Marc Kaufmann | Peter Andre | Botond Köszegi

Understanding Markets with Socially Responsible Consumers

SAFE Working Paper No. 411 | December 2023

Leibniz Institute for Financial Research SAFE
Sustainable Architecture for Finance in Europe
info@safe-frankfurt.de | www.safe-frankfurt.de

Understanding Markets with Socially Responsible Consumers*

Marc Kaufmann
Central European University

Peter Andre
SAFE and Goethe University Frankfurt

Botond Kőszegi
briq and University of Bonn

December 8, 2023

Abstract

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 non-trivially dampens a consumer's mitigation efforts, undermining responsible behavior. 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 dampening play a crucial role in our framework, we investigate them empirically via a tailored survey. Consistent with our model, consumers are predominantly consequentialist, and on average believe in dampening. Inconsistent with our model, however, many consumers fail to anticipate dampening. Third, therefore, we analyze how such "naive" consumers modify our theoretical conclusions. Naive consumers behave more responsibly than rational consumers in a single-good economy, but may behave less responsibly in a multi-good economy with cross-market spillovers. A mix of naive and rational consumers may yield the worst outcomes.

Keywords: socially responsible consumers, social preferences, climate change, externalities, competitive equilibrium, regulation, taxes, caps

JEL Codes: D01, D11, D50, D62, D64, D91

*Anik De, Khanh Hoang, and Wenjun Zheng provided fantastic research assistance. We thank 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, Paul Voss, as well as seminar and conference audiences for insightful discussions. Kőszegi thanks the European Research Council for financial support under Grant #788918. Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A01) is gratefully acknowledged. We also gratefully acknowledge research support from the Leibniz Institute for Financial Research SAFE.

1 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 paper 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 non-trivially dampens a person's mitigation efforts, eroding responsible behavior. 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 dampening play a crucial role in our framework, we assess their empirical relevance via a tailored survey. We find that consumers are predominantly consequentialist and on average believe in dampening, but many do not anticipate dampening. Third, therefore, we return to our framework and analyze how the latter "naive" consumers modify our conclusions. We establish that they behave more responsibly than rational consumers in some situations, but less responsibly in others.

We begin the first part of our paper in Section 2 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

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

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 $1/I$, 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 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 dampening generated by the market demand curve, and that demand curve is consistent with consumer behavior. Further, we define social welfare as total consumption utility (consumers' $u(c)$'s) minus production costs and K times the externality g.

Section 3 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 4, 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 hence 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. And we 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 5, 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 clean product and the dirty product 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, i.e., 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 6, we turn to the second major part of our paper: an empirical investigation of two crucial ingredients for the results above. First, due to rationality, consumers expect markets to exert a dampening effect. Instead, consumers may be "naive" in that they only think about their own direct effect. Second, due to consequentialism, consumers care about their total, dampened effect on the externality. Instead, consumers may be "deontological" in that they care about their action or direct effect. To distinguish these possibilities, we survey 2,000 US consumers. We measure dampening beliefs by asking consumers how they think a change in their own consumption would affect aggregate consumption. Using eight specific externality-generating goods, we study both reductions in consumption, including of fuel and meat, and reallocations of consumption, including from brown to green electricity and from new to second-hand clothing. We measure the nature of social concerns by comparing consumers' values for a specific action, such as a reduction in one's own CO_2 emissions, when it does versus when it does not reduce the aggregate externality.

We document three main facts. First, beliefs in dampening are widespread. For reductions in consumption, 43% of consumers believe in full or partial dampening. For reallocations of consumption, 34% believe in dampening, of which many — 59% — believe in full dampening. Second, however, even more consumers — 54% — express the naive view that a change in their own consumption affects aggregate consumption one-to-one. Third, most consumers care about the externalities they cause in a consequentialist way.

These findings imply that an empirically realistic model must include consequentialist consumers who understand dampening. 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 impact of naive consumers, especially when they coexist with rational consumers in the market.

Motivated by this last observation, the third part of our paper in Section 7 begins analyzing how naive consumers affect equilibrium outcomes. Since deontological consumers behave identically to naive consumers, our analysis also applies to this minority. We begin by observing that 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, encouraging responsible behavior. But in multi-market settings, naive consumers may underestimate their impact and hence behave less responsibly than rational consumers. As an illustration, consider our 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. More generally, the failure to appreciate cross-market effects brings naive consumers to markets in which they underestimate their effect, encouraging higher consumption. Hence, in general the market fails for both consumer types, but in different ways.

We conclude in Section 8 by emphasizing two points. First, if read as advice to consumers wondering about their market impacts, the message of our theory 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, since 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, it opens the possibility for studying the behavior of socially conscious consumers in myriad other settings, some of which we mention.

The appendix provides background on our framework and survey, and contains all our proofs.

Related Literature Our paper 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, Pastor et al., 2021, Dewatripont and Tirole, 2022, Herweg and Schmidt, 2022, Piccolo et al., 2022, Aghion et al., 2023, Arnold, 2023). 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 et al. (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 et al. (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.[3]

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 decisionmakers. In some recent models of socially responsible investment, investors — worrying about the "additionality" of their efforts — are cognizant of their price impacts and modify their behavior accordingly (e.g., Moisson, 2020, Green and Roth, 2021, Hakenes and Schliephake, 2021, Krahnen et al., 2023, Oehmke and Opp, 2022). 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. In their model, this is because consumers' motive to mitigate (β^R) is exogenously assumed to be different under different policies. In our model, the motive to mitigate depends on the degree of dampening, which we derive endogenously from the economic situation.

[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.

[3] At a technical level, Broccardo et al. assume that agents commit to their actions and firms can observe these actions before making their own decisions. Our equilibrium and its microfoundations are based on standard notions of a Walrasian auctioneer, who asks for demand schedules to find the equilibrium price.

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 et al., 2020, Ziegler et al., 2022). 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.

2 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 will reconsider these assumptions in Sections 6 and 7. The insights we uncover in the rational consequentialist model will remain pertinent there too.

2.1 A Single Consumer's Perspective

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, i.e., 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 $g(c) = q(c)$.[4]

The consumer correctly predicts the above outcomes, and maximizes

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

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 $-k\left(g(c) - g(0)\right)$ represents social concerns. Capturing our notion of social responsibility — being willing to modify one's consumption to mitigate the associated externality — the consumer derives disutility from her effect on the aggregate externality relative to when she consumes nothing. Since 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 Appendix A.1, we show how to derive the objective (1), as well as the social-welfare function (5) below, 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 (1) excludes it to start with.

[4] In this model, the notation g is redundant. Below, we will 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.

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

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

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,

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

II. The same consumer has a non-negligible 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. 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 non-vanishing extent. In particular, others have a dampening effect of

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

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 $1/I$, there are I other consumers, so their total response is comparable to c. Importantly, this property of the market derives from market clearing, so it is independent of whether consumers know about it or care about it. Nevertheless, while the property is irrelevant for a selfish consumer, it matters for the externality part of a socially responsible consumer's optimization problem. In the sense of having to consider her effect on the price, therefore, the consumer is not a price taker.[5]

Based on the above, an infinitesimally small consumer chooses c to solve[6]

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

Two immediate points follow. First, because the consumer has an impact on the externality that 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 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 et al., 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 believes in dampening, 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 non-market behavior). If the

[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, but the fundamental logic of our model appears unchanged. The consumer's impact on the probability of a price change is likely to vanish at rate $1/I$, but I others respond if the price change does occur. Hence, the consumer's effect is unlikely to vanish in expectation.

[6] Since $u(\cdot)$ is strictly concave and $\lim_{c\to\infty} u'(c) \leq 0$, Equation (4) has a unique solution for any p^*. The strict concavity of $u(\cdot)$ also ensures that the solution satisfies the second-order condition.

consumer is consequentialist with respect to externalities in markets, and she believes in dampening for whatever reason, then she behaves according to (4).

The Degree of Dampening By (3), dampening is an increasing function of the responsiveness or elasticity of demand relative to that of supply, $-\dfrac{D'(p^*)}{S'(p^*)} = -\dfrac{D'(p^*)/D(p^*)}{S'(p^*)/S(p^*)}$. 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.

Notably, the above relative elasticity also plays a central role in the incidence of a commodity tax in the same market. 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, we can use empirical work on tax incidence to gauge the level of dampening.

As an example, consider CO_2 emissions. The relevant dampening then equals the producer incidence of a broad carbon or other tax that covers a wide range of polluting consumption. Although estimates are sparse and vary, they suggest a degree of dampening well above zero. For Finland, Harju et al. (2022) measure producer incidence of fuel carbon taxes at 0.2. Using a broader range of products with energy as an input, Ganapati et al. (2020) estimate 0.3 for the US. Across the full range of taxed products, Benzarti et al. (2020, Table 1) find that the average producer incidence of VAT in the EU is 0.66-0.94. Furthermore, these studies use national taxes to estimate incidence. Since suppliers — but not consumers — can readily substitute between countries, the producer incidence of international taxes can be higher. To the extent that consumers care about global warming or other worldwide externalities, the latter estimates would be relevant, so the preceding estimates are likely to be lower bounds for the dampening of interest.[7]

[7] Incidence is often estimated based on state or other local tax changes, on tax changes for a specific product, or, typically, both. For instance, many researchers investigate the incidence of state gasoline taxes. These estimates map to dampening in our model if consumers care only about pollution in their own state, and this is caused by driving

Given that the degree of dampening can be non-trivial, its effect on the behavior of a socially responsible consumer can be non-trivial too. Suppose that the consumer values a one-ton reduction in atmospheric CO_2 at \$50 (the median valuation in our survey). Then, the above lower-bound estimates imply that the consumer chooses the same c in the market as she would in home production with a \$10-47/ton subsidy on polluting behavior. In this sense, the market can be seen as providing a potentially substantial incentive for generating externalities.

2.2 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 $s > 0$.[8] 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^*. Since 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.

(a) We introduce the equilibrium market responsiveness to the consumer's demand, q_c^*.

(b) We write demand as a function of q_c^*. From a consumer's first-order condition (4), we get $u'(q^*) = p^* + kq_c^*$.

in the state. For such settings, the modal study finds producer incidence near zero (e.g., Alm et al., 2009, Marion and Muehlegger, 2011, Silvia and Taylor, 2016), although some estimate higher numbers even here (e.g., Barron et al., 2004, Doyle and Samphantharak, 2008). Our main interest, however, is in situations where consumers care about externalities caused by a range of products, and also care about damage beyond their state borders. For the above reason of supply-side substitution between jurisdictions and products, producer incidence and hence dampening in the latter cases is likely to be higher.

[8] 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.

(c) We write q_c^* as a function of supply and demand. From (2), we have $q_c^* = \dfrac{S'(p^*)}{S'(p^*) - D'(p^*)}$. In the expression, $S'(p^*) = s$, but $D'(p^*)$ is not a primitive of the model. Hence, we introduce the equilibrium consumer price responsiveness $q_p^* \equiv D'(p^*)$.

(d) We derive q_p^* from consumer optimization, imposing that q_c^* does not change in response to an infinitesimal price change. Then, differentiating (4) with respect to the price yields a standard expression for price responsiveness, $q_p^* = 1/u''(q^*)$.

Combining these considerations yields the following definition of competitive equilibrium.

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: (1) supply equals q^*: $q^* = S(p^*)$; (2) demand equals q^*: $u'(q^*) = p^* + k \cdot q_c^*$; (3) market responsiveness is consistent with consumer price responsiveness: $q_c^* = s/(s - q_p^*)$; (4) consumer price responsiveness is consistent with optimization: $q_p^* = 1/u''(q^*)$.

Steps (a)-(d) can be applied mechanically to define competitive equilibrium in new situations, as we do in Sections 4 and 5 below. Step (a) recognizes the crucial market responsiveness the consumer cares about, and Step (b) expresses demand as a function of that market responsiveness. These steps follow directly from consumer preferences. Then, Steps (c) and (d) express market responsiveness as a function of demand. Step (c) uses market clearing; in Sections 4 and 5, just applying Equation (2) to the situation at hand. Finally, Step (d) derives price responsiveness from the consumer's optimality condition in Step (b).

