}{I(S'(p(c)) - D'(p(c)))}.$$

The total quantity produced is $q(c) = IS(p(c))$, on which the consumer's demand has an effect of

$$q'(c) = IS'(p(c))p'(c) = \frac{S'(p(c))}{S'(p(c)) - D'(p(c))}.$$

Taking limits yields:

PROPOSITION 1 (Violation of Price Taking). Take any $D(\cdot)$ and $S(\cdot)$.

i. A vanishingly small consumer has a negligible impact on the price: for any c,

$$\lim_{I\to\infty} p(c) = p^* \quad \text{and} \quad \lim_{I\to\infty} p'(c) = 0.$$

ii. The same consumer has a nonnegligible impact on others' consumption: for any c,

$$\lim_{I\to\infty} q'(c) = \frac{S'(p^*)}{S'(p^*) - D'(p^*)} \in (0, 1). \tag{2}$$

Part i replicates the insight of Roberts and Postlewaite (1976) and others that in a large economy, a consumer has a negligible price impact. This means that in terms of evaluating her expenditure, she takes the price as given.

Even so, part ii says that when a vanishingly small consumer raises her demand c, the "market responsiveness" $q_c \equiv \lim_{I\to\infty} q'(c)$ to it is less than one, which means that she impacts others' consumption to a nonvanishing extent. In particular, others have a dampening effect of

$$\lim_{I\to\infty}\left|\frac{\mathrm{d}[ID(p(c))]}{\mathrm{d}c}\right| = \lim_{I\to\infty}[-ID'(p(c))p'(c)]$$
$$= \frac{-D'(p^*)}{S'(p^*)-D'(p^*)} = 1 - q_c. \tag{3}$$

Intuitively, an increase in the consumer's demand c raises the price, which leads others to consume less. Although the price impact vanishes at rate $\frac{1}{I}$, there are I other consumers, so their total response is comparable to c. This logic derives from market clearing, so it is independent of whether consumers know or care about it. Unlike a selfish consumer, however, a socially responsible consumer does care about it. Indeed, an infinitesimally small consumer chooses c to solve[4]

$$\lim_{I\to\infty}\frac{\mathrm{d}}{\mathrm{d}c}[u(c)-p(c)c-kq(c)] = \lim_{I\to\infty}[u'(c)-p'(c)c-p(c)-kq'(c)]$$
$$= u'(c) - p^* - k\cdot q_c = 0, \tag{4}$$

so she cares about her dampened effect on the externality. In the sense that her price impact is relevant for optimization, therefore, the consumer is not a price taker.[5]

Two immediate points follow. First, because the consumer has an impact on the externality she derives disutility from ($q_c > 0$), she consumes less than a selfish consumer ($k = 0$). Hence, in our specification social responsibility affects market behavior for a broad range of market environments (any $D(\cdot)$ and $S(\cdot)$ with $S'(p^*) > 0$ and $D'(p^*) < \infty$). In contrast, some previous models

4. The conditions we have imposed on $u(\cdot)$ ensure that equation (4) has a unique solution for any p^*, and the solution satisfies the second-order condition.

5. In reality, firms do not constantly reoptimize prices, and when they do, they make discrete adjustments. Under such frictions, a small consumer has a discrete price impact with a small probability rather than a small price impact with certainty. The fundamental logic of our model, however, appears unchanged. The consumer's impact on the probability of a price change is likely to vanish at rate $\frac{1}{I}$, but I others respond if the price change does occur. Hence, the consumer's effect is unlikely to vanish in expectation.

predict that an agent with social preferences chooses her market consumption selfishly (Dubey and Shubik 1985; Dufwenberg et al. 2011; Arnold 2023).

Second, however, dampening ($q_c < 1$) implies that consumption is higher than it would be under home production with the same private marginal cost, p^*. With home production, the consumer would not affect others' behavior, so she would solve $u'(c) - p^* - k = 0$. This erosion of moral behavior by the market also occurs in experimental research (e.g., Falk and Szech 2013; Bartling, Weber, and Yao 2015) using different trading rules. Proposition 1 shows that the erosion is a fundamental property of the standard price-based market mechanism. Further, the degree of dampening $(1 - q_c)$ identifies how much of the consumer's social motivation is eliminated by the market.

The optimality condition (4) also clarifies how rationality and consequentialism matter for the consumer's behavior. They enter through her response to dampening: rationality implies that she expects dampening to occur, and consequentialism implies that she cares about her dampened effect. Hence, our conclusions do not require that the consumer understand the precise mechanism behind dampening, nor that she is consistently consequentialist (e.g., also in nonmarket behavior). If the consumer is consequentialist with respect to externalities in markets, and she expects to have a dampened effect ($q_c < 1$) for whatever reason, then she behaves according to equation (4).

Using equation (3) and the market-clearing condition $D(p^*) = S(p^*)$, dampening is an increasing function of the responsiveness or elasticity of demand relative to that of supply, $-\frac{D'(p^*)}{S'(p^*)} = -\frac{\frac{D'(p^*)}{D(p^*)}}{\frac{S'(p^*)}{S(p^*)}}$. If this relative elasticity is low, then it is mostly supply that responds to an increase in the consumer's demand, so the dampening effect is small. An example is when suppliers can easily adjust production because their technology approximates constant returns to scale. But if the relative elasticity is high, then it is mostly demand from other consumers that responds to an increase in the consumer's demand, so the dampening effect is large. An example is when producers cannot flexibly expand production due to capacity or input constraints.[6]

6. Notably, the above relative elasticity also plays a central role in the incidence of a commodity tax. In that context, it determines the relative ease with which consumers and producers can avoid the tax. In fact, the degree of dampening $(1 - q_c)$ equals the incidence on producers. Given this identity, one can

II.B. Competitive Equilibrium

We turn to markets with many socially responsible consumers, each of whom behaves like the consumer above. For simplicity, we assume that consumers are homogeneous and that the supply curve $S(\cdot)$ is exogenous and linear: $S(p) = sp$, with $s > 0$.[7] To ensure that the competitive equilibrium defined below exists, we also impose that $u'(0) > k$. Again, our interest is in what happens in the limit as the number of consumers approaches infinity.

A version of the classical definition of competitive equilibrium involves a quantity q^* and a price p^* such that supply at p^* and demand at p^* both equal q^*. Because supply is exogenously given, the condition on supply is immediate from primitives. With socially responsible consumers, however, the condition on demand is more complicated because the market responsiveness q_c both affects and is affected by consumer behavior. We build on our analysis of individual behavior to extend the classical definition using the following steps.

i. We introduce the equilibrium market responsiveness to the consumer's demand, q_c^*.
ii. We write demand as a function of q_c^*. From a consumer's first-order condition (4), we obtain $u'(q^*) = p^* + kq_c^*$.
iii. We write q_c^* as a function of supply and demand. From equation (2), we have $q_c^* = \frac{S'(p^*)}{S'(p^*)-D'(p^*)}$. In the expression, $S'(p^*) = s$, but $D'(p^*)$ is not a primitive of the model. Hence, we introduce the equilibrium consumer price responsiveness $q_p^* \equiv D'(p^*)$.
iv. We derive q_p^* from consumer optimization, imposing that q_c^* does not change in response to an infinitesimal price change. Then, differentiating equation (4) with respect to the price yields a standard expression for price responsiveness, $q_p^* = \frac{1}{u''(q^*)}$.

Combining these considerations yields the following definition of competitive equilibrium.

---

use empirical work on tax incidence in an appropriately chosen market to gauge the level of dampening. For a broad externality effect like the impact of CO_2 emissions on climate change, for instance, one would use the producer incidence of a broad carbon or other tax that covers a wide range of polluting consumption.

7. Incorporating consumers who are heterogeneous in their consumption utilities u or their weights k is straightforward but requires additional notation. Because our insights derive from demand-side considerations, the simplifying assumption regarding the supply curve does not affect any of our points.

DEFINITION 1. A competitive equilibrium consists of a quantity $q^* > 0$, price $p^* > 0$, consumer price responsiveness $q_p^* < 0$, and market responsiveness $q_c^* > 0$ that satisfy the following conditions:

i. supply equals q^*: $q^* = S(p^*)$;
ii. demand equals q^*: $u'(q^*) = p^* + k \cdot q_c^*$;
iii. market responsiveness is consistent with consumer price responsiveness: $q_c^* = \frac{s}{s - q_p^*}$;
iv. consumer price responsiveness is consistent with optimization: $q_p^* = \frac{1}{u''(q^*)}$.

Steps i–iv can be applied mechanically to define competitive equilibrium in new situations, as we do in Sections IV and V. Step i recognizes the crucial market responsiveness the consumer cares about, and step ii expresses demand as a function of that market responsiveness. These steps follow directly from consumer preferences. Then, Steps iii and iv express market responsiveness as a function of demand. Step iii uses market clearing; in Sections IV and V, just applying equation (2) to the situation at hand. Finally, Step iv derives price responsiveness from the consumer's optimality condition in step ii.

In Online Appendix A, we provide some foundational analysis for our equilibrium concept. In Section A.1, we outline a way to think about equilibrium determination graphically. In Sections A.2 and A.3, we develop microfoundations that are analogous to the replicator economy we have used for an individual consumer. Specifically, we start from the interaction of finitely many socially responsible consumers. To solve for equilibrium, we adapt methods that Kyle (1989) and the subsequent literature have developed to model financial-market participants with nontrivial price impacts (see Rostek and Yoon 2023 for a review). We establish that Definition 1 describes the limit of equilibrium outcomes as the number of players approaches infinity.

As a basic point, we note:

OBSERVATION 1. A competitive equilibrium exists.

To complete our setup, we define social welfare when everyone consumes an amount q as

$$u(q) - \int_0^q S^{-1}(x)\mathrm{d}x - Kq. \tag{5}$$

One part of social welfare is total consumption utility net of the costs of production. In addition, the social planner puts an exogenously given weight $K > 0$ on the externality. In the Appendix, we motivate this specification in more detail using the same foundation as that for the consumer's objective. In particular, we argue that in the context of social responsibility, it is inappropriate to equate consumer welfare with the sum of individual utilities, as this would lead to multiple-counting utilities. We also argue that a natural assumption is $k \leq K$. This assumption means that consumers internalize their effect on others through the externality at most fully; $k = K$ corresponds to the extreme where they do so fully.