In Appendix A, we provide some foundational analysis for our equilibrium concept. In Section A.2, we outline a way to think about equilibrium determination graphically. In Sections A.3 and A.4, 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 non-trivial price impacts (see Rostek and Yoon, 2020, 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 Appendix A.1, 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 social 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.

3 Failures of Socially Responsible Consumerism

In this section, we demonstrate ways in which 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 $k = K$ — i.e., consumers attach the same value to the externality as the social planner. Imagine, for example, 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 subtle pecuniary externality that emerges in the presence of social responsibility. Since consumers are restraining consumption, their marginal private utilities 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 $g = \sum_{i'} c_{i'}$ as before. Social welfare equals $[\sum_i u_i(c_i)] - Kg$. Then, the following is obvious:

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

Since 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. As a prominent example, Klaus Schwab, founder of the World Economic Forum in Davos, proposes that individuals should affect change as consumers and investors through "stakeholder capitalism," rather than rely on "shareholder capitalism" (pure profit maximization) or "state capitalism" (government intervention).[9] Our result that the social optimum does not obtain even with $k = K$ shows that stakeholder capitalism fails on its own terms: dealing with externalities cannot be successfully outsourced even to extremely responsible individuals.

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 $(-1/u''(c))$ increases fast. As an example, consumers' utility function for air travel may decrease sharply below, but flatten out quickly above, one flight. Then, if a consumer expects everyone to take one flight, 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 takes one flight. But if the consumer expects everyone to take multiple flights, 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 takes multiple flights.

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.

[9] See, e.g., "Why we need the 'Davos Manifesto' for a better kind of capitalism" (Schwab 2019, `https://www.weforum.org/agenda/2019/12/why-we-need-the-davos-manifesto-for-better-kind-of-capitalism/`).

The detrimental effects of many externalities, such as global warming or ocean pollution, have not been 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 take just one flight. This reduces dampening, so as the shock dissipates and prices drop, consumers may naturally stay in the low-consumption equilibrium.

4 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.

4.1 Permit-Supply Policies

We first 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 (a)-(d) before Definition 1 above. It is immediately apparent that only Step (c), the expression for market responsiveness q_c^*, requires modification. Here, we let p be the consumer price, and show that once we impose market clearing for permits, we can express supply as a function $S_{\text{net}}(p)$. Then, Equation (2) implies $q_c^* = \dfrac{S'_{\text{net}}(p^*)}{S'_{\text{net}}(p^*) - D'(p^*)}$.

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

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

For the full definition of competitive equilibrium, see the appendix.

We consider situations in which the policymaker 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, and is the scenario under which the contribution of socially responsible consumers to mitigation is most crucial. We then 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 $q_A^* < q_B^*$.[10] By market clearing for the product ($q_i^* = S(p_i^* - \tau^*)$

[10] 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^*.

for $i = A, B$), the A-type policy also leads to a lower consumer price ($p_A^* < p_B^*$), making it extra preferable if we deem consumer surplus more important than producer surplus.

While different policies can achieve the same outcomes with selfish consumers,[11] 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. Herweg and Schmidt, however, reach that conclusion by exogenously imposing that consumers are more willing to mitigate under a tax. Our framework endogenizes this willingness, and thereby allows us to study other policy questions without new exogenous assumptions about how consumers think. We do so next.

4.2 Other Policy Examples

Improving on Taxes We show that a regulator can do better than with a unit tax.[12] 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.*

While 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:

[11] To see this, consider any demand curve $D(p)$ and supply curve $S(p)$ satisfying the conditions in Section 2. In equilibrium $q^* = S(p^* - \tau^*) = D(p^*)$. Hence, there is a unique pair p^*, q^* that is consistent with a given τ^*.

[12] Adapting the definition of competitive equilibrium to these settings is straightforward, and hence 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.

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.[13] 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.[14] For simplicity in stating our results, we assume that 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

[13] The results are identical if consumers can distinguish but are in equilibrium indifferent between home and foreign-sourced purchases. Analogously to Section 5 below, such an equilibrium always exists.

[14] Our formulation assumes that the consumer cares equally about home and foreign pollution. This should hold true for externalities with world-wide 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 production. Then, e_f may be lower than e_h even if foreign production is more polluting.

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 $g_A^* < g_B^*$.[15]

Proposition 6 (Cap versus Tax Under Trade)**.**

I. If

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

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 the argument (e.g., Intergovernmental Panel on Climate Change, 2022, page 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

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

consumer prices of traded goods, leading foreign suppliers to decrease production.[16]

5 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, i.e., that every consumer purchases exactly one unit. Later, we reintroduce quantity choice into the picture.

5.1 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 $[\underline{k}, \overline{k}] \subset [0, K]$ according to the continuous, positive-valued density function f.[17] 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.

[16] 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.

[17] 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.

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 appendix. In Step (a), we introduce a market responsiveness $\mathrm{d}Q$ that determines how much the equilibrium quantities respond if a consumer moves her consumption from one market to the other. In Step (b), we identify an optimality condition for demand from $\mathrm{d}Q$. 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 (c), we express $\mathrm{d}Q$ 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 (d) 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 $\mathrm{d}Q$ 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 hence can be ignored when making welfare comparisons.

5.2 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, i.e., 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 1. 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 above 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 non-trivial. 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 non-trivial, kickstarting the unraveling.

Part III states that if the products generate sufficiently different externalities, then at least one other equilibrium exists. In such a non-selfish 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.[18] 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.[19]

For situations in which the selfish equilibrium is played, two additional 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 et al. (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 et al.'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

[18] 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 paper.

[19] 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 as we have mentioned, 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.

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.

5.3 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 2, 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

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

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 2.1 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 non-organic 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 above selfish equilibrium still exists, but now prices of the products equal after normalizing by v: $p^{c*} = vp^{d*}$.

6 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.[20]

6.1 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 re-weighting and correcting for these imbalances. Appendix C.1 presents further details on the sample.

[20] 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).

6.2 Design

Beliefs about Dampening We measure beliefs about dampening 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 8 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 second-hand 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 context. Next, the respondent is asked to consider two scenarios that differ in her personal consumption level and to estimate how the yearly total global consumption of the good would be affected.[21] 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.

Then, 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.[22,23] Finally, the respondent explains her prediction in an open-ended

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

[22] 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 (Appendix C.5.2).

[23] Our scenario-based approach induces controlled variation but renders it difficult to incentivize respondents'

text box. The open-ended format allows respondents to express themselves freely and unconstrained by the researcher, providing a direct lens into consumers' reasoning about dampening (Ferrario and Stantcheva, 2022, Andre et al., 2022). See Appendix C.6 for the full instructions of the survey.

Nature of Social Concerns We measure the relevance of consequentialist concerns by comparing consumers' valuations for effective versus ineffective mitigation actions. Respondents are randomly assigned to one of four practically relevant externalities — CO_2 emissions, non-recyclable 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 why 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.[24]

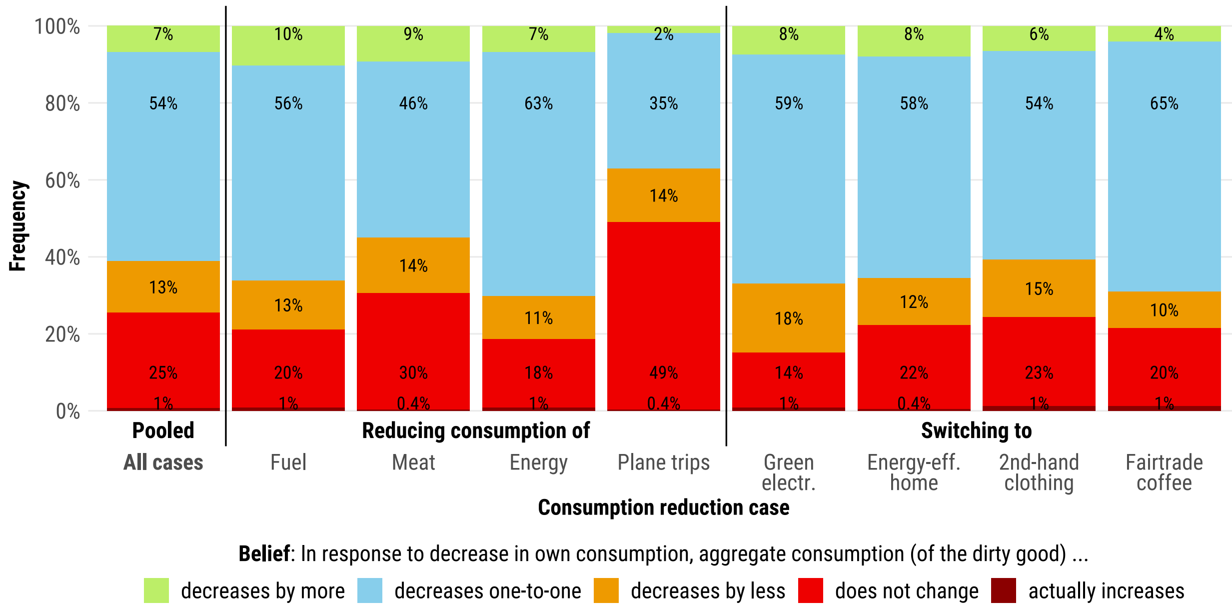
6.3 Results

Beliefs in Dampening Figure 1 displays the distribution of dampening beliefs. Across scenarios, 29% to 63% of consumers predict that their own consumption reduction has a dampened impact on aggregate consumption. On average, the share of consumers who believe in dampening is 38%, with 43% for reductions of consumption and 34% for reallocations of consumption. Among consumers who believe in dampening, most believe in full dampening: 69% for reductions of consumption and 59% for reallocations of consumption. 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 average

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 (Appendix C.5.3).

[24] Our focus on four real-world settings helps us to 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 et al., 2022, Hart et al., 2023). However, it implies that we cannot incentivize responses. Reassuringly, Bénabou et al. (2022) document that the strength of consequentialist behavior does not differ between incentivized and non-incentivized choices.

Figure 1: Consumers' Beliefs about Dampening



Notes: 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 each of the eight cases.

of 0.28.[25] The average consumer's belief in non-trivial quantity dampening, and many consumers' belief in full substitution dampening, are consistent with the beliefs rational consumers hold in our models above.

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, i.e., 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 direct effect, and (iii) explanations for a multiplier effect.

[25] 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 1: Reasoning about Dampening: Example Responses

Dampening	**Direct effect**
"Someone else would buy the meat I didn't buy and eat it."	"It makes sense that if I use less it means the total used will be less."
"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 would think that my personal consumption would be directly related to the global consumption, gallon for gallon."
	Positive multiplier
"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."	

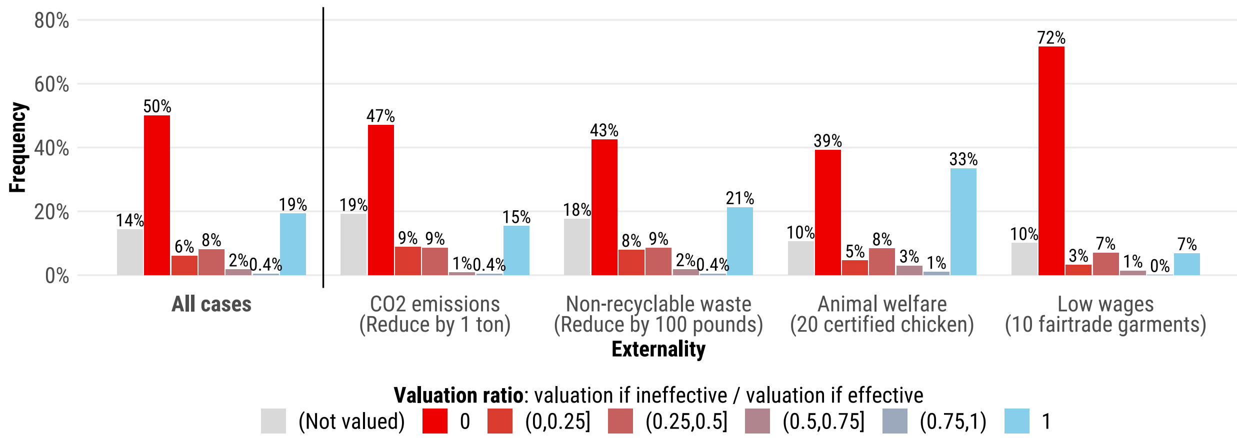
Notes: Examples of open-ended explanations from the consumer survey. Consumers make a prediction and are then asked, "Please explain why you chose this response."

78% of the responses can be classified.[26]

The results mirror our earlier conclusions (see Appendix Figure C.1). A significant share of consumers — 25% — provide an explanation consistent with dampening beliefs. They argue that their reduced consumption will be offset by others' consumption, or that being minuscule, they have little to no influence on global consumption. A few consumers identify exactly our price-based mechanism. As we have emphasized, however, even consequentialist consumers who believe in dampening without being able to spell out the precise dampening mechanism solve the same first-order condition as a rational consumer, and hence can from the perspective of our model be considered as-if rational. 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. Table 1 shows a few example responses for illustration.

[26] Each response is coded independently by two research assistants who have been extensively trained for this task. We cross-verify each classification. Reassuringly, the inter-rater reliability is high: the two coders agree in 87% of all cases. 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.

Figure 2: Consumers' Concern for Consequences



Notes: 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 1.

We present a series of additional analyses in Appendix C.2. We find that young and Republican-leaning consumers are most likely to voice dampening beliefs and explanations. 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 in Dampening)**.** *Consumers' beliefs are heterogeneous. A sizeable share believe in dampening, and so does the average consumer. But approximately half believe that they have a one-to-one effect on aggregate consumption.*

Nature of Social Concerns Next, we turn to the nature of consumers' social concerns. In line with a large literature showing that individuals care about the externalities they cause (e.g., Auger et al., 2003, Hainmueller et al., 2015), most consumers — 86% — report a positive willingness to pay to reduce global externalities. For example, the median valuation for a one-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 non-recyclable waste is \$20.

Table 2: 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."

Going beyond this basic finding, Figure 2 presents the distribution of valuation ratios, defined as a consumer's valuation for ineffective externality reduction divided by her valuation for effective externality reduction. A ratio close to zero (red colors) means that consumers primarily care about the consequences of their reduction, consistent with a consequentialist concern for externalities. A ratio close to one (blue colors) means that consumers value their own reduction irrespective of its net consequences, consistent with deontological or warm-glow preferences. The ratio is not defined (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. Most 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 reduction 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. 69% of consumers who report a positive valuation for effective actions voice only consequentialist arguments (see 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 it ultimately may 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 simultaneously. Table 2 gives a few example responses for illustration.

We present a series of additional analyses in Appendix C.3. We find that consequentialist concerns are most common among younger and politically independent consumers. As before, our results replicate in a series of sensitivity analyses. Importantly, we also find that consequentialism is the predominant attitude among consumers who believe in dampening. We summarize:

Empirical Result 2 (Nature of Social Concerns)**.** *Consumers mostly care about the externalities they cause in a consequentialist way.*

We draw three main conclusions from our survey. First, consequentialism is a good approximation of consumers' social concerns, although there is a significant minority of deontological consumers as well. Second, an empirically realistic model must take rational consequentialist consumers seriously. Third, however, it is also important to understand the impact of naive consumers, especially when they coexist with rational consumers in the population. These conclusions motivate the next section.

7 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

[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. Appendix C.4 describes the coding procedure in detail. 12% of responses cannot unambiguously be classified or do not contain a clear explanation.

direct effect is what she cares about, and a naive consequentialist consumer does so because she fails to anticipate dampening. While the two types therefore behave identically, we mostly describe the results in terms of naive consumers, who we found are more numerous.

Single-Good Setting Consider first the single-good model of Section 2, 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 (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, i.e., an optimal Pigouvian tax.

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 5.3 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.

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.*

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

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 non-monotonic effect on welfare, with intermediate shares generating the worst outcomes. This intermediate case, featuring bimodal dampening beliefs at 0 and 1, provides the most accurate theoretical representation of our survey findings.

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 two 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, lowering social welfare. In the context of clothing, for example, 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]

[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 5 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 believe in 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

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 impact: 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 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.

---

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.

Multiple Inputs To conclude, we discuss another instance in which naiveté about cross-market effects is harmful. This time, the effect runs through supply-side responses. Suppose that there are two production resources that can be used to produce two goods. Resource 1 is clean, and is available in a fixed supply of 1. Resource 2 is dirty with an externality of 1 per unit, and is available in completely elastic supply at a price of 1. Good 1 is produced with the clean resource with constant returns: $q_1 = r_1^c$, where r_1^c is the amount of the clean resource devoted to the production of good 1. Good 2 can be produced with both resources, with a perfect-substitutes production technology: $q_2 = r_2^c + r_2^d$, where r_2^c and r_2^d are the amounts of the two resources devoted to the production of good 2. Consumption utility from consuming an amount $c_i \geq 0$ of good i is $u_i(c_i)$, where $u_i(\cdot)$ is a strictly concave utility function with $u_1'(0) > 1 + k$, $u_1'(1) < 1$, $u_2'(1) > 1 + k$, and $\lim_{c\to\infty} u_2'(c) < 1$. Consumers are homogeneous in k. Naive consumers perceive the externalities from consuming goods 1 and 2 to be 0 and 1, respectively. The share of naive consumers is α.

As an example, goods 1 and 2 could be organic and non-organic food, respectively. The clean resource is land and the dirty resource is pesticides and oil-based fertilizers, the latter of which cannot be used for organic food. Good 1 could alternatively be biofuel, which — even if produced with pesticides — can be seen as clean because it saves on the use of fossil fuel (good 2).

Proposition 9 (Naiveté and Supply-Side Substitution)**.** *The competitive-equilibrium price of the clean resource, as well as of both goods, is 1. Naive consumers choose c_1 satisfying $u_1'(c_1) = 1$, while rational consumers choose c_1 satisfying $u_1'(c_1) = 1 + k$. Both types choose c_2 satisfying $u_2'(c_2) = 1 + k$. Social welfare is strictly decreasing in α.*

In this economy, buying more of good 1 uses up more of the clean resource, forcing producers to substitute more of the dirty resource in the production of good 2. Rational consumers understand this cross-market effect, but naive consumers do not, so the latter consumers purchase more of good 1. Hence, the more naive consumers there are, the lower is social welfare.

We now return to our assumption that a naive consumer perceives the externality from consuming good 2 to be 1. This perception is correct. Good 2, however, is produced using a mix of the two resources, and it is plausible that a naive consumer fails to understand her marginal effect. For example, suppose that the naive consumer perceives the externality to be the average amount of the dirty resource used for producing a unit of good 2 in equilibrium. Then, she consumes too much of good 2 as well as good 1, lowering social welfare further. Intuitively, although she understands the direct externality associated with her equilibrium consumption, she fails to appreciate that increasing her consumption of good 2 changes the mix of inputs used by firms. This example also illustrates that it can be challenging to specify exactly how naive consumers perceive their impact, and predictions can be sensitive to such assumptions.

Our results regarding naive or, equivalently, deontological consumers reinforce the message that an individual-responsibility-based approach to dealing with externalities is necessarily incomplete. Whatever the mix of rational and naive consumers, market outcomes can be far from optimal even if consumers are extremely responsible. Nevertheless, the sources of distortion depend crucially on consumers' perceptions of dampening and the nature of their social concerns.

8 Conclusion

Our paper 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 5.

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 non-local 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," i.e., 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.

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*, 2023, *5* (1), 1–20.

Alm, James, Edward Sennoga, and Mark Skidmore, "Perfect Competition, Urbanization, and Tax Incidence in the Retail Gasoline Market," *Economic Inquiry*, 2009, *47* (1), 118–134.

Andre, Peter, Ingar Haaland, Christopher Roth, and Johannes Wohlfart, "Narratives about the Macroeconomy," *Working Paper*, 2022.

Arnold, Lutz G., "On the Neutrality of Socially Responsible Investing," *Theoretical Economics*, 2023, *18* (1), 65–95.

Arrow, Kenneth J. and Frank Hahn, *General Competitive Analysis*, 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*, 2003, *42*, 281–304.

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*, 2020, *117* (5), 2332–2337.

Babiker, Mustafa H., "Climate Change Policy, Market Structure, and Carbon Leakage," *Journal of International Economics*, March 2005, *65* (2), 421–445.

Bardsley, Nicholas, "Control Without Deception: Individual Behaviour in Free-Riding Experiments Revisited," *Experimental Economics*, 2000, *3*, 215–240.

Barron, John M., Kelly Hunt Blanchard, and John R. Umbeck, "An Economic Analysis of a Change in an Excise Tax," *Journal of Economic Education*, 2004, *35* (2), 184–196.

Bartling, Björn, Roberto A. Weber, and Lan Yao, "Do Markets Erode Social Responsibility?," *Quarterly Journal of Economics*, 2015, *130* (1), 219–266.

Benzarti, Youssef, Dorian Carloni, Jarkko Harju, and Tuomas Kosonen, "What Goes Up May Not Come Down: Asymmetric Incidence of Value-Added Taxes," *Journal of Political Economy*, 2020, *128* (12), 4438–4474.

Bergstrom, Theodore C., "Benefit-Cost in a Benevolent Society," *American Economic Review*, 2006, *96* (1), 339–351.

Bordalo, Pedro, John Conlon, Nicola Gennaioli, Spencer Kwon, and Andrei Shleifer, "How People Use Statistics," *Working Paper*, 2023.

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales, "Exit versus Voice," *Journal of Political Economy*, December 2022, *130* (12), 3101–3145.

Burniaux, Jean-Marc and Joaquim Oliveira Martins, "Carbon Leakages: a General Equilibrium View," *Economic Theory*, 2012, *49* (2), 473–495.

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

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*, 2023, *46*, E147.

Chinoy, Sahil, Nathan Nunn, Sandra Sequeira, and Stefanie Stantcheva, "Zero-Sum Thinking and the Roots of U.S. Political Divides," *Working Paper*, 2023.

Debreu, Gerard, *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, Yale University Press, 1959.

Dechezleprêtre, Antoine and Misato Sato, "The Impacts of Environmental Regulations on Competitiveness," *Review of Environmental Economics and Policy*, 2017, *11* (2), 183–206.

Dewatripont, Mathias and Jean Tirole, "The Morality of Markets," 2022. Working Paper.

Doyle, Joseph J. and Krislert Samphantharak, "$2.00 Gas! Studying the Effects of a Gas Tax Moratorium," *Journal of Public Economics*, 2008, *92* (3-4), 869–884.

Dubey, Pradeep and Martin Shubik, "Perfect Competition in Strategic Market Games with Interlinked Preferences," *Economics Letters*, 1985, *17* (1-2), 3–4.

Dufwenberg, Martin, Paul Heidhues, Georg Kirchsteiger, Frank Riedel, and Joel Sobel, "Other-Regarding Preferences in General Equilibrium," *Review of Economic Studies*, 2011, *78* (2), 613–639.

Falk, Armin and Nora Szech, "Morals and Markets," *Science*, 2013, *340* (6133), 707–711.

_ , Thomas Neuber, and Nora Szech, "Diffusion of Being Pivotal and Immoral Outcomes," *Review of Economic Studies*, 2020, *87* (5), 2205–2229.

Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

Felder, Stefan and Thomas F. Rutherford, "Unilateral CO2 Reductions and Carbon Leakage: The Consequences of International Trade in Oil and Basic Materials," *Journal of Environmental Economics and Management*, September 1993, *25* (2), 162–176.