III. Failures of Socially Responsible Consumerism

In this section, we demonstrate ways the market's ability to coordinate socially responsible behavior is limited. Our first result describes a general market failure:

Proposition 2 (Overconsumption). There is a unique socially optimal quantity q^{FB}. For any $k \leq K$, any competitive-equilibrium quantity q^* is strictly greater than q^{FB}.

Proposition 2 adds to our understanding of the basic economics of markets. A fundamental Econ-1 lesson is that when there are no externalities or other frictions, markets perform efficiently despite everyone favoring their own private consumption. Another fundamental Econ-1 lesson is that when each person's consumption creates an externality she does not care about, markets perform poorly. Going beyond these insights, Proposition 2 says that when each person's consumption contributes to an externality she does care about, markets still perform poorly.

As an extreme but illustrative special case, suppose that $k = K$—that is, consumers attach the same value to the externality as the social planner. Imagine that the planner values a \$200 increase in a citizen's funds the same as reducing atmospheric CO_2 by a ton. Then, $k = K$ means that the citizen is also willing to give \$200 to reduce atmospheric CO_2 by a ton. As everyone's preferences internalize the externality, the only disagreement between individuals regards their private consumption. One would then think that—as without externalities—the social optimum again obtains. Indeed, such an internalization logic is exactly the rationale economists typically give for the efficiency of Pigouvian taxes (e.g., Gruber 2005). Yet Proposition 2 says that a similar logic does not apply to socially responsible consumers. While consumption is lower in their presence than with selfish consumers, it is still too high.

The first-pass intuition for the market failure in Proposition 2 is extremely simple: dampening reduces each consumer's incentive to cut back, leading to overconsumption. A deeper perspective is based on a pecuniary externality that emerges in the presence of social responsibility. Since consumers are restraining consumption, their marginal private utilities $u'(c_i) - p$ are positive. Hence, when a consumer reduces others' consumption through her price impact, she reduces the private part of social welfare. Even a socially responsible consumer disregards this effect.

Like the erosion of moral behavior in the previous section, overconsumption is facilitated by the market environment. To see this formally, consider the following modification of our model. The number of individuals $I + 1$ is finite, and there are no suppliers. Consumer i obtains utility $u_i(c_i) - kg$, where c_i is a home-produced good, such as cattle raised and grazed on her property, $u_i(\cdot)$ satisfies the same conditions as $u(\cdot)$ above, and $g = \Sigma_j c_j$ as before. Social welfare equals $[\Sigma_i u_i(c_i)] - Kg$. Then the following is obvious:

OBSERVATION 2. With nonmarket consumption, there is a unique socially optimal consumption profile $(c_1^{\text{FB}}, \ldots, c_{I+1}^{\text{FB}})$. If $k = K$, then in the unique Nash equilibrium, consumer i chooses c_i^{FB}.

Because there is no dampening, a fully responsible consumer ($k = K$) trades off the private benefit of consumption with the full externality generated. Equivalently, she exerts no pecuniary externality on others' private utility. Hence, she chooses the socially optimal level of consumption.

Paralleling its message for economics, the market failure we have identified is relevant for society's vision for dealing with externalities. The observation that many consumers care raises the hope that we can use a decentralized, voluntary approach to overcome social problems. Indeed, Giesler and Veresiu (2014) and Chater and Loewenstein (2023) argue that there has been a shift in emphasis toward solutions based on consumer responsibility rather than systemic reform.[8] Our result that the social optimum does not obtain even with $k = K$ implies that dealing with externalities cannot be successfully outsourced even to extremely responsible individuals, so it highlights a fundamental weakness of consumer-driven solutions.

We now show that beyond overconsumption in any equilibrium, multiple equilibria can arise.

PROPOSITION 3 (Multiple Equilibria). Fix any $u(\cdot)$, k, and s, and take a resulting competitive equilibrium (q^*, p^*, q_p^*, q_c^*). If $u'''(q^*)$ is sufficiently high, then there are competitive equilibria (i) q^+, p^+, q_p^+, and q_c^+ with $q^+ > q^*, p^+ > p^*, |q_p^+| > |q_p^*|$, $q_c^+ < q_c^*$; and (ii) q^-, p^-, q_p^-, and q_c^- with $q^- < q^*$, $p^- < p^*$, $|q_p^-| < |q_p^*|, q_c^- > q_c^*$. Among multiple equilibria, social welfare is strictly decreasing in the equilibrium quantity.

The condition for multiple equilibria to exist is that $u'''(c)$ is high over a range, so that consumer price responsiveness $\frac{-1}{u''(c)}$ increases fast. As an example, consumers' utility function for air travel may decrease sharply below but flatten out quickly above 1,000 km, the distance to a nice vacation destination. Then, if a consumer expects everyone to fly 1,000 km, she expects demand to be price-insensitive, so by equation (3) she thinks that dampening is low. Consistent with equilibrium, therefore, the consumer is motivated to mitigate and flies 1,000 km. But if the consumer expects everyone to fly much further for vacation, she expects demand to be price-sensitive and hence dampening to be high. Again consistent with equilibrium, therefore, she is not motivated to mitigate and flies far.

While we have not developed formal selection criteria, informal arguments based on path dependence suggest that in stable supply conditions, the worst equilibrium is the most likely outcome. The detrimental effects of many externalities, such as climate change or ocean pollution, have not been

8. As a prominent example, consumer responsibility appears to be an essential part of "stakeholder capitalism" as advocated by Klaus Schwab, founder of the World Economic Forum in Davos. In this view, individuals should affect change as consumers, investors, and business leaders, rather than rely on "shareholder capitalism" (pure profit maximization) or "state capitalism" (government intervention). See Klaus Schwab, "Why We Need the 'Davos Manifesto' for a Better Kind of Capitalism, World Economic Forum, December 1, 2019, https://www.weforum.org/agenda/2019/12/why-we-need-the-davos-manifesto-for-better-kind-of-capitalism/).

appreciated until recently. Therefore, the market has been in a high-consumption equilibrium (that with $k = 0$). Even as consumers find out about the high social cost of consumption, they also realize that the current equilibrium is one in which everyone consumes a lot. Then, it is plausible that society equilibrates at the highest-consumption equilibrium. Reinforcing this miscoordination issue is that different consumers may recognize the problem at different points in time, creating no obvious focal point for switching to a better equilibrium.

At the same time, temporary shifts in supply may permanently change the equilibrium. Suppose, for example, that an oil shock increases the price of air travel to a level where even selfish consumers would fly just 1,000 km. This reduces dampening, so as the shock dissipates and prices drop, consumers may naturally stay in the low-consumption equilibrium.

IV. Policy

In this section, we assess the effectiveness of different policies in mitigating overconsumption in our model. We focus on market-based approaches in which producers must purchase permits at a fee τ per unit of the good, and proceeds from permit sales are lump-sum redistributed to consumers. Policies differ in how τ is determined.

IV.A. Permit-Supply Policies

First we analyze policies under which the planner acts as a supplier of permits. We define such a "permit-supply policy" as a curve $\pi g - (1 - \pi)\tau + \pi_0 = 0$, where $g \geq 0$ is the amount of permits sold, $\tau \geq 0$ is the permit fee, and $\pi \in [0, 1]$ and $\pi_0 \in \mathbb{R}$ are exogenous constants that parameterize the policy. Two commonly analyzed types of policies arise as extreme cases. If $\pi = 1$, then $g = -\pi_0$, so a binding cap of $-\pi_0$ is in place. This creates a completely inelastic supply of permits. If $\pi = 0$, then $\tau = \pi_0$, so a fixed unit tax of π_0 is in place. With the planner willing to supply any number of permits at the same fee, this creates an infinitely elastic supply of permits. Intermediate cases can capture, in reduced form, a hybrid system that may be closest to current political reality. For instance, even if the nominal policy features a fixed supply of permits, the planner may promise to increase the supply should the permit fee—and thus the associated economic cost—increase. The lower is π, the more permit supply g responds to changes in the permit fee τ.

We define the demand side of competitive equilibrium by following the steps i–iv outlined before Definition 1. It is immediately apparent that only step iii, the expression for market responsiveness q_c^*, requires modification. Here, we let p be the consumer price, and show that there is a unique level of supply, which we denote by $S_{\text{net}}(p)$, that is consistent with market clearing for permits and price p. Then, equation (2) implies $q_c^* = \frac{S'_{\text{net}}(p^*)}{S'_{\text{net}}(p^*)-D'(p^*)}$.

To obtain $S_{\text{net}}(p)$, note that if the consumer price is p, then the producer price is $p - \tau$, so supply is $S(p - \tau) = s \cdot (p - \tau)$. Hence, market clearing for permits ($g = S(p - \tau)$) requires

$$\pi s(p - \tau) - (1 - \pi)\tau + \pi_0 = 0. \tag{6}$$

Consider a $p^* > 0$ and $\tau^* > 0$ with $p^* - \tau^* > 0$ that solve equation (6). Then, for any p in the neighborhood of p^*, there is a unique $\tau(p) > 0$ that solves equation (6), and $\tau(p)$ is differentiable with $\frac{\mathrm{d}\tau}{\mathrm{d}p} = \frac{\pi s}{1-\pi+\pi s}$. We define $S_{\text{net}}(p) \equiv S(p - \tau(p))$, so that

$$S'_{\text{net}}(p^*) = S'(p^* - \tau^*)\left(1 - \frac{\mathrm{d}\tau(p^*)}{\mathrm{d}p}\right) = \frac{(1-\pi)s}{(1-\pi)+\pi s}. \tag{7}$$

For the full definition of competitive equilibrium, see the Online Appendix.

We consider situations in which the policy maker is prevented from implementing the socially optimal outcome, so the equilibrium quantity remains too high. This scenario could arise from prohibitive monitoring, enforcement, or political costs linked to high permit fees. It is also the scenario under which the contribution of socially responsible consumers to mitigation is most crucial. We think of a policy type as superior to another if it can achieve a lower equilibrium externality level $g^* = q^*$ with the same degree of intervention. More precisely, policy type A (e.g., a tax) is strictly better than policy type B (e.g., a cap) if for any $q_B^*, \tau^* > 0$ with $u'(0) > k + \tau^*$ that is part of an equilibrium with a B-type policy, there is an A-type policy and a corresponding equilibrium with permit fee τ^* and quantity $q_A^* < q_B^*$.[9] By market clearing for the product ($q_i^* = S(p_i^* - \tau^*)$ for $i = A, B$), the A-type

9. We impose the condition $u'(0) > k + \tau^*$ to guarantee that with both types of policies, an equilibrium (defined to feature positive consumption) exists with fee τ^*.

policy also leads to a lower consumer price ($p^*_A < p^*_B$). This makes it extra preferable if we deem consumer surplus more important than producer surplus.