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

Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker, "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes," *American Economic Journal: Applied Economics*, 2020, *12* (2), 303–342.

Giesler, Markus and Ela Veresiu, "Creating the Responsible Consumer: Moralistic Governance Regimes and Consumer Subjectivity," *Journal of Consumer Research*, 2014, *41* (3), 840–857.

Green, Daniel and Benjamin Roth, "The Allocation of Socially Responsible Capital," 2021. Working Paper.

Grethe, Harald, "The Economics of Farm Animal Welfare," *Annual Review of Resource Economics*, 2017, *9* (1), 75–94.

Grubb, Michael, Nino David Jordan, Edgar Hertwich, Karsten Neuhoff, Kasturi Das, Kaushik Ranjan Bandyopadhyay, Harro van Asselt, Misato Sato, Ranran Wang, William A. Pizer, and Hyungna Oh, "Carbon Leakage, Consumption, and Trade," *Annual Review of Environment and Resources*, 2022, *47* (1), 753–795.

Gruber, Jonathan, *Public Finance and Public Policy*, first ed., 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*, 2015, *97* (2), 242–256.

Hakenes, Hendrik and Eva Schliephake, "Responsible Investment and Responsible Consumption," 2021. Working Paper.

Harju, Jarkko, Tuomas Kosonen, Marita Laukkanen, and Kimmo Palanne, "The Heterogeneous Incidence of Fuel Carbon Taxes: Evidence from Station-Level Data," *Journal of Environmental Economics and Management*, 2022, *112.*

Hart, Oliver D., David Thesmar, and Luigi Zingales, "Private Sanctions," *Working Paper*, 2023.

Heinkel, Robert, Alan Kraus, and Josef Zechner, "The Effect of Green Investment on Corporate Behavior," *Journal of Financial and Quantitative Analysis*, 2001, *36* (4), 431–449.

Herweg, Fabian and Klaus M. Schmidt, "How to Regulate Carbon Emissions with Climate-Conscious Consumers," *Economic Journal*, 2022, *132* (648), 2992–3019.

Intergovernmental Panel on Climate Change, "Climate Change 2022: Mitigation of Climate Change," Technical Report, World Meteorological Organization and United Nations Environment Programme 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*, 2020, *95* (July), 101950. Publisher: Elsevier.

Krahnen, Jan, Jörg Rocholl, and Marcel Thum, "A Primer on Green Finance: From Wishful Thinking to Marginal Impact," *Review of Economics*, April 2023, *74* (1), 1–19.

Kyle, Albert S., "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, 1989, *56* (3), 317–355.

Marion, Justin and Erich Muehlegger, "Fuel Tax Incidence and Supply Conditions," *Journal of Public Economics*, 2011, *95* (9-10), 1202–1212.

Mas-Colell, Andreu, "Noncooperative Approaches to the Theory of Perfect Competition: Presentation," *Journal of Economic Theory*, 1980, *22* (2), 121–135.

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 University Press, 2011.

Oehmke, Martin and Marcus M. Opp, "A Theory of Socially Responsible Investment," 2022. Working Paper.

Pasek, Josh, Matthew Debell, and Jon A. Krosnick, "Standardizing and Democratizing Survey Weights: The ANES Weighting System and anesrake," *Working Paper*, 2014.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, "Sustainable Investing in Equilibrium," *Journal of Financial Economics*, November 2021, *142* (2), 550–571.

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*, 2015, *2* (3), 469–495. Publisher: The University of Chicago Press.

Piccolo, Alessio, Jan Schneemeier, and Michele Bisceglia, "Externalities of Responsible Investments," *Working Paper*, 2022.

Roberts, Donald John and Andrew Postlewaite, "The Incentives for Price-Taking Behavior in Large Exchange Economies," *Econometrica*, 1976, *44* (1), 115–127.

Rostek, Marzena and Ji Hee Yoon, "Equilibrium Theory of Financial Markets: Recent Developments," 2020. Working Paper.

Shubik, Martin, "Commodity Money, Oligopoly, Credit and Bankruptcy in a General Equilibrium Model," *Economic Inquiry*, 1973, *11* (1), 24–38.

Silvia, Louis and Christopher T. Taylor, "Gasoline and Diesel Tax Incidence: The 2003 Washington State Nickel Funding Package," *The B.E. Journal of Economic Analysis & Policy*, 2016, *16* (3), 1539–1562.

Sobel, Joel, "Do Markets Make People Selfish," *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*, 2023, *15*, 205–234.

Sundt, Swantje and Katrin Rehdanz, "Consumers' Willingness to Pay for Green Electricity: A Meta-Analysis of the Literature," *Energy Economics*, September 2015, *51*, 1–8.

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

Trammell, Philip, "Ethical Consumerism," 2023. Working Paper.

Weitzman, Martin L., "Prices vs. Quantities," *Review of Economic Studies*, 1974, *41* (4), 477–491.

Ziegler, Andreas, Giorgia Romagnoli, and Theo Offerman, "Morals in Multi-Unit Markets," *Working Paper*, 2022.

Appendix

A Background on Competitive Equilibrium

A.1 Private Utility, Social Concerns, and Social Welfare

In this subsection, we motivate the consumer utility function (1) as well as the social welfare function (5) we have assumed in the text.

Suppose that there are $I+1$ identical consumers. Individual i consumes c_i, and her private utility is

$$u(c_i) - pc_i - K \cdot \frac{\sum_{i'=1}^{I+1} c_{i'}}{I+1}.$$

Hence, each consumer's utility is decreasing in the average consumption in society, for instance because of the economic, health, and social consequences of global warming or ocean pollution. Consumer i realizes that her consumption contributes to the externality, which affects others negatively. The total disutility others suffer from the externality is

$$K \cdot \frac{I}{I+1} \cdot \sum_{i'=1}^{I+1} c_{i'}.$$

Being socially responsible, the consumer partially or fully takes into account this disutility, or her effect on it. Including this social concern, then, her utility is

$$u(c_i) - pc_i - K \cdot \frac{\sum_{i'=1}^{I+1} c_{i'}}{I+1} - k \cdot \frac{I}{I+1} \cdot \sum_{i'=1}^{I+1} c_{i'} = u(c_i) - pc_i - \left(k \cdot \frac{I}{I+1} + K \cdot \frac{1}{I+1}\right) \cdot \sum_{i'=1}^{I+1} c_{i'},$$

where $k \leq K$. As $I \to \infty$, the weight the consumer attaches to the total market quantity $\sum_{i'=1}^{I+1} c_{i'}$ approaches k, as we have assumed in the text. Note also that if $k = 0$, then in the limit the consumer ignores her effect on the externality, as in the classical case of a selfish consumer.[30]

[30] To simplify matters, we assume that the consumer's utility is linear in both money and the externality. This could lead to the non-existence of a utility-maximizing choice if the consumer could exchange directly between the two, e.g., if she had access to offsets. To analyze such situations, it is necessary to adjust the consumer's problem slightly, for instance by assuming non-linear utility in money.

Our formulation assumes that in choosing her consumption, a socially responsible consumer cares about her effect on the externality, but not about her effect on others' private utilities. Hence, for instance, in choosing how much to fly the consumer thinks about global warming and its effect on humanity, but does not internalize others' enjoyment of flying. 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. For simplicity, we also assume that the weight k does not depend I. Our points remain unchanged as long as k does not vanish when $I \to \infty$. A non-vanishing 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 the social-welfare function equals the average of individuals' *private* utilities. First-pass conventional logic might dictate that the social welfare function equals the average of individuals' total utilities. But the weight a consumer puts on the externality 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 consumer is consuming $c_i = 1$, which generates an aggregate externality of $g = I + 1$, and hence a private harm of K on each person. How much should society be willing to pay per person to eliminate the harm from the externality? The natural answer is K. If we were willing to pay more than K, 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 disutility from the externality to justify a policy that makes everyone privately worse off.

Hence, the social welfare function is

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

When all consumers choose the same c_i, the above reduces to the social welfare function in the text.

A.2 Graphical Illustration of Competitive Equilibrium

Figure 3 shows how to think about equilibrium determination using modifications of a standard supply-demand diagram, and illustrates the possibility of multiple equilibria. With slight abuse of notation, we write demand and supply curves as functions of the quantity rather than of the price (as we do in the rest of the paper). We can start by drawing the exogenous and linear supply curve $S(q)$. To proceed, we start from the demand curve $D_0(q)$ that would obtain if consumers were selfish ($k = 0$). This is a standard demand curve given by $D_0(q) = u'(q)$. For simplicity in making our points, we assume that this demand curve is two-piece linear and kinked. As indicated on the figure, the intersection of the supply curve and the selfish demand curve gives the unique selfish competitive equilibrium.

Now assume $k > 0$. Based on Definition 1, the equilibrium quantity q and price p must satisfy $u'(q) - p - kq_c = 0$, or $p = u'(q) - ks/(s - 1/u''(q))$. Using that $u'(q) = D_0(q)$ and therefore $u''(q) = D_0'(q)$, we define a "virtual demand curve"

$$D_k(q) = D_0(q) - k \cdot \frac{s}{s - 1/D_0'(q)}.$$

It is easy to see that intersections of this virtual demand curve with the supply curve correspond to competitive equilibria. We refer to $D_k(q)$ as virtual because it cannot be easily observed in its entirety. For instance, a common way of identifying the demand curve, looking at changes in the equilibrium due to exogenous shifts in supply, does not in general trace out any virtual demand curve as the virtual demand curve depends on the supply curve.

Multiple competitive equilibria can occur because even for a downward-sloping selfish demand curve $D_0(q)$, the virtual demand curve $D_k(q)$ can increase. Indeed, for our kinked $D_0(q)$, $D_k(q)$ jumps up. Intuitively, at the kink the price sensitivity of demand and hence dampening jump up, so a consumer's willingness to mitigate jumps down. This results in the two competitive equilibria identified in the figure. Both feature lower consumption than the selfish competitive equilibrium, but the consumption levels are quite different.

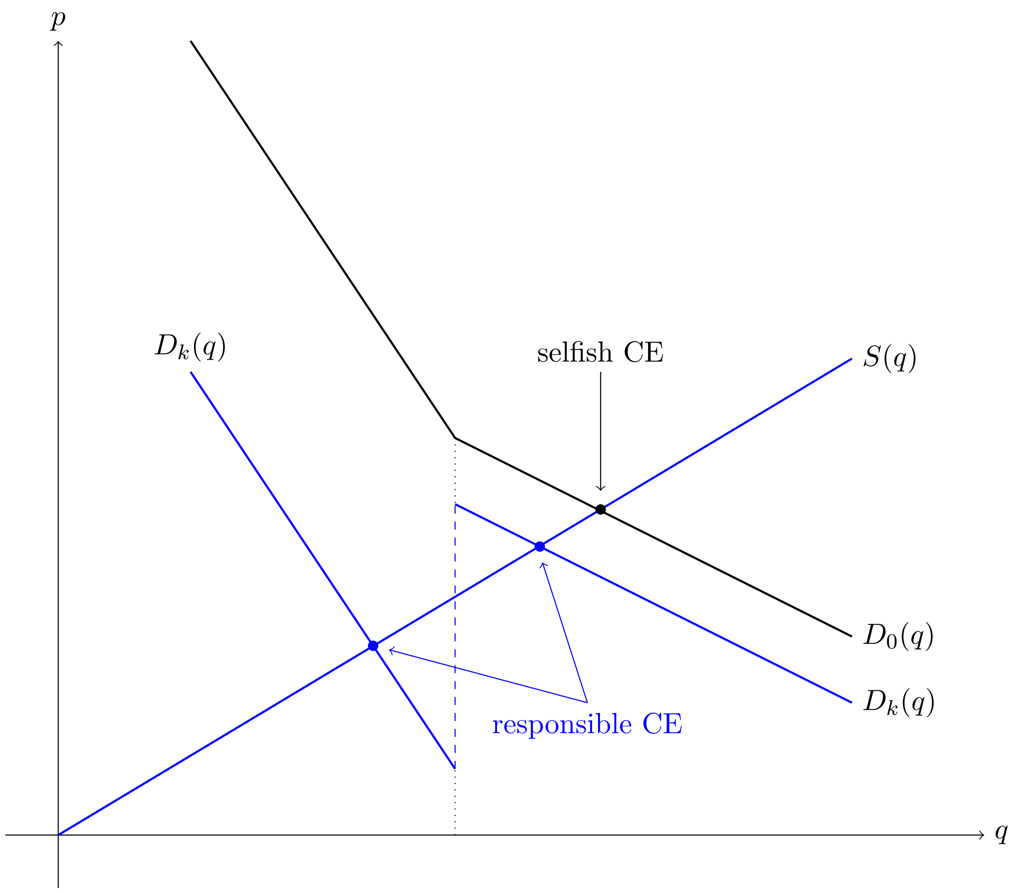


Figure 3: Illustration of Competitive Equilibrium and Multiple Equilibria

A.3 Microfoundation I

We consider an economy with $I+1$ identical consumers and I identical suppliers. Consumer i's utility is

$$U = Ac_i - \frac{1}{2}Bc_i^2 - pc_i - k\sum_j c_j,$$

where c_i denotes consumer i's consumption and $A, B > 0$. The supply of each seller is $S(p) = sp$ with $s > 0$. In addition, there is a shock s_0 to total supply, where s_0 is a non-degenerate random variable whose support includes 0. Hence, total supply is $s_0 + Isp$.

The timing of the game is as follows. First, each consumer submits a (weakly) decreasing demand schedule $c_i(p)$. Then, given all submitted schedules, the market-clearing price is determined. Since supply is linear with a strictly positive slope and demand is weakly decreasing, the market-clearing price exists and is unique. Finally, outcomes are determined and utilities are realized.

We look for linear symmetric Nash equilibrium in which all consumers submit the same linear demand schedule $c_i(p) = a - bp$ with $a,\, b \geq 0$.

Proposition 10. *There is a unique linear symmetric Nash equilibrium. In this equilibrium, each consumer chooses the schedule $c_i(p) = a - bp$, where*

$$b = \frac{-(IBs - I + 1) + \sqrt{\Delta}}{2IB}, \text{ with } \Delta = (IBs - I + 1)^2 + 4BI^2 s, \text{ and}$$
$$a = \left(A - \frac{ks}{b+s}\right) \cdot \frac{1}{B + \frac{1}{I(b+s)}}.$$

We now use Proposition 10 to provide a microfoundation for Definition 1. Notice that in Proposition 10, the parameters describing the equilibrium, a and b, are functions of I, so that we can write them as $a(I)$ and $b(I)$. We define:

Definition 2. A limiting Kyle equilibrium is a strategy profile in which a consumer chooses $c_i(p) = a_\infty - b_\infty p$, where $a_\infty = \lim_{I\to\infty} a(I)$ and $b_\infty = \lim_{I\to\infty} b(I)$.

We also say that A and B quadratically approximate the utility function $u(\cdot)$ around c if $B = -u''(c)$ and $A - Bc = u'(c)$.

Proposition 11. *The following are equivalent.*

I. The quantity $q^* > 0$, price $p^* > 0$, and consumer price responsiveness $q_p^* \in \mathbb{R}$ are part of a competitive equilibrium.

II. The pair q^*, p^* constitutes the realized outcome in the limiting Kyle equilibrium of the economy in which consumption utility $u(\cdot)$ is quadratically approximated around q^*, and $s_0 = 0$ (i.e., when $s_0 = 0$, the realized quantity and price are q^* and p^*). In this limiting Kyle equilibrium, consumers' price responsiveness equals $c_i'(p) = -b_\infty = q_p^*$.

Proposition 11 says that there is a one-to-one correspondence between competitive equilibria as we defined in the text for an economy with vanishingly small consumers and Nash equilibria of our game in this section with a diverging number of players. To map the economy with a general $u(\cdot)$ to a game with quadratic utility, we quadratically approximate $u(\cdot)$ around the equilibrium consumption level. The one-to-one correspondence means that the quantity, price, and consumer price responsiveness are the same in the competitive equilibrium of the economy and the limiting Nash equilibrium of the game. We do not explicitly include the market responsiveness q_c^* because it is a function of the supply responsiveness s and q_p^*.

A.4 Microfoundation II

We consider the game as in Section A.3 with the following modifications. Consumer i's utility is $u(c_i)$ — as in the text, i.e., not necessarily quadratic. At the same time, a consumer can only submit a linear demand schedule, i.e., she submits $c_i(p) = a_i - b_i p$, where she can specify $a_i \in \mathbb{R}$ and $b_i > 0$. We consider symmetric pure-strategy equilibria when $s_0 = 0$, but within that class impose a kind of robustness requirement with respect to shocks to supply. This is defined in the following way. Let $\bar{s}_0$ be a random variable with mean 0 that is continuously distributed with support $[-1, 1]$. When optimizing with respect to the strategies of other consumers, consumer i assumes that $s_0 = \varepsilon \bar{s}_0$, where $\varepsilon > 0$ and $E(s_0) = 0$. Consumer i's strategy must be the limit of optimal responses as $\varepsilon \to 0$.

Proposition 12. *The pair a, b, price p, and resulting consumption level $q = a - bp$ constitute a robust equilibrium if and only if market clearing is satisfied with q and p,*

$$0 = u'(q) - p - \frac{1}{I(b+s)} q - k \frac{s}{b+s}, \text{ and}$$
$$b = \frac{1}{(-u''(q)) + 1/(I(b+s))}.$$

Notice that as $I \to \infty$, the conditions of a robust equilibrium approach those of a competitive equilibrium where the price responsiveness of demand is $-b$ and the market responsiveness is $s/(s+b)$. Hence, competitive equilibria are in a sense close to robust equilibria with large I. Our final proposition makes this relationship precise.

Proposition 13. *The following are equivalent.*

I. The quantity $q^* > 0$, price $p^* > 0$, and consumer price responsiveness $q_p^* \in \mathbb{R}$ are part of a competitive equilibrium.

II. There is a sequence of thrice differentiable utility functions u_I and a sequence of robust equilibria $a(I)$, $b(I)$, $p(I)$ for u_I such that as $I \to \infty$, u_I and its derivatives converge uniformly on bounded intervals to u, $p(I) \to p^*$, $[a(I) - b(I)p(I)] \to q^*$, and $[-b(I_n)] \to q_p^*$.

Proposition 13 says that there is a one-to-one correspondence between competitive equilibria and limits of robust equilibria with large I and utility functions close to u.

B Proofs

Proof of Proposition 1. Let us prove that $p(c) \to p^*$ as $I \to \infty$. The rest follows from the derivations in the text. We know that $p(c)$ is determined by $c + ID(p(c)) = IS(p(c))$, while p^* satisfies $D(p^*) = S(p^*)$. We know that $D'(p) < 0$ and $S'(p) > 0$ everywhere, and $\lim_{p\to\infty}(S(p) - D(p)) = \infty$. Now fix some c, $S(p(c)) - D(p(c)) = c/I \to 0$ as $I \to \infty$, hence by continuity $p(c) \to p^*$.

This proves the proposition.

Proof of Observation 1. Define the following:

$$q(p) = S(p)$$
$$q_p(p) = 1/u''(q(p))$$
$$q_c(p) = s/(s - q_p(p))$$

Thus all the equilibrium conditions hold by construction for p, $q(p)$, $q_p(p)$, and $q_c(p)$ except for the consumer's utility maximization. Note that at $p = 0$, we have $u'(0) > k$ (by assumption), so $u'(0) > 0 + kq_c(0)$, since $q_c(p) \leq 1$ for any p. Clearly, as $p \to \infty$, we have that $q(p) \to \infty$, and that $u'(q(p))$ is eventually strictly less than p. Hence for sufficiently large p, we have $u'(q(p)) < p + kq_c(p)$. Therefore by continuity, there is some $p^* > 0$ s.t. $u'(q(p^*)) = p^* + kq_c(p^*)$. Hence $p^* > 0$, $q^* = q(p^*) > 0$, $q_c^* = q_c(p^*) = s/(s - q_p^*) > 0$, $q_p^* = q_p(p^*) = 1/u''(q^*) < 0$ is an equilibrium.

Further note, that the quantity consumed must be strictly less than that under a selfish equilibrium. Suppose, by contradiction, that it is not, and we have a selfish equilbrium with equilibrium price and quantity p_0^* and q_0^*. So we assume $q_0^* \leq q^*$. Then supply must be larger, which requires that prices must be larger, so that $p_0^* \geq p^*$. But then from the first order condition we have $u'(q^*) = p^* + kq_c^* > p^* \geq p_0^* = u'(q_0^*)$, which implies that $q^* < q_0^*$, a contradiction.

This proves the observation.

Proof of Proposition 2. The social welfare is:

$$u(q) - \int_0^q S^{-1}(x)dx - Kq = u(q) - \int_0^q \frac{x}{s}dx - Kq$$

since $S(p) = s \cdot p$. The socially optimal quantity q^{FB} satisfies the following first order condition:

$$u'(q^{FB}) = \frac{q^{FB}}{s} + K$$

On the other hand, from the definition of competitive equilibrium, the consumer's FOC is

$$u'(q^*) = p + kq_c^* = \frac{q^*}{s} + kq_c^*$$

with $q_c^* = s/(s - 1/u''(q^*))$. Since $u''(\cdot) \leq 0$, q_c^* is in the range $[0, 1)$: it is equal to 0 when $s = 0$ (fixed supply), and converges to 1 as s goes to infinity — for perfectly elastic supply, as would be the case when the marginal cost of production is constant.

We now show that $q^* > q^{FB}$. Suppose by contradiction that $q^* \leq q^{FB}$. Then $u'(q^*) \geq u'(q^{FB}) = \frac{q^{FB}}{s} + K \geq \frac{q^*}{s} + K \geq \frac{q^*}{s} + k > \frac{q^*}{s} + kq_c^*$, where the last inequality holds since $q_c^* \in [0, 1)$.

Proof of Observation 2. Choosing the first-best consumption level is obviously a dominant uniquely optimal strategy.

Proof of Proposition 3. Substituting the equilibrium conditions 1, 3, and 4 into condition 2, we obtain:

$$u'(sp) - p = k\frac{su''(sp)}{su''(sp)-1} \tag{10}$$

Let $g(p) \equiv u'(sp) - p$ and $h(p) \equiv \frac{su''(sp)}{su''(sp)-1}$. Then a necessary and sufficient condition for an (interior) equilibrium is $g(p) = k \cdot h(p)$. Let us assume that $u'''(c) > 0$, $u''(c) < 0$, and $u''(c) \to 0$ as $c \to \infty$. (Note that this rules out the case of quadratic utility, which does not allow for multiple equilibria.)

Note the following derivatives:

$$\begin{aligned} h'(p) &= \frac{d}{dp}\left(1 + \frac{1}{su''(p)-1}\right) = -\frac{s^2u'''(sp)}{(su''(sp)-1)^2} \\ g'(p) &= su''(sp) - 1. \end{aligned} \tag{11}$$

Then the following Lemma holds:

Lemma 1. *Fixing some $k > 0$ and $s > 0$, with $u'(0) > k$, if there is an equilibrium at $p^* > 0$ s.t. $k \cdot h'(p^*) < g'(p^*)$, then there is (i) an equilibrium at p^+ with $p^+ > p_*$ and (ii) an equilibrium at p^- with $p^- \in (0, p^*)$.*

Proof. Since $u''(c) < 0$ and $u''(c) \to 0$ as $c \to \infty$, $h(p) > 0$ and $h(p) \to 0$ as $p \to \infty$. Second, $g'(p) = su''(sp) - 1 < -1$ for all p, so that $g(p) \leq g(0) - p \to -\infty$ as $p \to \infty$. Hence, $h(p) > g(p)$ and therefore $k \cdot h(p) > g(p)$ for sufficiently large p.