While different policies can achieve the same outcomes with selfish consumers,[10] we obtain:

PROPOSITION 4 (Permit-Supply Policies). More responsive permit-supply policies are superior: policies with parameter $\pi \in [0, 1)$ are strictly better than policies with parameter $\pi' > \pi$.

The intuition follows from a broad "dampening principle": policies that induce lower dampening provide a greater motivation for consumers to mitigate and hence yield better outcomes. Indeed, by equation (7) a more elastic supply of permits translates into a higher price responsiveness of product supply, which by equation (3) leads to lower dampening. An implication is that a tax ($\pi = 0$) is the best permit-supply policy, while a cap ($\pi = 1$) is the worst. Replicating Herweg and Schmidt's (2022) main insight, therefore, a tax is better than a cap. We now use the flexibility of our framework to study other policy questions.

IV.B. Other Policy Examples

1. *Improving on Taxes.* We show that a regulator can do better than with a unit tax.[11] A conceptually interesting example is a fixed-price policy: the planner fixes the consumer price p^*, and chooses the fee τ^* to clear the market. Because dampening operates through changes in the price, it now equals zero. Hence, by the above dampening principle:

OBSERVATION 3. A fixed-price policy is strictly better than a tax.

Although a fixed-price policy is not realistic in practice, its logic suggests more plausible alternatives. We analyze taxes that respond to changes in the consumer price linearly: $\tau = \tau_0 + \tau_1 p$. The case $\tau_1 = 0$ corresponds to a unit tax we have considered above. We get:

10. To see this, consider any demand curve $D(p)$ and supply curve $S(p)$ satisfying the conditions in Section II. In equilibrium $q^* = S(p^* - \tau^*) = D(p^*)$. Hence, there is a unique pair p^*, q^* that is consistent with a given τ^*.

11. Adapting the definition of competitive equilibrium to the settings in this subsection is straightforward and thus omitted.

PROPOSITION 5. Taxes that are decreasing in the consumer price ($\tau_1 < 0$) are better than fixed taxes ($\tau_1 = 0$), which are better than taxes that are increasing in the consumer price ($\tau_1 > 0$).

Although most existing tax regimes feature constant or increasing taxes, a tax that increases when the consumer price decreases is better at motivating socially responsible consumers. Such a tax can also be a reduced-form representation of a dynamic policy rule for emissions markets in which the number of future permits is lowered in response to a low consumer price on externality-generating goods. Like a tax increase, the withdrawal of permits puts an upward pressure on the fee. The intuition can again be understood from the dampening principle. If a consumer cuts her consumption, the resulting decrease in the price raises the tax. This attenuates the price drop, so the response of other consumers is lower. Hence, dampening is weaker than under a fixed tax.

2. *International Trade.* We compare cap and tax policies when there is trade, and the policy applies only at home. We assume that there is a single product with both home and foreign producers and only home consumers. A consumer does not observe her purchase's source, so there is a single consumer price p that applies to all purchases.[12] The foreign supply curve, which is a function of p because there is free trade and foreign producers are not subject to regulation, equals $S^f(p) = s^f p$ with $s^f > 0$. The home supply curve, which instead is a function of $p - \tau$ because home producers must pay the permit fee τ to supply a unit of the good, equals $S^h(p - \tau) = s^h(p - \tau)$ with $s^h > 0$. Market activities generate a total externality of $g = e^h q^h + e^f q^f$, where q^h and q^f are the quantities produced at home and abroad, respectively, and $e^f > e^h > 0$. This means that foreign suppliers are more polluting, for instance, because their technology is different due to the laxer regulation.[13] For simplicity in stating our results, we assume that

12. The results are identical if consumers can distinguish but are in equilibrium indifferent between home and foreign-sourced purchases. Analogously to Section V, such an equilibrium always exists.

13. Our formulation assumes that the consumer cares equally about home and foreign pollution. This should hold true for externalities with worldwide effects, such as global warming, but may not hold true for externalities with more localized effects, such as air pollution. We can capture the latter case by assuming that e_f denotes the relative weight the consumer puts on pollution from foreign

u is quadratic, and denote its second derivative by u_{cc}. As our proof establishes, this implies that there is a unique equilibrium. Then we say that policy type A (a cap or tax) is strictly better than policy type B (a tax or cap) if for any equilibrium g^*_B, τ^* with $u'(S^f(\tau^*)) > \tau^* + ke^f$ that obtains under a B-type policy, there is an A-type policy that in equilibrium generates the same permit fee τ^* and an externality $g^*_A < g^*_B$.[14]

PROPOSITION 6 (Cap versus Tax under Trade).

i. If

$$\frac{-1}{u_{cc}} < \frac{e^f - e^h}{e^h} \cdot s^f, \tag{8}$$

then a cap is strictly better than a tax.

ii. If inequality (8) goes strictly the other way, then a tax is strictly better than a cap.

Unlike in a closed economy, in an open economy a cap does not keep production constant. In particular, when a consumer reduces her consumption under a cap, her effect on home production is fully dampened, but her effect on foreign production is not. When the consumer reduces her consumption under a tax, the dampening effect on total production is lower, but this stems from lower dampening in clean home production at the expense of higher dampening in dirty foreign production. As a result, the consumer may have a greater effect on total pollution, and hence she may be more motivated to mitigate, under a cap. This happens if foreign supply is sufficiently dirty or sufficiently responsive relative to demand. The latter is especially likely to be the case when foreign supply dominates production in the market.

The logic of Proposition 6 clarifies a misperception regarding a consumer's or even country's impact on emissions. We have often heard that for consumption under a cap-and-trade system, for instance, buying steel in the EU, the environmental impact is zero because emissions will be at the cap anyway. As has already been understood at least for large buyers under the rubric of carbon leakage, this argument is incorrect with international trade. Nevertheless, there is a continuing policy debate regarding

production. Then, e_f may be lower than e_h even if foreign production is more polluting.

14. We impose the condition on u' to guarantee that under both policies, home production remains positive.

the argument (e.g., Intergovernmental Panel on Climate Change 2022, 1396). Our analysis says that in case i, the argument is incorrect in a major way: it is exactly because of the cap that a consumer's impact on pollution is large. Furthermore, the logic extends to non-traded products, such as flying within the EU, covered by the same cap. This is because a decrease in the permit fee due to the reduction in demand for the non-traded good lowers the consumer prices of traded goods, leading foreign suppliers to decrease production.[15]

V. Substitute Products

In our basic model, there is a single market. We now analyze effects acting across markets for substitute products. To study the scope for mitigation through product selection rather than quantity reduction, we first impose unit demand, that is, that every consumer purchases exactly one unit. Later, we reintroduce quantity choice into the picture.

V.A. Setup: Unit Demand

We assume that there are two products, a clean one and a dirty one. They are perfectly interchangeable in consumption utility, but generate different externalities denoted by $e^c \geq 0$ and $e^d > e^c$, respectively. For example, a consumer can power her appliances equally well with green and brown electricity, but the former is more environmentally friendly to produce. Letting the market quantities of the two products be q^c and q^d, the total externality is $g = e^c q^c + e^d q^d$. Consumers have unit demand, and are heterogeneous in their social concerns k, with k distributed on

15. Our analysis assumes completely free trade. But the European Union recently enacted the Carbon Border Adjustment Mechanism (or CBAM, colloquially known as a carbon tariff). Under this system, an importer wishing to sell in the domestic market must buy a "CBAM certificate" for the pollution caused abroad, paying a price equal to that of a domestic permit. If the system was symmetric, exporters serving the foreign market would not have to buy permits. With such a symmetric CBAM, a cap fixes domestic consumption rather than domestic production, so that a consumer does not have an effect on total pollution. As with a closed economy, therefore, a tax is always superior to a cap. The EU's system, however, is not symmetric: exporters do have to buy permits. In this case, a cap may again be superior. Intuitively, if a consumer reduces her consumption under a cap, she lowers the permit price and thereby lowers clean exporters' competitive disadvantage over dirty foreign producers.

$[\underline{k}, \overline{k}] \subset [0, K]$ according to the continuous, positive-valued density function f.[16] Suppliers provide the clean and dirty products according to the exogenous supply curves $S^c(p^c) = s^c p^c$ and $S^d(p^d) = s^d p^d$, respectively, where p^c, $p^d \geq 0$ are the product prices and s^c, $s^d > 0$. Typically, one would expect $s^c > s^d$ because the clean good is more difficult to produce, but we do not impose this.

We focus on situations in which both markets are active, and follow the steps preceding Definition 1 to define equilibrium. We describe the logic here, but relegate the formal treatment to the Online Appendix. In step i, we introduce a market responsiveness dQ that determines how much the equilibrium quantities respond if a consumer moves her consumption from one market to the other. In step ii, we identify an optimality condition for demand given dQ. Without loss of generality, we can think of demand in terms of an indifferent consumer type $k^* \in (\underline{k}, \overline{k})$ such that those with $k > k^*$ choose the clean product and those with $k < k^*$ choose the dirty product. Since moving one's demand from the dirty to the clean market raises the price paid from p^d to p^c but lowers the externality by $\mathrm{d}Q \cdot (e^d - e^c)$, the cutoff k^* must satisfy $p^c - p^d = k^* \cdot \mathrm{d}Q \cdot (e^d - e^c)$. In step iii, we express dQ as a function of the price responsivenesses of supply and demand. Again, we apply equation (2), this time to supply and demand curves expressed as functions of the price difference $p^c - p^d$ rather than a single price p. Finally, in step iv we derive the responsiveness of demand to $p^c - p^d$ from the above consumer indifference condition combined with the distribution of k. Here, we impose that dQ does not change in response to infinitesimal price changes.

To evaluate outcomes, we define social welfare as the negative of suppliers' production costs, minus K times the externality g. Since consumers have unit demand and the products are interchangeable in consumption, consumption utility is always the same and thus can be ignored when making welfare comparisons.