Part i). Take the equilibrium at price p^*, so that $k \cdot h(p^*) = g(p^*)$. Since $k \cdot h'(p) < g'(p)$, we have $k \cdot h(p^* + \varepsilon) < g(p^* + \varepsilon)$ for sufficiently small $\varepsilon > 0$. Since for sufficiently large p, we have $k \cdot h(p) > g(p)$, this implies that there is some $p^+ > p^* + \varepsilon$ s.t. $k \cdot h(p^+) = g(p^+)$ – and hence there is another equilibrium at $p^+ > p^*$.

Part ii). Since $u'(0) > k$, we have that $g(0) = u'(0) > k \geq kh_0$, where $h_0 = \lim_{x \to 0^+} h(x) \in [0, 1]$. Therefore $g(p) > k \cdot h(p)$ for p sufficiently close to 0. Since $k \cdot h'(p) < g'(p)$, we have $k \cdot h(p - \varepsilon) > g(p - \varepsilon)$ for sufficiently small $\varepsilon > 0$. So by continuity there is some $p^- \in (0, p^*)$ s.t. $k \cdot h(p)^-) = g(p^-)$, and there is an equilibrium at p^-. □

Now suppose that we have a given k (with $k > 0$) and s with a competitive equilibrium

(q^*, p^*, q_p^*, q_c^*). Since this is an equilibrium, it must satisfy equation 10, which puts constraints on $u'(q^*)$ and $u''(q^*)$ but leaves $u(\cdot)$ otherwise unconstrained.

Next let us note that if we can show that $u'''(q^*)$ is sufficiently large implies that $k \cdot h'(p^*) < g'(p^*)$, then we are done. In that case, by Lemma 1 we know that there must be a strictly higher equilibrium at price $p^+ > p^*$ and a strictly lower equilibrium at $p^- < p^*$. At the strictly higher equilibrium, we have that $q^+ = sp^+ > sp^* = q^*$; that $|u''(q)|$ is strictly decreasing in q, so that $|q_p^+| = 1/|u''(q^+)| > 1/|u''(q^*)| = |q_p^*|$; and that $q_c^+ = \frac{s}{s-q_c^+} = \frac{s}{s+|q_c^+|} < \frac{s}{s+|q_c^*|} = \frac{s}{s-q_c^*} = q_c^-$. The inequalities for the case $p^- < p^*$ are obtained similarly.

Therefore, we now show that that when $u'''(p^*)$ is sufficiently large, then $k \cdot h'(p^*) < g'(p^*)$. Using equations 11, we have:

$$k \cdot h'(p) < g'(p) \iff k \left(\frac{s^2 u'''(sp)}{(su''(sp) - 1)^2} \right) < su''(sp) - 1 \iff u'''(sp) > -\frac{(su''(sp) - 1)^3}{ks^2}$$

which completes the proof of Proposition 3.

Proofs of Propositions 4 and 5. First, we give the formal definition of the equilibrium with policy:

Definition 3. A competitive equilibrium with a permit-supply policy consists of a quantity $q^* > 0$, consumer price $p^* > 0$, permit fee $\tau^* > 0$, consumer price responsiveness $q_p^* \in \mathbb{R}$, and market responsiveness $q_c^* \in \mathbb{R}$ that satisfy the following conditions.

1. a. Supply equals q^*: $q^* = S(p^* - \tau^*)$.

   b. The market for permits clears: $\pi S(p^* - \tau^*) - (1 - \pi)\tau^* + \pi_0 = 0$.

2. Demand equals q^*: $u'(q^*) = p^* + k \cdot q_c^*$.

3. Market responsiveness is consistent with the responsiveness of consumers and net supply: $q_c^* = s_{\text{net}}/(s_{\text{net}} - q_p^*)$, where $s_{\text{net}} \equiv \frac{(1-\pi)s}{(1-\pi)s+\pi s}$ based on Equation (7).

4. Consumer price responsiveness is consistent with optimization: $q_p^* = 1/u''(q^*)$.

For a general policy, including our permit-supply policies, we can define an equilibrium with policy by p^*, q^*, q_c^*, q_p^*, τ^* that satisfy the equations in Definition 3, but subject to different policy constraints. Let us label the two policy types that we want to compare by π' and π.

To show that policy of type π is better than the policy of type π', for every equilibrium under policy of type π' satisfying $u'(0) > k + \tau^*$, we need to find an equilibrium under policy of type π with the same τ^* and strictly lower pollution. Consider an equilibrium under policy of type π', which satisfies

$$\begin{aligned} q^* &= S(p^* - \tau^*) \\ u'(q^*) &= p^* + kq_c^* \\ q_c^* &= \frac{s_{\text{net}}(\pi')}{s_{\text{net}}(\pi') - q_p^*} \\ q_p^* &= \frac{1}{u''(q^*)} \end{aligned}$$

where $s_{\text{net}}(\pi') = \frac{d}{dp} S_{\text{net}}(p - \tau^*)$, where the latter is evaluated under policy π' and at price p^*. Let us define the following:

$$\begin{aligned} p(q) &= S^{-1}(q) + \tau^* \\ q_c(q) &= \frac{s_{\text{net}}(\pi)}{s_{\text{net}}(\pi) - q_p(q)} \\ q_p(q) &= \frac{1}{u''(q)} \end{aligned}$$

By construction, q, $p(q)$, $q_c(q)$, $q_p(q)$, and τ^* satisfy all the conditions for equilibrium except utility maximization and the policy constraint.

Lemma 2. *Suppose that $s_{net}(\pi) > s_{net}(\pi')$ holds for every τ^* and p^*. Then there are values q^-, p^-, q_c^-, and q_p^- that, together with τ^* satisfy $q^- < q^*$ and the equilibrium conditions, with the possible exception of the policy constraint.*

Proof. Since $s_{\text{net}}(\pi) > s_{\text{net}}(\pi')$, we have $q_c^* = \frac{s_{\text{net}}(\pi')}{s_{\text{net}}(\pi')-q_p^*} < \frac{s_{\text{net}}(\pi)}{s_{\text{net}}(\pi)-q_p^*} = q_c(q^*)$ and $p^* = p(q^*)$, so

$$u'(q^*) = p^* + kq_c^* = p(q^*) + kq_c^* < p(q^*) + kq_c(q^*).$$

Since $p(0) = \tau^*$, $q_c(0) \leq 1$, and $u'(0) > \tau^* + k$ all hold, we also have

$$u'(0) > \tau^* + k = p(0) + k \geq p(0) + kq_c(0).$$

By continuity of $u'(q)$, $p(q)$, and $q_c(q)$, there exists a $q \in (0, q^*)$ with

$$u'(q) = p(q) + kq_c(q)$$

so that q, $p(q)$, $q_c(q)$, $q_p(q)$, and τ^* satisfy the equilibrium conditions with $q < q^*$, with the possible exception of the policy constraint. □

To show that policies of type π are better than those of type π', it is enough to show that $s_{\text{net}}(\pi) > s_{\text{net}}(\pi')$ and that the policy constraint holds.

For policies satisfying equation (6), when $\pi' > \pi$, then we showed that $s_{\text{net}}(\pi) = \frac{(1-\pi)s}{(1-\pi)+\pi s}$, hence $s_{\text{net}}(\pi) > s_{\text{net}}(\pi')$. Since for every π we can pick π_0 such that the policy constraint holds, this proves Proposition 4.

Similarly for Proposition 5, we have that $\tau^* = \tau_0 + \tau_1 p^*$. Fixing τ^*, for any p^* and τ_1, we can pick τ_0 such that $\tau^* = \tau_0 + \tau_1 p^*$ such that the policy constraint holds. And since $s_{\text{net}}(\tau_1) = \frac{d}{dp}S(p - \tau_0 - \tau_1 p) = s \cdot (1 - \tau_1)$ is strictly decreasing in τ_1 we have $s_{\text{net}}(\tau_1) < s_{\text{net}}(0)$ when $\tau_1 < 0$; and $s_{\text{net}}(\tau_1) > s_{\text{net}}(0)$ when $\tau_1 > 0$. This proves the Proposition, showing that decreasing taxes are better than fixed ones which are better than increasing ones. This proves Proposition 5.

Proof of Observation 3. Obvious.

Proof of Proposition 6. We first define the competitive equilibrium with policy when there is trade. To adapt Definition 3 for this situation, we distinguish between the market responsivenesses of home- and foreign-supplied quantities, q_c^h and q_c^f. These can be calculated similarly to q_c in the

Definitions 1 and 3. Denoting the price responsiveness of demand by q_p, they are

$$\begin{array}{lcc} & q_c^h & q_c^f \\ \text{tax} & \dfrac{s^h}{s^h + s^f - q_p} & \dfrac{s^f}{s^h + s^f - q_p} \\ \text{cap} & 0 & \dfrac{s^f}{s^f - q_p} \end{array} \tag{12}$$

A consumer's effect on the externality is then $q_c^h e^h + q_c^f e^f$, and this is what she takes into account when choosing her consumption. This leads to the following definition of competitive equilibrium with trade:

Definition 4. A competitive equilibrium with trade and permit-supply policy consists of home and foreign quantities $q^{h*}, q^{f*} > 0$, consumer price $p^* > 0$, tax τ^*, a consumer price responsiveness $q_p^* \in \mathbb{R}$, and home and foreign market responsivenesses $q_c^{h*}, q_c^{f*} \in \mathbb{R}$ such that

1. a. $q^{h*} = S^h(p^* - \tau^*)$ and $q^{f*} = S^f(p^*)$.

   b. $\pi S^h(p^* - \tau^*) - (1 - \pi)\tau^* + \pi_0 = 0$.

2. $u'(q^{h*} + q^{f*}) = p^* + k \cdot (e^h q_c^{h*} + e^f q_c^{f*})$.

3. The responsivenesses q_c^{h*} and q_c^{f*} are given by Equation (12).

4. $q_p^* = 1/u''(q^{h*} + q^{f*})$.

From the producer side, Definition 4 can be seen as a mixture between Definitions 1 and 3, where the former applies to the foreign market and the latter to the home market. Hence, foreign suppliers receive the consumer price p^*, while domestic suppliers receive only $p^* - \tau^*$. In addition, the market for permits at home must clear. From the consumer side, equilibrium accounts for the fact that supply comes from two sources. Hence, consumers' total consumption equals total production, and a consumer takes into account her impact on both sources of supply.

We can now prove Proposition 6. Consider a policy-maker who compares a tax and a cap. First we show that, for every τ^* that satisfies $u'(S^f(\tau^*)) > ke^f + \tau^*$, a unique equilibrium exists for both a tax and a cap that leads to such a τ^*; then we compare these equilibria. Let $q^f(p) = S^f(p)$, $q^h(p) = S^h(p - \tau^*)$. Note that due to the quadratic utility, the price responsiveness is constant and equal to $1/u_{cc}$, while the consumer responsiveness in the home and foreign market depend on the type of policy, but are constant with respect to the price p.

At a price $p = \tau^*$, where $q^h(p) = 0$, we have $u'(q^f(p) + q^h(p)) = u'(q^f(\tau^*)) = u'(S^f(\tau^*)) > \tau^* + ke^f = p + ke^f \geq p + k(e^h q_c^{h*} + e^f q_c^{f*})$. Letting p increase also increases the quantity provided and decreases the marginal utility, until a point where that quantity satiates the consumer: $u'(q^f(p) + q^h(p)) = 0$. Denote this price by $\overline{p}$. At the same time, $p + k(e^h q_c^{h*} + e^f q_c^{f*})$ strictly increases, since all the terms except the first are constant in p. So there is a unique $p^* \in (\tau^*, \bar{p})$ such that the $u'(q^f(p^* + q^h(p^*) = p^* + k(e^h q_c^{h*} + e^f q_c^{f*})$ holds. At this price, all other equilibrium conditions hold, so we have a unique equilibrium for each policy with strictly positive supply in the home and foreign market.