16. Assuming that k is continuously distributed facilitates our definition of competitive equilibrium, as it allows us to define an indifferent type k^* that separates consumers who buy the clean good from consumers who buy the dirty good. But the forces we identify do not depend on this assumption. For instance, a population with homogeneous k can be approximated arbitrarily closely by a uniform distribution with high density; then the selfish equilibrium we identify below is the unique equilibrium. Alternatively, our definition extends to distributions with atoms with minor adjustments. Similarly, the key possibilities we identify can also occur when f equals zero at $\underline{k}$ and $\overline{k}$, but the conditions for them are then more difficult to state.

V.B. Selfish Equilibria with Unselfish Consumers

Proposition 7 characterizes the key features of competitive equilibria.

PROPOSITION 7 (Substitute Products).

i. There is a competitive equilibrium in which the two products have the same price ($p^{c^*} = p^{d^*}$), and all consumers are indifferent between them.
ii. If $e^d - e^c$ or s^c is sufficiently small, then there is no other competitive equilibrium.
iii. Suppose $\underline{k} = 0$. If $e^d - e^c$ is sufficiently large, then there is a competitive equilibrium in which the clean product is more expensive ($p^{c^*} > p^{d^*}$), yet some consumers strictly prefer it.
iv. Among multiple competitive equilibria, the greater is the clean product's price premium ($p^{c^*} - p^{d^*}$), the greater is social welfare.

Part i says that there is always an equilibrium in which the two products sell at the same price, and consumers are indifferent between them. This outcome is identical to that when all consumers are selfish. To see the intuition, suppose that a consumer expects everyone else to behave selfishly, that is, to always choose the cheaper product. Then, the consumer believes that the two prices always equalize, in particular at the price p^* where total supply $S^c(p^*) + S^d(p^*)$ equals one. The consumer concludes that the quantities in the two markets, and therefore total pollution, do not depend on which product she chooses. The same conclusion can be seen from the fact that with others behaving selfishly and the products being interchangeable in consumption, the price elasticity of demand is infinite, so dampening is full. Consistent with equilibrium, therefore, the consumer chooses the cheaper product herself.

Worse, part ii implies that if the products are not too different in the externalities they generate ($e^d - e^c$ is sufficiently small), then the selfish equilibrium is the only equilibrium. The intuition derives from a mutually reinforcing interaction between dampening and consumers' price sensitivity. Suppose, to start, that consumers expect dampening to be zero, and consider the social concern k that makes a consumer indifferent between the two products. Since the products generate similar externalities, a small change in their price difference implies a large change in this cutoff k. Hence, the price sensitivity of demand is quite high, so dampening must be nontrivial. Taking dampening into account, then, purchasing the two products generates even more similar externalities. As a result, a small price change implies an even greater change in the cutoff k. In this fashion, the equilibrium unravels to the point where dampening is full.

A similar unraveling logic holds when s^c is low, but in that case the first step acts in part through supply. Analogously to equation (3), a low s^c implies that when a consumer moves her demand to the clean sector, it is mostly other consumers rather than suppliers who respond to the price increase. This means that dampening is nontrivial, kickstarting the unraveling.

Part iii states that if the products generate sufficiently different externalities, at least one other equilibrium exists. In such a nonselfish equilibrium, the clean product is more expensive, but dampening is not full, so consumers with a sufficiently high concern k are willing to pay the higher price to mitigate.[17] Part iv says that this results in higher social welfare.

The existence and potential uniqueness of the selfish equilibrium represents a more drastic market failure than that in our single-good model. There, market failure is partial: dampening erodes consumers' social motives, but the outcome is better than with selfish consumers. Here, market failure is complete: although each consumer is willing to pay to mitigate the externality stemming from her own consumption, the equilibrium is identical to that when all consumers are selfish. Because consumers' social preferences are not reflected in their behavior or aggregate outcomes, the market does not serve its role of aggregating information at all.[18]

17. To ensure the existence of an equilibrium, which we have defined for situations when both markets are active, Part iii imposes that $\underline{k} = 0$. If $\underline{k} > 0$ and $e^d - e^c$ is sufficiently large, then under our definition an equilibrium does not exist. Under natural generalizations, an equilibrium with only the clean market being active exists in that case (but not for small $e^d - e^c$ or s^c). For simplicity, we do not analyze such clean equilibria in the current article.

18. The selfish equilibrium is superficially related to what happens in previous models where consumers with social preferences act selfishly in a competitive equilibrium (Dubey and Shubik 1985; Dufwenberg et al. 2011; Arnold 2023). But in previous theories, this requires that consumers are not socially responsible by our definition. Similarly, Fehr and Schmidt's (1999) inequity-averse agents act selfishly in "proposer competition" because they are in a disadvantageous position where they are unwilling to sacrifice for others.

For situations in which the selfish equilibrium is played, two more observations follow. First, since socially responsible consumers do not induce a price premium for the clean product, they provide no incentive to develop cleaner technologies. Second, when observers see that the clean product is no more successful than the dirty one, they may naturally conclude that consumers are selfish. They may then, for example, underestimate support for policies to mitigate externalities.

Furthermore, these observations about selfish equilibria apply in weaker form to non-selfish equilibria as well. Even in such equilibria, dampening implies that individuals' choice between the products provides a lower bound on how much they care about the externality. Hence, the market partially fails in aggregating consumers' social preferences. This reduces firms' incentive to innovate and means that observers may underestimate consumers' social concerns.

Our model in this section is related to the model of consumer boycotts by Broccardo, Hart, and Zingales (2022). In both settings, socially responsible consumers choose between a dirty product and a clean product, and the main question is how the two products fare. In Broccardo, Hart, and Zingales's model, however, the clean product's price premium is fixed by the positive cost firms must pay to be clean. This means that there are no analogues to our main results, which depend on how demand affects the premium. In particular, there is no analogue to our finding that a socially responsible population may in the unique equilibrium not favor a cleaner product selling at a zero premium.

V.C. Quantity Choice and Cross-Market Effects

We now briefly outline what happens when a consumer chooses not only the product but also the quantity she purchases. As in Section II, we let the utility function over the quantity be $u(\cdot)$, and for simplicity impose that consumers are homogeneous in their social concern k. We focus on the analogue of the selfish equilibrium, in which the products sell at the same price and consumers are indifferent between them. This implies that we can think of equilibrium as featuring a single price and of consumers as choosing a single quantity. Generalizing Definition 1, then, we specify an equilibrium as a price p^*, consumption level q^*, consumer price responsiveness q_p^*, and clean and dirty market responsivenesses q_c^{c*} and q_c^{d*} that satisfy

$$u'(q^*) = p^* + k\left(q_c^{c*}e^c + q_c^{d*}e^d\right), \quad q_c^{c*} = \frac{s^c}{s^c + s^d - q_p^*},$$

$$q_c^{d*} = \frac{s^d}{s^c + s^d - q_p^*}, \quad \text{and} \quad q_p^* = \frac{1}{u''(q^*)}. \tag{9}$$

The logic of this equilibrium derives from a combination of the forces we have seen. First, full "substitution dampening" similar to that in the selfish equilibrium holds: because prices always equalize at the market-clearing level, outcomes are unchanged if a consumer moves a given quantity from one market to the other. This means that the consumer is indeed indifferent between the products and behaves as if she was buying a composite good whose externality is a weighted sum of the externality levels e^c and e^d. Second, on this composite good, "quantity dampening" similar to that in Section II.A holds: if the consumer raises her consumption, the (uniform) price rises, leading others to consume less. The consumer therefore consumes less than a selfish consumer, but dampening lowers her motive to mitigate.

The above equilibrium identifies an economically important new consideration. In a single-good setting, consuming more of a product causes an externality that is at most the direct externality of the product itself. In the two-good setting, in contrast, consuming the clean product also raises production of the dirty product, so it can cause a greater externality than the clean product's direct impact. For instance consuming the clean product is harmful even if $e^c = 0$. Consumers choosing the clean product recognize this cross-market effect and lower their consumption in response.

While our model assumes that the two products are perfectly interchangeable in terms of consumption utility, consumers often perceive a difference between a clean product and a substitute dirty product. For instance, a consumer may consider organic food as healthier or more nutritious than nonorganic food, and used clothes as less stylish than new clothes. Our model easily extends to such situations. Suppose that one unit of the clean good provides the same consumption utility as $v > 0$ units of the dirty good. Then the analogue of the selfish equilibrium still exists, but now prices of the products are equal after normalizing by v: $p^{c^*} = vp^{d^*}$.

VI. Evidence on Consumers' Beliefs and Social Concerns

In our benchmark framework, consumers anticipate dampening and evaluate options based on their dampened effects. But other assumptions about socially responsible consumers are also plausible. First, a consumer may be naive in that she fails to recognize the dampening effect of markets. Second, a consumer may have deontological or warm-glow preferences in that she cares about her action or direct effect rather than her dampened effect. We explore the empirical relevance of these possibilities via a preregistered survey.[19]

VI.A. Sample

We conducted the survey with 2,000 US consumers in October 2023 using the online survey company Prolific. The sample approximates the adult US population in terms of gender, age, income, and region, but overrepresents college-educated and Democratic consumers and underrepresents Hispanic consumers. All of our results are robust to reweighting and correcting for these imbalances. Online Appendix C.1 presents further details on the sample.

VI.B. Design

1. *Beliefs about Aggregate Impact of Own Consumption.* We measure beliefs about one's impact by asking respondents to estimate the change in total global consumption that would result from a change in their personal consumption. Our questions concern eight practically relevant markets. Four ask about quantity dampening—the implications of reducing consumption of fuel by 200 gallons, of meat by 100 pounds, of flights by eight trips, or of energy by 10,000 kWh. In turn, four questions ask about substitution dampening—the implications of reallocating consumption from brown to green electricity (10,000 kWh), from new to secondhand clothing (40 garments), from energy-inefficient to energy-efficient housing (10,000 kWh), or from conventional to fair-trade coffee (10 pounds).

Each respondent is randomly assigned to one of the eight cases and is first presented with a short introduction to the

19. The preregistration is available at www.doi.org/10.17605/osf.io/btz5p. We received ethics approval from the German Association for Experimental Economic Research (no. es3dPMfa).

context. Next the respondent is asked to consider two scenarios that differ in her personal consumption level and estimate how the yearly total global consumption of the good would be affected.[20] For example, the instructions for fuel consumption read:

> Your consumption of fuel is part of the **total global consumption of fuel**. We would like to know what you think would happen to the global consumption of fuel if you reduced your own consumption of fuel. Would it make a difference to the total consumption of fuel worldwide?
>
> Consider these two scenarios:
>
> Scenario 1: You consume 400 gallons of fuel every year.
> Scenario 2: You consume 200 gallons of fuel every year.
>
> In contrast to scenario 1, you would permanently reduce your yearly fuel consumption by 200 gallons in scenario 2.

The respondent indicates whether she thinks global consumption would (i) decrease by more than 200 gallons, (ii) decrease by 200 gallons, (iii) decrease by less than 200 gallons, (iv) not change at all, or (v) actually increase.[21] Finally, the respondent explains her prediction in an open-ended text box. The open-ended format allows respondents to express themselves freely and unconstrained by the researcher, providing a lens into their reasoning (Andre et al. 2022; Ferrario and Stantcheva 2022). See Online Appendix C.6 for the full instructions of the survey.

2. *Nature of Social Concerns.* We measure the relevance of consequentialist concerns by comparing consumers' val-

20. We obtain even stronger evidence for beliefs in dampening if we alternatively ask for the yearly total global *production* of the good (Online Appendix C.5.1).

21. These categorical response options focus on the differences that are most interesting from the theoretical perspective. The contrast they create also facilitates the subsequent measurement of open-ended explanations. We obtain similar results in a robustness study with an unrestricted numeric response field (Online Appendix C.5.2).

Our scenario-based approach induces controlled variation but renders it difficult to incentivize respondents' predictions in our main study. However, studies often find at most weak differences in the answers to incentivized and non-incentivized questions (Stantcheva 2023). Consistent with this, we replicate our results in a robustness study with a probabilistic incentivization approach (Online Appendix C.5.3).

uations for effective versus ineffective mitigation actions. Respondents are randomly assigned to one of four practically relevant externalities—CO_2 emissions, nonrecyclable waste, animal welfare, and low wages for workers—and are asked how much they would be willing to pay to reduce their own contribution to the externality in two scenarios. In the first scenario, reducing their contribution is effective and translates one-to-one into a reduction of the total global level of the externality. In the second scenario, reducing their contribution has no effect on the total global level of the externality. We provide our dampening mechanism as one potential reason one's reduction could be ineffective. Respondents indicate their valuations for reducing the externality in the two scenarios and then explain their responses in an open-ended text box.[22]

VI.C. Results

1. *Beliefs about Aggregate Impact of Own Consumption.* Figure I displays the distribution of respondents' beliefs about their impact. Across scenarios, 29% to 63% of consumers predict that their own consumption reduction has a dampened effect on aggregate consumption. On average, the share of consumers who expect to have a less than one-to-one effect is 38%, with 43% doing so for reductions of consumption and 34% doing so for reallocations of consumption. Among these consumers, most (69% for reductions of consumption and 59% for reallocations of consumption) expect to have a zero effect. Using a follow-up question, we find that the average perceived degree of dampening ranges from 0.19 to 0.55 across scenarios, with an overall

22. Our focus on four real-world settings helps us measure concerns in contexts that matter for our model, which is important because the strength of consequentialist behavior can vary across settings (Awad et al. 2020; Bénabou, Falk, and Henkel 2022; Hart, Thesmar, and Zingales 2024). However, it implies that we cannot incentivize responses. Reassuringly, Bénabou, Falk, and Henkel (2022) document that the strength of consequentialist behavior does not differ between incentivized and non-incentivized choices. We measure consumers' willingness to pay more money rather than their willingness to change consumption because consumers can be hesitant to change their consumption even if they care about impact. For example, they might not eat more conventional chicken in a scenario where doing so has no effect because they would not derive sufficient consumption utility from doing so. This problem is lessened for consumers' willingness to pay.

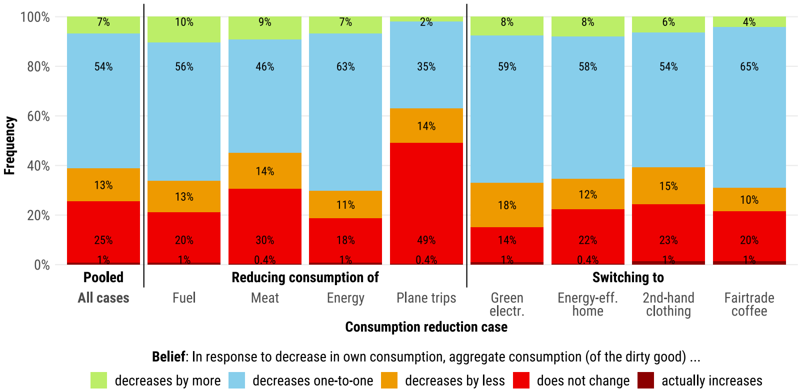

FIGURE I

Consumers' Beliefs about the Aggregate Impact of Their Own Consumption

This figure displays the distributions of consumers' beliefs about their own impact on aggregate consumption. The first column displays results pooled across all eight cases, the other columns present the results for the eight cases.

average of 0.28.[23] The average consumer's belief that she has a dampened effect, and many consumers' belief that they have a zero effect, are consistent with the beliefs rational consumers hold in our models.

At the same time, an average of 54% of consumers, ranging from 35% to 65% across scenarios, believe that their own consumption changes translate one-to-one to aggregate consumption changes. We also find a 7% share of consumers who believe in a multiplier effect, that is, that aggregate consumption falls by more than their own consumption. Fewer than 1% of consumers believe that aggregate consumption actually increases when their own consumption decreases. These three types of beliefs are largely inconsistent with our rational models.

The open-ended text data allow us to shed light on how consumers reason about their predictions. We manually classify each response into one of three categories: (i) explanations for a dampened effect, (ii) explanations for a one-to-one effect, and (iii)

23. To approximate the perceived dampening factor of consumers who respond that aggregate consumption decreases "by more," "by less," or "actually increases," a follow-up question asks consumers to narrow down their prediction.

TABLE I

REASONING ABOUT DAMPENING: EXAMPLE RESPONSES

Dampened effect (25%)	**One-to-one effect (50%)**
Own effect is offset by others (38% of cases)	"It makes sense that if I use less it means the total used will be less."
"Someone else would buy the meat I didn't buy and eat it."	"I would think that my personal consumption would be directly related to the global consumption, gallon for gallon."
"My switch to used clothing would mean less demand for new clothing, albeit a very small 'less'. This would cause a (very small) reduction in price. This would cause a (very small) increase in consumption."	"I am part of the total, so if I use 10,000 less, the total will be 10,000 less."
Too small to matter (62% of cases)	**Positive multiplier (3%)**
"I'm only one person, and any changes I make would not show up in the big picture."	"I think if I reduced my own consumption of fuel, I would likely mention that to a number of people. A few of those people are likely to be at least a little influenced by my new habit and feel inspired by it.…
"The shopping habits of one individual are not going to affect global production."	

Notes. Examples of open-ended explanations from the consumer survey. Consumers make a prediction and are then asked, "Please explain why you chose this response."

explanations for a multiplier effect. Seventy-eight percent of the responses can be classified.[24]

Online Appendix Figure C.1 summarizes the results, and Table I shows a few example responses for illustration. A significant share of consumers—25%—explain why they have a dampened effect on markets. Of these, 38% argue that their reduced consumption will be offset by others' consumption, with a few identifying exactly our price-based mechanism. The others (62%) argue that being minuscule, they have little to no influence on aggregate consumption. Such statements could be incomplete explanations of a fuller mechanism; for example, that producers will not notice a small consumer's change in habits, and hence will produce and sell the same amount. The statements might

24. Each response is coded independently by two research assistants who have been extensively trained for this task. We cross-verify each classification. Reassuringly, the interrater reliability is high: the two coders agree in 87% of all cases. Online Appendix C.4 describes the coding procedure in detail. As is common in research with qualitative text data, a subset of the responses (22%) cannot unambiguously be classified or do not contain a clear explanation.

also reflect a heuristic view—not tied to a precise mechanism—that only large entities can make a difference in markets. But as we emphasized earlier, a consumer's behavior depends on her perception of her aggregate impact, not on her perception of the underlying mechanism. Hence, the erosion of moral behavior due to dampening that is at the heart of our theory applies to all these consumers.[25]

A larger group of consumers—50%—argue that they have a direct, one-to-one impact on aggregate consumption. For them, the case is often a straightforward matter of math—that their own consumption simply adds to others' consumption. A much smaller share of consumers—3%—argue for a positive multiplier, typically by referring to spillover effects on others.

We present a series of additional analyses in Online Appendix C.2. We find that young and Republican-leaning consumers are most likely to voice dampening beliefs and explanations.[26] We also confirm the robustness of our results in multiple sensitivity analyses, including using post-stratification weights, dropping respondents who "speed" through the survey, and focusing on consumers who consume the relevant goods and hence can actually reduce consumption. We summarize:

EMPIRICAL RESULT 1 (Beliefs about Impact). Consumers' beliefs are heterogeneous. A sizable share believe that they have a dampened effect on markets, and so does the average consumer. But approximately half believe that they have a one-to-one effect on aggregate consumption.

2. *Nature of Social Concerns.* Next we turn to the nature of consumers' social concerns. In line with a large literature

25. Beliefs about the mechanism could matter from an intervention perspective. For example, a consumer who believes she has no impact simply because she is "too small to matter" could potentially be persuaded that in expectation, she still has a positive, though dampened effect. Likewise, consumers who do not understand the dampening mechanism might be slower to respond to differences in dampening across policy regimes, though they might eventually catch up by following the public debate. These matters are beyond the scope of this article, but promise to be interesting avenues for future research.

26. One potential explanation for why Republicans and Democrats have different views is motivated reasoning. Skepticism about climate policies and private mitigation efforts would favor a worldview that rates these steps as ineffective. The reverse could apply to Democrats.