By market clearing, we have $q^{h*} = S^h(p^* - \tau^*)$ and $q^{f*} = S^f(p^*)$ for both tax and cap. We have that q^{h*} and q^{f*}, and hence $q^* = q^{h*} + q^{f*}$ and $g^* = e^h q^{h*} + e^f q^{f*}$, are all strictly increasing in p^*. Since τ^* is the same for both cap and tax, the total pollution is therefore lower for the policy that generates the lower p^* and q^*.

From the FOC for utility maximization we have $u'(q^*) - p^* = k(q_c^h e^h + q_c^f e^f)$. A cap is strictly better than a tax if and only if $q^*_{\text{cap}} < q^*_{\text{tax}}$ and $p^*_{\text{cap}} < p^*_{\text{tax}}$, which from the FOC is equivalent to $q^h_{c,\text{cap}} e^h + q^f_{c,\text{cap}} e^f > q^h_{c,\text{tax}} e^h + q^f_{c,\text{tax}} e_f$. Plugging in the values from equation (12), we get that a cap is strictly better than a tax iff:

$$
\begin{aligned}
& e^h q^h_{c,\text{tax}} + e^f q^f_{c,\text{tax}} - e^f q^f_{c,\text{cap}} < 0 \\
\iff & \frac{e^h s^h + e^f s^f}{s^h + s^f - q_p} < \frac{e^f s^f}{s^f - q_p} \\
\iff & (e^h s^h + e^f s^f)(s^f - q_p) < e^f s^f (s^h + s^f - q_p) \\
\iff & - q_p e^h s^h + e^h s^h s^f < e^f s^f s^h \\
\iff & - q_p < s^f \frac{e^f - e^h}{e^h} \\
\iff & - \frac{1}{u_{cc}} < s^f \frac{e^f - e^h}{e^h}
\end{aligned}
$$

and a tax is strictly better if the same inequality holds strictly in the opposite direction.

This proves the proposition.

Notation and Setup for Propositions 7. Let p^c and p^d be the consumer prices for the clean and dirty product respectively. With some abuse of notation, we denote by p the difference between the consumer prices of the two products: $p = p^c - p^d$. We now show that for any such relative price p, there exist unique consumer prices $p^c(p)$ and $p^d(p)$ such that $p = p^c(p) - p^d(p)$, and market clearing holds on the supply side: $S^c(p^c(p)) + S^d(p^d(p)) = 1$.

Fix p, and consider p^d and p^c s.t. $p^c - p^d = p$. As $p^d \to -\infty$, we also have $p^c \to -\infty$ and thus $S^c(p^c) + S^d(p^d) \to -\infty$. Similarly, when $p^d \to \infty$, then $p^c \to \infty$ and hence $S^c(p^c) + S^d(p^d) \to \infty$. Next, as p^d increases, p^c strictly increases. Moreover, when $S^c(p^c) + S^d(p^d) > 0$, either $S^c(p^c)$ or $S^d(p^d)$ is strictly increasing in p^d. Therefore, by continuity, there is a unique p^d solving $S^c(p^c) + S^d(p^d) = 1$, so that $p^c(p)$ and $p^d(p)$ are well-defined.

Market clearing implies that $S^c(p^c) + S^d(p^d) = 1$, so that $s^c p^c + s^d p^d = 1$. Since $p^c - p^d = p$, we have $p^c = p + p^d$, so that $s^c(p + p^d) + s^d p^d = 1$. Taking derivatives with respect to p, we get that $s^c(1 + p^{d\prime}) + s^d p^{d\prime} = 0$, so that $p^{d\prime} = -\frac{s^c}{s^c + s^d}$.

Again abusing notation somewhat, we define $S(p)$ as the net *relative* supply of the clean product relative to the dirty product that is consistent with market clearing and p: $S(p) \equiv S^c(p^c(p)) -$

$S^d(p^d(p))$.

For ease of exposition, we drop the explicit dependence of $p^c(p)$ and $p^d(p)$ on p and write p^c and p^d.

We now derive $S(p)$. Note that $S(p)$ is continuous in p, since the consumer and production prices are continuous in p, and the supply functions are continuous in the consumer prices. Moreover, p^c and p^d are differentiable in p, hence so is $S(p)$.

We know that $p^{d\prime}(p) = -\frac{s^c(1+\psi^c)}{s^c(1+\psi^c)+s^d(1-\psi^d)}$. Moreover, $S(p) - 1 = S^c(p^c) - S^d(p^d) - 1 = S^c(p^c) - S^d(p^d) - (S^c(p^c) + S^d(p^d)) = -2S^d(p^d)$, so $S'(p) = -2s^d p^d$, i.e.

$$S'(p) = 2\frac{s^d s^c}{s^c + s^d} \equiv s. \tag{13}$$

Hence we have that $S(p)$ is linear with slope s.

Definition of Competitive Equilibrium. Since each consumer has unit demand, aggregate demand is uniquely determined by knowing which consumers buy the clean good and which buy the dirty good. Without loss of generality, we describe demand by an indifferent consumer $k^* \in (\underline{k}, \overline{k})$ such that every consumer with $k > k^*$ buys the clean good and every consumer with $k < k^*$ buys the dirty good. To describe equilibrium, we follow the steps outlined in the text.

In Step (a), we introduce the market responsiveness dQ, which denotes by how much the total quantity of a good increases when a consumer switches to it.

In Step (b), we identify an optimality condition for demand given dQ. Since the consumer with k^* is indifferent between the products, the gain in money from switching from clean to dirty must equal the costs from increased externalities. The monetary gain is the equilibrium price difference $p^* = p^c - p^d$, while the cost is $k^*(e^d - e^c)dQ$. So k^* must satisfy

$$p^* = k^*(e^d - e^c)dQ. \tag{14}$$

In Step (c), we solve for dQ in terms of price responsiveness of demand and supply. We denote by $\mathrm{d}k^* \in [0, \infty]$ the price responsiveness of the cutoff consumer k^*. Note that $\mathrm{d}k^* = \infty$ captures the possibility that consumers respond perfectly elastically to price changes. This is relevant because the products are perfect substitutes.

Applying the formula for the quantity effect from Equation (2), we have that $dQ = S'(p^*)/(S'(p^*) - D'(p^*))$, where $S(\cdot)$ is the relative supply of the clean good and $D(\cdot)$ its relative demand. We have that $S'(p^*) = s$, and we know that $D(p^*) = 1 - 2F(k^*)$, so that $D'(p^*) = 2f(k^*)dk^*$. Putting these together, we get the market responsiveness:

$$dQ = \frac{s}{s + 2f(k^*)dk^*}.$$

Finally, in Step (d), we solve for the price responsiveness of demand captured in dk^*. For this, we totally differentiate the indifference equation (14) with respect to p^*, assuming (analogously to Definition 1) that over the infinitesimal range in question, the market responsiveness $\mathrm{d}Q$ is constant. This gives $\mathrm{d}k^* = 1/((e^d - e^c)\mathrm{d}Q)$.

We can now put the above together with the market clearing condition to define competitive equilibrium. To do so, let F denote the distribution function corresponding to f. We define a <u>competitive equilibrium with substitutes</u> as a relative price $p^* \geq 0$, an indifferent consumer $k^* \in (\underline{k}, \overline{k})$, and a consumer price responsiveness $\mathrm{d}k^* \in [0, \infty]$ such that:

1. The market clears: $S(p^*) = 1 - 2F(k^*)$.

2. The consumer $k = k^*$ is indifferent between the two products: $p^* = k^*(e^d - e^c)\mathrm{d}Q$, with

$$\mathrm{d}Q = \frac{s}{s + 2f(k^*)\mathrm{d}k^*}.$$

3. Consumer price responsiveness is consistent with consumer k^*'s indifference:

$$\mathrm{d}k^* = \frac{1}{(e^d - e^c)\mathrm{d}Q}.$$

Lemmas and Preliminary results for Proposition 7. First, we characterize equilibria with $dk^* = \infty$.

Lemma 3. *Suppose that $S(0) \in (-1, 1)$. Then:*

i) there is an equilibrium with $dk^* = \infty$;

ii) in any equilibrium with $dk^* = \infty$, we have $p^* = 0$, $dQ = 0$, and $S(0) = 1 - 2F(k^*)$, and $k^* \in (\underline{k}, \overline{k})$.

Proof. First, suppose $S(0) \in (-1, 1)$. Then let us show that $dk^* = \infty$, $p^* = 0$, and k^* s.t. $1 - 2F(k^*) = S(0) \in (-1, 1)$ is an equilibrium. Since $f(k) > 0$ for all $k \in (\underline{k}, \overline{k})$, $k^* \in (\underline{k}, \overline{k})$ is uniquely determined. Market clearing holds, since $S(p^*) = S(0) = 1 - 2F(k^*)$. We have that $f(k^*) > 0$ and $dQ = \frac{S'(p^*)}{S'(p^*)+2f(k^*)dk^*}$, so that $dQ = 0$. Hence $p^* = k^*(e^d - e^c)dQ$ holds, since $p^* = 0$ and $dQ = 0$. This proves part i).

Suppose that we have an equilibrium with $dk^* = \infty$. Then by consumer price responsiveness, we have $dk^* = \frac{1}{(e^d - e^c)dQ}$, hence $dQ = 0$. For the indifferent consumer, we have $p^* = k^*(e^d - e^c)dQ$, so $p^* = 0$. By market clearing, we must have $S(p^*) = S(0) = 1 - 2F(k^*) \in (-1, 1)$, which also pins down $k^* \in (\underline{k}, \overline{k})$. This proves part ii).

□

Next we look at equilibria that satisfy $dk^* < \infty$.

Lemma 4. *Let $\underline{f} \equiv \inf_{[\underline{k}, \overline{k}]} f(k)$ and $\overline{f} \equiv \sup_{[\underline{k}, \overline{k}]} f(k)$. Let $\overline{p} = 1/s^c$.*

i) When $s(e^d - e^c) < 2\underline{f}$, there is no equilibrium with $k^* \in (\underline{k}, \overline{k})$ and $dk^* < \infty$.

ii) When $s(e^d - e^c) > 2\overline{f}$, and $\overline{p} > \underline{k}(e^d - e^c)$, there is a unique equilibrium with $k^* \in (\underline{k}, \overline{k})$ and $dk^* < \infty$. In this equilibrium, $p^* > 0$.

iii) For any equilibrium with $k^* \in (\underline{k}, \overline{k})$ and $dk^* < \infty$, we have that $dQ \leq dQ(\underline{f}) \equiv \frac{s(e^d - e^c) - 2\underline{f}}{s(e^d - e^c)} < 1$.

iv) If $\overline{p}s < \underline{k}s(e^d - e^c) - 2f(\underline{k})$, then there is no equilibrium with dk^* and $k^* \in (\underline{k}, \overline{k})$.

Proof. Since we are looking for equilibria with $k^* \in (\underline{k}, \overline{k})$, there has to be some supply of each product and hence $f(k^*) > 0$. Combining the equilibrium conditions 2 and 3, we get that:

$$\frac{1}{(e^d - e^c)\mathrm{d}k^*} = dQ = \frac{s}{s + 2f(k^*)\mathrm{d}k^*} \tag{15}$$

Multiplying out and rearranging, we obtain:

$$(s \cdot (e^d - e^c) - 2f(k^*))dk^* = s \tag{16}$$

If $s \cdot (e^d - e^c) < 2\underline{f}$, then the left-hand side of (16) is negative, while the right-hand side is strictly positive, so the inequality cannot hold. This proves part i).