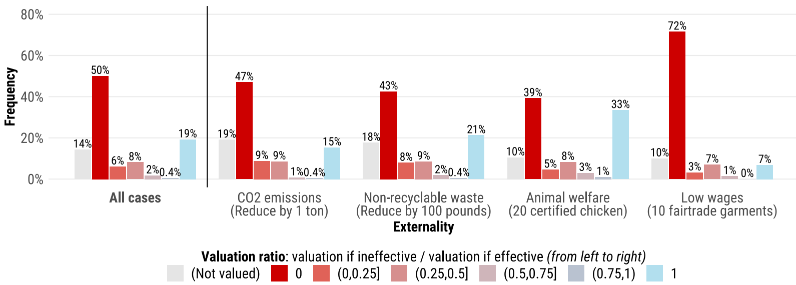

FIGURE II

Consumers' Concern for Consequences

This figure presents the distribution of the valuation ratio for each externality. The ratio is derived by dividing consumers' valuation of the ineffective externality reduction by their valuation of the effective externality reduction. The ratio is not defined (gray bars) for consumers who do not value an effective reduction in the externality. In the few cases (less than 2%) where consumers indicate a higher valuation for ineffective than effective mitigation, we set the ratio to one.

showing that individuals care about the externalities they cause (e.g., Auger et al. 2003; Hainmueller, Hiscox, and Sequeira 2015), most consumers—86%—report a positive willingness to pay to reduce global externalities. For example, the median valuation for a 1 ton reduction in personal CO_2 emissions that is effective in reducing global emissions is \$50, and the median valuation for an effective 100-pound reduction in nonrecyclable waste is \$20.

Our main interest, however, is in consumers' valuations for ineffective externality reduction relative to their valuations for effective externality reduction. Figure II presents the distribution of these "valuation ratios" across consumers. Ratios close to zero (red colors/darker gray; color version available online) mean that consumers care about their ultimate impact, consistent with a consequentialist concern for externalities. Ratios close to one (blue colors/lighter gray) mean that consumers value their own reduction irrespective of its net consequences, consistent with deontological or warm-glow preferences. The ratio is not defined (light gray bars) for consumers who do not value an effective reduction in the externality.

The figure shows that the majority of consumers care about the consequences of their actions. Many consumers—50%—only value effective reductions, placing \$0 value on ineffective reductions. A smaller group of consumers positively value ineffective reductions but strictly prefer effective reductions. Typically, their valuations for effective reductions are at least twice as high. Only 19% equally value effective and ineffective reductions. The figure also illustrates that the nature of consumers' concerns can vary with the externality at stake. Yet consequentialism remains the dominant motive in all four cases.

The qualitative text data on consumers' reasoning confirm this pattern. Sixty-nine percent of consumers who report a positive valuation for effective actions voice only consequentialist arguments (see Online Appendix Figure C.3).[27] In fact, many consumers still refer to consequences when they justify a positive valuation for an ineffective reduction, offering various arguments for why their action may ultimately still have a positive impact. By contrast, arguments that focus on the action and deliberately ignore the consequences are much rarer (10%). Here, consumers refer to their personal responsibility, moral principles, or the desire to feel good about their own behavior. An additional 9% of explanations mention both consequence- and action-based arguments. Table II gives a few example responses for illustration.

Interestingly, we observe some variation in consumers' reasoning across the different types of externalities. For example, we detect more strict deontologist arguments for animal welfare (16%) and nonrecyclable waste (16%) than for carbon emissions (7%) and worker wages (0.5%). One plausible reason is that in the former cases, there is more direct contact with the externality one causes. This suggests that while consequentialist concerns dominate in the settings we study, deontologist motives could play an important role in other settings and situations with more personal involvement.

We present a series of additional analyses in Online Appendix C.3. We find that consequentialist concerns are most common among younger and politically independent consumers. Our results replicate in a series of sensitivity analyses. Importantly, we also find that consequentialism is common among consumers who believe that they have a dampened effect. We summarize:

27. As before, each response is coded independently by two research assistants who have been extensively trained for this task. The inter-rater reliability is high: the two coders agree in 84% of all cases. Online Appendix C.4 describes the coding procedure in detail. Twelve percent of responses cannot unambiguously be classified or do not contain a clear explanation.

TABLE II

REASONING ABOUT CONSEQUENCES: EXAMPLE RESPONSES

Consequentialist arguments	**Deontologist / warm-glow arguments**
"If it makes no difference, then there is no point in spending money on it."	"I would feel better about myself if I at least gave it some effort."
"If I could help workers, I would definitely be willing to pay more. However, if there was no effect, why would I pay more?"	"It would not have a positive effect, but from a moral standpoint it would have some significance to me. I'd be willing to pay for that."
"I wouldn't want to pay any money if it does not have any global or 'big picture' impact."	"I do not want to be associated with clothing companies that have bad business practices regarding worker wages."

Notes. Examples of open-ended explanations from the consumer survey. Consumers indicate their valuations and, depending on their responses, are then asked "Please explain why you gave different answers in the two situations," "Please explain why you gave the same answer in the two situations," and/or "Please explain why you would be willing to pay money in situation 2 where the total impact would be zero."

EMPIRICAL RESULT 2 (Nature of Social Concerns). Consumers often care about the externalities they cause in a consequentialist way.

We conclude that an empirically realistic model must take rational consequentialist consumers seriously. However, it is also important to understand the impact of naive or deontologically minded consumers, especially when they coexist with rational consumers in the population. This conclusion motivates the next section.

VII. IMPLICATIONS OF NAIVE OR DEONTOLOGICAL CONSUMERS

In this section, we identify ways in which the presence of naive consequentialist or deontological consumers modifies our insights. Our starting assumption is that these two types maximize the same utility function as a rational consequentialist consumer, except that they evaluate an action based only on its direct consequences. A deontological consumer does so because the action or its direct effect is what she cares about, and a naive consequentialist consumer does so because she fails to anticipate dampening. Although the two types therefore behave identically, we mostly describe the results in terms of naive consumers, who we found are more numerous.

VII.A Single-Good Setting

Consider first the single-good model of Section II, assuming for simplicity that competitive equilibrium is unique. Since a naive consumer maximizes $u(c) - p^*c - kc$, she solves $u'(c) = p^* + k$. Recalling equation (4), a rational consumer solves $u'(c) = p^* + k \cdot q_c^*$, where $q_c^* < 1$. Hence, the naive consumer chooses a lower c than the rational consumer. Intuitively, a person who ignores dampening is overoptimistic about her ability to reduce output, which encourages more responsible behavior. Fixing consumers' social concern k, therefore, welfare is increasing in the share of naive consumers. Furthermore, a population of naive, fully socially responsible consumers ($k = K$) achieves the social optimum. These consumers solve $u'(c) = p^* + K$, so their social concern acts akin to a tax equal to the social cost of the externality, that is, an optimal Pigouvian tax. Summarizing:

OBSERVATION 4. Let q^R be the equilibrium quantity when all consumers are rational, q^N the quantity when all consumers are naive, and q^{FB} the unique socially optimal quantity. Then, for any $k \leq K$, we have $q^{FB} \leq q^N < q^R$, with $q^{FB} = q^N$ if and only if $k = K$.

VII.B Two-Good Setting

Unlike with a single good, in a multi-good context naive consumers can underestimate their impact and hence behave less responsibly than rational consumers. To illustrate one source of such underestimation, we use the model of Section V.C with the following simplifications. The clean good is not polluting ($e^c = 0$), and the pollution from the dirty good is normalized to 1 ($e^d = 1$). Furthermore, the dirty good has perfectly elastic supply at price P^d, while the clean good has fixed supply S^c satisfying $u'(S^c) \geq P^d + k$.[28] A share $\alpha \in [0, 1]$ of consumers is naive, perceiving the externalities from the clean and dirty goods to be 0 and 1, respectively. We look for equilibria in which rational consumers assume that substitution dampening is full.

28. This condition ensures that the clean product cannot supply the entire market.

As a potential example, the two markets could be used and new goods. Used goods are in fixed supply, but producers can readily supply more new goods. In another example, the clean good could be taking the train, and the dirty good could be driving.

PROPOSITION 8 (The Effects of Naiveté). There are $\underline{\alpha}, \overline{\alpha} \in (0, 1)$ with $\underline{\alpha} < \overline{\alpha}$ such that the competitive equilibrium has the following properties.

i. For $\alpha \in [0, \underline{\alpha})$, $p^{c^*} = P^d$, rational consumers are indifferent between the goods, and naive consumers strictly prefer the clean good. Social welfare is strictly decreasing in α.
ii. For $\alpha \in (\underline{\alpha}, \overline{\alpha})$, $p^{c^*} > P^d$, rational consumers strictly prefer the dirty good, and naive consumers strictly prefer the clean good. Social welfare is strictly increasing in α.
iii. For $\alpha \in (\overline{\alpha}, 1]$, $p^{c^*} = P^d + k$, rational consumers strictly prefer the dirty good, and naive consumers are indifferent between the goods. Social welfare is constant and the same as for $\alpha = 0$.

Proposition 8 shows that naive consumers have a nonmonotonic effect on welfare, with intermediate shares generating the worst outcomes. For an intuition, imagine starting from a rational population ($\alpha = 0$), and gradually replacing it with naive consumers. When the share of naive consumers is low (part i), product prices are (as in the rational model) equalized by rational consumers' indifference between the products. Believing that they can act responsibly for free, therefore, naive consumers buy the clean good. Furthermore, because they fail to understand the cross-market effect—that clean consumption raises dirty output—they buy more of the clean good than rational consumers do, lowering social welfare. In the context of clothing a naive consumer fails to realize that if she buys used clothes, others will shift to buying new clothes. As a result, she buys too many used clothes.

When the share of naive consumers reaches a critical threshold ($\underline{\alpha}$), their demand for the clean good exhausts the fixed level of supply S^c. At this point, buying pressure from naive consumers starts raising the clean good's price p^c. If there are not that many naive consumers (part ii), the increase in p^c is not too large. This implies that naive consumers still strictly prefer the clean good. Since the rise in p^c lowers their consumption, social welfare increases.[29]

At some point ($\alpha = \overline{\alpha}$), the premium on the clean good reaches k, and then remains constant (part iii). Due to this premium, naive consumers internalize the cross-market effect, exactly counteracting their naiveté. Hence, they are indifferent between the products, and buy exactly as much as rational consumers do in the rational equilibrium (that with $\alpha = 0$). Rational consumers, who all choose the dirty good, do so as well. Total welfare is therefore the same as with a rational population. Interestingly, however, naive consumers do have a distributional effect: by inducing a premium for the clean product, they create a windfall for clean suppliers and a loss for themselves.

The role of two simplifying assumptions is worth mentioning. First, perfectly elastic supply for the dirty good ensures that quantity dampening is zero. If this is not the case, then—as in our basic model—a naive population ($\alpha = 1$) achieves higher social welfare than a rational population ($\alpha = 0$), although a mix is still worst. Second, homogeneity in k ensures that the price premium a population of naive consumers induces for the clean good exactly counteracts their naiveté. Under heterogeneity, the price premium, set by a consumer with intermediate k, is insufficient to counteract the naiveté of high-k consumers. In

29. In this parameter range, both consumer types have a strict preference for one of the products. Since the definition of substitution dampening we have used in Section V relies on the existence of an indifferent consumer, it does not apply here. We proceed by noting that if a rational consumer chooses the clean product, she is making an off-equilibrium choice. Our approach to dealing with this off-equilibrium situation is to allow for arbitrary beliefs about dampening. Proposition 8 assumes that rational consumers expect full dampening. In the spirit of sequential equilibrium, this can be microfounded by introducing an arbitrarily small share of consumers who have exogenously given demand that generates full dampening at the equilibrium prices. Alternatively, we can assume that rational consumers' social concerns are distributed (say) normally around k with a small variance. We conjecture that an equilibrium outcome approximating the one above will result, in which a small share of rational consumers choose the clean product. These consumers create sufficient substitution dampening for other rational consumers to choose the dirty product. Another sensible approach is to impose that if both consumer types have a strict preference, then dampening must be zero. Indeed, if a rational consumer chooses the clean product and thereby slightly raises its price, she does not induce any naive consumer to switch in the other direction. If we insist on such a strict approach to tying down beliefs about dampening in an off-equilibrium situation, then in this parameter range a competitive equilibrium does not exist.

this case, therefore, a rational population achieves higher social welfare than a naive population.

This second point is also relevant for understanding the role of naiveté when only a share of the population is socially responsible. Suppose that a share β of consumers is socially responsible with a homogeneous k, and the rest are fully selfish. For a sufficiently low β, naive consumers leave the fully selfish equilibrium unaffected: they choose the clean product and, thinking that they are not causing an externality, consume the same amount as a selfish consumer. Rational consumers are instead indifferent between the two products, and understanding their effect on the production of the dirty good, consume less than selfish consumers. When the share of socially responsible consumers is not sufficiently high, therefore, social welfare is higher if those consumers are rational.

These results indicate that naive or, equivalently, deontological consumers do not necessarily improve outcomes. More generally, it is clear that the sources of distortion depend on consumers' perceptions of their impact and the nature of their social concerns. How different types of consumers behave in other environments is therefore an important area for future research.

VIII. Conclusion

This article introduces a novel, portable framework for thinking about markets with socially responsible consumers. Unlike in many previous models of social preferences in markets, in our theory a consumer's social concerns are typically reflected in her behavior. This translates into a simple message for consumers wondering whether they can mitigate market-produced externalities: yes, you can, and if you care, you should modify your market behavior (being cognizant of cross-market effects). But because our theory predicts a market failure, it has a different message for policymakers: reliance on consumer responsibility cannot fully address problems due to externalities, even if consumers are extremely responsible. Hence, systemic solutions are necessary.

While we investigated some basic market settings, there are many other natural environments to which our framework can be applied. As an example, consumers can purchase offsets, or buy and delete permits, to lower the externality. How these options affect equilibrium outcomes and policy comparisons is an important question for future research. In addition, the logic of dampening is likely to be different when a consumer is facing firms with market power, especially if the firm's profits are used for harmful purposes (e.g., sponsoring an oppressive regime). Dampening might even be reversed in some dynamic settings, in which the consumer's actions have an effect on a firm's or policymaker's future choices. Finally, with appropriate modifications our framework also applies to financial markets. In particular, investments that have identical cash flows but different externalities are like our substitute products in Section V.

Our framework also suggests questions that are not about classical externalities but require similar principles to analyze. As a case in point, a consumer may care about price changes due to other social concerns, for instance, sensitivity to the distribution of income between rich and poor or local and nonlocal producers. We conjecture that for some natural classes of such motives, a small consumer will again violate price taking. To go further, similar forces can arise in the presence of aggregate factors other than market prices. Suppose, for instance, that individuals care about how their own action deviates from a social norm, which in turn is given by the population's average action. Again, we conjecture that in such situations, vanishingly small individuals violate "norm taking," that is, they think about their effect on the social norm. Unlike dampening, however, this effect on others is beneficial, and thus leads to more responsible behavior. To understand the precise implications of such motives, variants of our framework are necessary.

Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

Data Availability

The data and code underlying this article are available in the Harvard Dataverse, https://doi.org/10.7910/DVN/49CETN (Kaufmann, Andre, and Kőszegi 2024).

APPENDIX: PRIVATE UTILITY, SOCIAL CONCERNS, AND SOCIAL WELFARE

In this appendix, we motivate the consumer utility function (1) and the social-welfare function (5) we assumed in the text.

Suppose that there are $I+1$ identical consumers. We denote individual i's consumption by c_i, and the vector of consumption levels by $\boldsymbol{c} = (c_1, ..., c_{I+1})$. Individual i's private utility is $u(c_i) - pc_i - K \cdot A(\boldsymbol{c})$, where $A(\boldsymbol{c})$ is a harmful consequence of market trade that lowers the consumer's utility. The outcome $A(\boldsymbol{c})$ could represent air pollution or climate change, for instance, which harms each consumer through effects on her health or well-being. Consumer i realizes that her consumption contributes to the harmful consequence and thus the negative effect on others. She calculates that the total disutility others suffer from $A(\boldsymbol{c})$ is $I \cdot K \cdot A(\boldsymbol{c})$. Being socially responsible, she cares about this disutility, or her effect on it, with weight κ, adding a term $-\kappa \cdot I \cdot K \cdot A(\boldsymbol{c})$ to her utility. Because it is plausible to assume that the consumer internalizes the harm she causes others at most fully, we impose that $\kappa \leq 1$. Including her social concern, then, her total utility is $u(c_i) - pc_i - K \cdot A(\boldsymbol{c}) - k \cdot I \cdot A(\boldsymbol{c})$, where we have substituted $k \equiv \kappa \cdot K \leq K$.[30]

To conclude the setup, we suppose that the harmful consequence in question equals the average consumption in the population: $A(\boldsymbol{c}) = \left(\sum_{j=1}^{I+1} c_j\right)/(I+1)$. For large I, this captures a classical externality situation in which one individual consumer has a negligible impact on each person's private utility. Plugging this into consumer i's utility function gives

$$u(c_i) - pc_i - K \cdot \frac{\sum_{j=1}^{I+1} c_j}{I+1} - k \cdot \frac{I}{I+1} \cdot \sum_{j=1}^{I+1} c_j$$

$$= u(c_i) \; - \; pc_i - \left(k \cdot \frac{I}{I+1} + K \cdot \frac{1}{I+1}\right) \cdot \sum_{j=1}^{I+1} c_j.$$

30. To simplify matters, we assume that the consumer's utility is linear in both money and the harmful consequence A. This could lead to the nonexistence of a utility-maximizing choice if the consumer could exchange directly between the two, for example, if she had access to offsets. To analyze such situations, it is necessary to adjust the consumer's problem slightly, for instance by assuming nonlinear utility in money.

As $I \to \infty$, the weight the consumer attaches to the total market quantity $\sum_{j=1}^{I+1} c_j$ approaches k, as we have assumed in the text.

Our formulation assumes that in choosing her consumption, a socially responsible consumer cares about her effect on others' private disutilities from A, but not about her effect on others' private surpluses from consumption. Hence, for instance, in choosing how much to fly, the consumer thinks about climate change and its effect on humanity but does not internalize others' enjoyment of traveling. Beyond realism, this assumption is helpful in connecting our basic market-failure results to previous ones. In classical settings, the efficiency of markets does not require consumers to internalize others' private utilities from consumption. For simplicity, we also assume that the weight k does not depend on I. Even if k is not constant in I, our points remain unchanged as long as k converges to a strictly positive number as $I \to \infty$. A nontrivial k, in turn, follows from our definition of a socially responsible consumer—that she is willing to modify her consumption to mitigate the externality associated with it.

Crucially, we assume that consumer welfare equals the average of individuals' private utilities $u(c_i) - pc_i - K \cdot A(\boldsymbol{c})$. First-pass conventional logic might dictate that consumer welfare equals the average of individuals' total utilities, assigning a larger weight to $A(\boldsymbol{c})$. But the weight a consumer puts on $A(\boldsymbol{c})$ already incorporates a concern for society, so including each such term in the social welfare function amounts to multiple-counting the same concern. To illustrate this in another way, consider the following example, adapted to our setting from Bergstrom (2006). Suppose that each individual is consuming $c_i = 1$, which generates $A(\boldsymbol{c}) = 1$, and a private harm of K on each person. How much should society be willing to pay per person to eliminate the harm due to A? The natural answer is based on consumers' private utilities: K. Suppose that instead we were to use a criterion for consumer welfare that assigns a greater weight to $A(\boldsymbol{c})$, so that we would be willing to pay more than K per person. Then we would be willing to impose a tax greater than K and use it to eliminate the harm (e.g., by cleaning up air pollution). But this intervention would strictly lower all consumers' private utilities. It would be exceedingly odd to use consumers' concerns for others' private disutilities from an externality to justify a policy that makes everyone privately worse off.

Hence, consumer welfare is

$$\frac{\sum_{i=1}^{I+1}[u(c_i) - pc_i - KA(\boldsymbol{c})]}{I+1} = \frac{\sum_{i=1}^{I+1}[u(c_i) - pc_i] - (I+1)KA(\boldsymbol{c})}{I+1}$$

$$= \frac{\sum_{i=1}^{I+1} u(c_i) - pc_i - Kc_i}{I+1}.$$

When all consumers choose the same c_i, then the above, together with producer surplus, gives the social-welfare function in the text.

CENTRAL EUROPEAN UNIVERSITY, AUSTRIA
SAFE AND GOETHE UNIVERSITY FRANKFURT, GERMANY
UNIVERSITY OF BONN, GERMANY

REFERENCES

Aghion, Philippe, Roland Bénabou, Ralf Martin, and Alexandra Roulet, "Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?," *American Economic Review: Insights*, 5 (2023), 1–20. https://doi.org/10.1257/aeri.20210014

Andre, Peter, Ingar Haaland, Christopher Roth, and Johannes Wohlfart, "Narratives about the Macroeconomy," ECONtribute Discussion Paper No. 127, 2022.

Arnold, Lutz G., "On the Neutrality of Socially Responsible Investing," *Theoretical Economics*, 18 (2023), 65–95. https://doi.org/10.3982/TE4719

Arrow, Kenneth J., and Frank Hahn, *General Competitive Analysis* (Oakland, CA: Holden-Day, 1971).

Auger, Pat, Paul Burke, Timothy M. Devinney, and Jordan J. Louviere, "What Will Consumers Pay for Social Product Features?," *Journal of Business Ethics*, 42 (2003), 281–304. https://doi.org/10.1023/A:1022212816261

Awad, Edmond, Sohan Dsouza, Azim Shariff, Iyad Rahwan, and Jean-François Bonnefon, "Universals and Variations in Moral Decisions Made in 42 Countries by 70,000 Participants," *Proceedings of the National Academy of Sciences*, 117 (2020), 2332–2337. https://doi.org/10.1073/pnas.1911517117

Babiker, Mustafa H., "Climate Change Policy, Market Structure, and Carbon Leakage," *Journal of International Economics*, 65 (2005), 421–445. https://doi.org/10.1016/j.jinteco.2004.01.003

Bartling, Björn, Roberto A. Weber, and Lan Yao, "Do Markets Erode Social Responsibility?," *Quarterly Journal of Economics*, 130 (2015), 219–266. https://doi.org/10.1093/qje/qju031

Bergstrom, Theodore C., "Benefit-Cost in a Benevolent Society," *American Economic Review*, 96 (2006), 339–351. https://doi.org/10.1257/000282806776157623

Bénabou, Roland, Armin Falk, and Luca Henkel, "Ends versus Means: Kantians, Utilitarians and Moral Decisions," University of Bonn Working Paper, 2022.

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales, "Exit versus Voice," *Journal of Political Economy*, 130 (2022), 3101–3145. https://doi.org/10.1086/720516

Burniaux, Jean-Marc, and Joaquim Oliveira Martins, "Carbon Leakages: A General Equilibrium View," *Economic Theory*, 49 (2012), 473–495. https://doi.org/10.1007/s00199-010-0598-y

Chater, Nick, and George Loewenstein, "The I-Frame and the S-Frame: How Focusing on Individual-Level Solutions Has Led Behavioral Public Policy

Astray," *Behavioral and Brain Sciences*, 46 (2023), E147. https://doi.org/10.1017/S0140525X22002023

Debreu, Gerard, *Theory of Value: An Axiomatic Analysis of Economic Equilibrium* (New Haven, CT: Yale University Press, 1959).

Dewatripont, Mathias, and Jean Tirole, "The Morality of Markets," *Journal of Political Economy*, forthcoming. https://doi.org/10.1086/729445

Dubey, Pradeep, and Martin Shubik, "Perfect Competition in Strategic Market Games with Interlinked Preferences," *Economics Letters*, 17 (1985), 3–4. https://doi.org/10.1016/0165-1765(85)90115-6

Dufwenberg, Martin, Paul Heidhues, Georg Kirchsteiger, Frank Riedel, and Joel Sobel, "Other-Regarding Preferences in General Equilibrium," *Review of Economic Studies*, 78 (2011), 613–639. https://doi.org/10.1093/restud/rdq026

Falk, Armin, Thomas Neuber, and Nora Szech, "Diffusion of Being Pivotal and Immoral Outcomes," *Review of Economic Studies*, 87 (2020), 2205–2229. https://doi.org/10.1093/restud/rdz064

Falk, Armin, and Nora Szech, "Morals and Markets," *Science*, 340 (2013), 707–711. https://doi.org/10.1126/science.1231566

Fehr, Ernst, and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114 (1999), 817–868. https://doi.org/10.1162/003355399556151

Felder, Stefan, and Thomas F. Rutherford, "Unilateral CO_2 Reductions and Carbon Leakage: The Consequences of International Trade in Oil and Basic Materials," *Journal of Environmental Economics and Management*, 25 (1993), 162–176. https://doi.org/10.1006/jeem.1993.1040

Ferrario, Beatrice, and Stefanie Stantcheva, "Eliciting People's First-Order Concerns: Text Analysis of Open-Ended Survey Questions," *American Economic Association: Papers & Proceedings*, 112 (2022), 163–169. https://doi.org/10.1257/pandp.20221071

Giesler, Markus, and Ela Veresiu, "Creating the Responsible Consumer: Moralistic Governance Regimes and Consumer Subjectivity," *Journal of Consumer Research*, 41 (2014), 840–857. https://doi.org/10.1086/677842

Green, Daniel, and Benjamin Roth, "The Allocation of Socially Responsible Capital," SSRN Working Paper, 2021. https://dx.doi.org/10.2139/ssrn.3737772

Grethe, Harald, "The Economics of Farm Animal Welfare," *Annual Review of Resource Economics*, 9 (2017), 75–94. https://doi.org/10.1146/annurev-resource-100516-053419

Gruber, Jonathan, *Public Finance and Public Policy* (New York: Worth Publishers, 2005).

Hainmueller, Jens, Michael J. Hiscox, and Sandra Sequeira, "Consumer Demand for Fair Trade: Evidence from a Multistore Field Experiment," *Review of Economics and Statistics*, 97 (2015), 242–256. https://doi.org/10.1162/REST_a_00467

Hakenes, Hendrik, and Eva Schliephake, "Responsible Investment and Responsible Consumption," ECONtribute Discussion Paper No. 134, 2021.

Hart, Oliver D., David Thesmar, and Luigi Zingales, "Private Sanctions," *Economic Policy*, 39 (2024), 203–268. https://doi.org/10.1093/epolic/eiad041

Heinkel, Robert, Alan Kraus, and Josef Zechner, "The Effect of Green Investment on Corporate Behavior," *Journal of Financial and Quantitative Analysis*, 36 (2001), 431–449. https://doi.org/10.2307/2676219

Herweg, Fabian, and Klaus M. Schmidt, "How to Regulate Carbon Emissions with Climate-Conscious Consumers," *Economic Journal*, 132 (2022), 2992–3019. https://doi.org/10.1093/ej/ueac045

Intergovernmental Panel on Climate Change, "Climate Change 2022: Mitigation of Climate Change," World Meteorological Organization and United Nations Environment Programme Technical Report, 2022.

Jalil, Andrew J., Joshua Tasoff, and Arturo Vargas Bustamante, "Eating to Save the Planet: Evidence from a Randomized Controlled Trial Using Individual-Level Food Purchase Data," *Food Policy*, 95 (2020), 101950. https://doi.org/10.1016/j.foodpol.2020.101950

Kaufmann, Marc, Peter Andre, and Botond Kőszegi, "Replication Data for: 'Understanding Markets with Socially Responsible Consumers'," (2024), *Harvard Dataverse*, https://doi.org/10.7910/DVN/49CETN.

Krahnen, Jan, Jörg Rocholl, and Marcel Thum, "A Primer on Green Finance: From Wishful Thinking to Marginal Impact," *Review of Economics*, 74 (2023), 1–19. https://doi.org/10.1515/roe-2023-0019

Kyle, Albert S., "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, 56 (1989), 317–355. https://doi.org/10.2307/2297551

Mas-Colell, Andreu, "Noncooperative Approaches to the Theory of Perfect Competition: Presentation," *Journal of Economic Theory*, 22 (1980), 121–135. https://doi.org/10.1016/0022-0531(80)90036-8

Moisson, Paul-Henri, "Ethics and Impact Investment," Toulouse School of Economics Working Paper, 2020.

Norwood, F. Bailey, and Jayson L. Lusk, *Compassion, by the Pound: The Economics of Farm Animal Welfare* (Oxford: Oxford University Press, 2011). https://doi.org/10.1093/acprof:osobl/9780199551163.001.0001

Oehmke, Martin, and Marcus M. Opp, "A Theory of Socially Responsible Investment," SSRN Working Paper, 2022. https://dx.doi.org/10.2139/ssrn.3467644

Pástor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, "Sustainable Investing in Equilibrium," *Journal of Financial Economics*, 142 (2021), 550–571. https://doi.org/10.1016/j.jfineco.2020.12.011

Perino, Grischa, "Climate Campaigns, Cap and Trade, and Carbon Leakage: Why Trying to Reduce Your Carbon Footprint Can Harm the Climate," *Journal of the Association of Environmental and Resource Economists*, 2 (2015), 469–495. https://doi.org/10.1086/682572

Piccolo, Alessio, Jan Schneemeier, and Michele Bisceglia, "Externalities of Responsible Investments," SSRN Working Paper, 2022. https://doi.org/10.2139/ssrn.4183855

Roberts, Donald John, and Andrew Postlewaite, "The Incentives for Price-Taking Behavior in Large Exchange Economies," *Econometrica*, 44 (1976), 115–127. https://doi.org/10.2307/1911385

Rostek, Marzena J., and Ji Hee Yoon, "Imperfect Competition in Financial Markets: Recent Developments," SSRN Working Paper, 2023. https://doi.org/10.2139/ssrn.3710206

Schulze Tilling, Anna, "Changing Consumption Behavior with Carbon Labels: Causal Evidence on Behavioral Channels and Effectiveness," University of Bonn Working Paper, 2023.

Shubik, Martin, "Commodity Money, Oligopoly, Credit and Bankruptcy in a General Equilibrium Model," *Economic Inquiry*, 11 (1973), 24–38. https://doi.org/10.1111/j.1465-7295.1973.tb00959.x

Sobel, Joel, "Do Markets Make People Selfish," University of California, San Diego Working Paper, 2007.

Stantcheva, Stefanie, "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible," *Annual Review of Economics*, 15 (2023), 205–234. https://doi.org/10.1146/annurev-economics-091622-010157

Sundt, Swantje, and Katrin Rehdanz, "Consumers' Willingness to Pay for Green Electricity: A Meta-Analysis of the Literature," *Energy Economics*, 51 (2015), 1–8. https://doi.org/10.1016/j.eneco.2015.06.005

Trammell, Philip, "Ethical Consumerism," University of Oxford Working Paper, 2023.

Weitzman, Martin L., "Prices vs. Quantities," *Review of Economic Studies*, 41 (1974), 477–491. https://doi.org/10.2307/2296698

Ziegler, Andreas G. B., Giorgia Romagnoli, and Theo Offerman, "Morals in Multi-Unit Markets," *Journal of the European Economic Association*, forthcoming. https://doi.org/10.1093/jeea/jvae001.