We can rearrange equation (16) to obtain

$$\frac{1}{dk^*} = \frac{1}{s}\left(s(e^d - e^c) - 2f(k^*)\right) \tag{17}$$

From condition 3 of the equilibrium, we know that $(e^d - e^c)dQ = 1/dk^*$, so we can replace $(e^d - e^c)dQ$ in $p^* = k^*(e^d - e^c)dQ$ to obtain $p^* = k^*/dk^*$, hence $1/dk^* = p^*/k^*$ which we can plug into equation (17):

$$p^* = \frac{k^*}{s}\left(s \cdot (e^d - e^c) - 2f(k^*)\right) \tag{18}$$

Notice that at price $\overline{p}$, the whole market is served by the clean good, since the prices for clean and dirty of $p^c = 1/s^c$ and $p^d = 0$ satisfy market clearing and $p^c - p^d = \overline{p}$. Hence $S(\overline{p}) = 1$, and for $p \leq \overline{p}$, we have $S(p) = S(\overline{p}) - s(\overline{p} - p) = 1 - s(\overline{p} - p)$, since we showed that $S'(p) = s$.

By market clearing we have $S(p^*) = 1 - 2F(k^*)$, so $1 - s(\overline{p} - p^*) = 1 - 2F(k^*)$, i.e. $\overline{p}s - 2F(k^*) = sp^*$, where we can replace p^* by its value from equation (18):

$$\overline{p}s - 2F(k^*) = k^*(s(e^d - e^c) - 2f(k^*)) \tag{19}$$

Claim: $\overline{p} > \underline{k}(e^d - e^c)$ and $s(e^d - e^c) > 2\overline{f}$ implies that equation (19) has a unique solution with $k^* \in (\underline{k}, \overline{k})$.

Proof of claim: Let $l(k^*) \equiv \overline{p}s - 2F(k^*)$ and $r(k^*) \equiv k^*(s(e^d - e^c) - 2f(k^*))$. Then $l(0) = \overline{p}s > 0$ and $l(\overline{k}) = \overline{p}s - 2$. Since $\overline{p}s = \frac{s}{s^c} = \frac{2s^d}{s^c+s^d} < 2$, we have $l(\overline{k}) = \overline{p}s - 2 < 0$. Next, $r(0) = 0$ and since $s(e^d - e^c) > 2\overline{f} \geq 2f(k^*)$, we have $r(k) > 0$ for $k > 0$.

Now we can show that the curves of $r(k)$ and $l(k)$ must cross exactly once. Since $\overline{p} > \underline{k}(e^d - e^c)$, we have that $l(0) = \overline{p}s > k^*(e^d - e^c)s \geq r(k^*)$ for all $k^* \leq \underline{k}$. When $k^* < \underline{k}$, $F(k^*) = 0$ so $l(k^*) = \overline{p}s - F(k^*) = \overline{p}s = r(k^*)$, so the two curves do not cross for any $k^* < \underline{k}$ (nor do they cross at $\underline{k}$, where $r(k^*)$ jumps discontinuously down). So they cross exactly once for some $k^* \in (\underline{k}, \overline{k})$, since $l(\cdot)$ is continuous and strictly decreasing in that range and ends below 0, while $r(\cdot)$ is continuous and strictly increasing in that range and ends up strictly above 0.

This proves the claim.

Note that in the above argument, if we have $l(\underline{k}) < r(\underline{k})$, then the curve of r lies above that of l for all $k \geq \underline{k}$ and we have no such competitive equilibrium with $k \in (\underline{k}, \overline{k})$, but instead have one with $k < \underline{k}$ if we appropriately define dQ outside of the range $(\underline{k}, \overline{k})$. This proves condition iv).

Plugging the value of k^* where the curves cross into equations (18) and (17) yields equilibrium values for p^* and dk^* that are strictly positive, since $s(e^d - e^c) > 2\overline{f} \geq 2f(k^*)$. This proves part ii).

To prove part iii), start with equation (17) which must hold for an arbitrary interior equilibrium:

$$\frac{1}{dk^*} < \frac{s(e^d - e^c) - 2\underline{f}}{s} \tag{20}$$

Since $dQ = \frac{1}{(e^d - e^c)dk^*}$, this implies that

$$dQ = \frac{1}{(e^d - e^c)dk^*} \leq \frac{s(e^d - e^c) - 2\underline{f}}{s(e^d - e^c)} = dQ(\underline{f}) < 1 \tag{21}$$

□

Proof of Proposition 7. For Part I, we have $S(0) \in (-1, 1)$, since some of both products sell when they go at the same price (since s^d, $s^c > 0$). Thus by Lemma 3, there is an equilibrium with $p^* = 0$, so $p^{c*} = p^{d*}$, and in this equilibrium $dQ = 0$, so consumers are indifferent between the two products.

For part II, we know by Lemma 4 that there is no equilibrium with $dk^* < \infty$ and $k^* \in (\underline{k}, \overline{k})$ if $s(e^d - e^c) < 2\underline{f}$ (which is strictly greater than 0), where $s = \frac{2s^c s^d}{s^c + s^d}$. So there is no other equilibrium in which both products are consumed if $e^d - e^c$ or s^c are sufficiently small.

For part III, by Lemma 4, there is a unique equilibrium with $k^* \in (\underline{k}, \overline{k})$, $dk^* < \infty$, and $p^* > 0$ if $\underline{k}(e^d - e^c) < \overline{p}$ and $s(e^d - e^c) > 2\overline{f}$. Since $\underline{k} = 0$, the first condition always holds. Hence when $e^d - e^c$ is sufficiently large, we have a unique equilibrium of this kind. Since $dk^* < \infty$, we have $dQ > 0$, so consumers with $k > k^*$ strictly prefer the clean product, and since the price $p^* > 0$, this means that $p^{c*} > p^{d*}$.

We prove Part IV directly. Since in any equilibrium, everyone consumes one unit and the two goods are perfect consumption substitutes, consumption utility is the same in any equilibrium, so we can ignore consumption utility when comparing the welfare of equilibria. Social welfare is thus given by:

$$W(q) = -\int_0^q (S^c)^{-1}(x) - (S^d)^{-1}(1-x)dx + Kq(e^d - e^c) - Ke^d \tag{22}$$

where q is the amount of the clean product, so that $1 - q$ is the amount of the dirty product. So $W'(q) = -\left((S^c)^{-1}(q) - (S^d)^{-1}(1-q)\right) + K(e^d - e^c) = -(p^c(q) - p^d(q)) + K(e^d - e^c) = -p + K(e^d - e^c)$, which is strictly positive as long as $p(q) < K(e^d - e^c)$, and maximized at quantity q^{FB} satisfying $p(q^{FB}) = K(e^d - e^c)$.

Suppose we have two equilibria with k^* taking the values $k^*_H < k^*_L$, with $k^*_i \in (\underline{k}, \overline{k})$, so that the consumption of the clean good is higher for the equilibrium with k^*_H than the one with k^*_L.

Then the high equilibrium must be a non-selfish equilibrium with $dQ \in (0, 1)$, so that $p^*_H = k^*_H(e^d - e^c)dQ$. Therefore $p^*_H < k^*_H(e^d - e^c) < \overline{k}(e^d - e^c) \le K(e^d - e^c)$, and hence social welfare is still increasing in the quantity of the clean product, therefore the equilibrium with higher consumption in the clean good has higher social welfare. This proves part IV and thus completes the proof of Proposition 7.

Proof of Proposition 8. The clean good is in fixed supply S^c, the dirty good has perfectly elastic supply at price P^d, and a share $\alpha \in [0, 1]$ of consumers are naive, which means that they perceive the externality from the clean and dirty good to be 0 and 1, respectively.

First notice that because $u'(S^c) > P^d + k$, even consumers who think that the externality from consuming the clean good is 1 will buy more than the clean good alone can provide. Therefore any equilibrium will feature some of both goods.

We look for equilibria in which rational consumers assume that substitution dampening is full: thus the impact in terms of externality is the same whether they buy a unit of the clean or the dirty good. Hence they always buy the cheapest good. They take into account quantity dampening, thinking that for each unit of consumption (no matter which product), they cause an increase of the externality of exactly e^d, which we will show later is equal to 1. The naive consumers on the other hand believe that the clean good has no externality and so strictly prefer it as long as the price premium compared to the dirty good is strictly below k.

This means that the price p of the clean good must be at least P^d. If not, then both the naive and the rational strictly prefer it and do not buy any of the dirty good, but this violates market clearing, since we showed that they consume more than the available supply of the clean good.

Similarly, we must have $p \leq P^d + k$: if $p > P^d + k$, not just rational but also naive consumers strictly prefer the dirty good, which means that no one would buy the clean good. This violates market clearing, given the fixed supply of the clean good.

To prove part I, let $\bar{c}$ be s.t. $u'(\bar{c}) = P^d$, i.e. the amount a naive person consumes of the clean good if it sells at price P^d. Let $\underline{\alpha}$ be s.t. $\underline{\alpha}\bar{c} = S^c$. Then for $\alpha < \underline{\alpha}$, in any equilibrium we must have that $p = P^d$. Suppose not, so that $p > P^d$. Then only the naive consumers prefer buying the clean good, and each of them buys strictly less than $\bar{c}$. But since $\alpha\bar{c} < S^c$, this means that demand is strictly less than supply, thus markets do not clear. We have a contradiction and thus must have $p = P^d$.

The fact that prices are constant and equal to P^d in any equilibrium also means that when one person buys one more unit, they cause the full externality from that unit, which means that $e^d = 1$. Therefore rational consumers consume $\underline{c}$ given by $u'(\underline{c}) = P^d + k$, whereas naive consumers consume $\bar{c}$ given by $u'(\bar{c}) = P^d$.

Social welfare is given by $W \equiv \alpha u(\bar{c}) + (1-\alpha)u(\underline{c}) - P^d q - Kq$, where $q = \alpha\bar{c} + (1-\alpha)\underline{c} - S^c$

is the amount produced that causes externalities (i.e. net of the supply of the clean good). Then social welfare changes with α as follows:

$$
\begin{aligned}
\frac{dW}{d\alpha} &= u(\bar{c}) - u(\underline{c}) - (P^d + K)\frac{dq}{d\alpha} \\
&= \int_{\underline{c}}^{\bar{c}} u'(x)dx - (P^d + K)(\bar{c} - \underline{c}) \\
&< u'(\underline{c})(\bar{c} - \underline{c}) - (P^d + K)(\bar{c} - \underline{c}) \\
&\leq (\bar{c} - \underline{c})(u'(\underline{c}) - (P^d + K)) \\
&= 0
\end{aligned}
$$

where the first inequality holds because u is strictly concave, and the last line holds since $u'(\underline{c}) = P^d + k \leq P^d + K$. Thus social welfare strictly decreases in α.

This proves the first part.

Now suppose that $\alpha > \underline{\alpha}$. This means that at the price of $p = P^d$, the naive consumers all want to buy $\bar{c}$ of the clean good, which exceeds its supply. Therefore any equilibrium will have to feature $p > P^d$. This means that rational consumers will strictly prefer the dirty good, which they perceive as causing the same externality at strictly lower monetary cost.

Let $\overline{\alpha}$ be s.t. $\overline{\alpha}\underline{c} = S^c$. Now consider $\alpha \in (\underline{\alpha}, \overline{\alpha})$. Then let $c(\alpha) = S^c/\alpha$, which is the amount each naive can consume for that α, based on market clearing and the fact that naive consumers all buy the clean good. We have that $c(\alpha) = S^c/\alpha > S^c/\overline{\alpha} = \underline{c}$. Then p must satisfy $u'(c(\alpha)) = p$: if it is larger, then the naive consumers consume too little, if it is smaller naive consumers consume too much; in both cases the market for the clean good doesn't clear. Note that since $c(\alpha) > \underline{c}$, we have that $p = u'(c(\alpha)) < u'(\underline{c}) = P^d + k$, thus the naive consumers strictly prefer to buy the clean good. Hence this is an equilibrium.

That this leads to a welfare improvement can be shown formally by computing $dW/d\alpha$, or as follows. Consider increasing α from α_L to $\alpha_H \in (\alpha_L, \overline{\alpha})$. Let us consider three groups of consumers: the naive consumers (those that are initially naive), the rational consumers (those rational before and after), and the switchers (those initially rational, later naive). The welfare depends only on the consumption of these three groups and total consumption.

Then note that the change from α_L to α_H can be considered as the naive consumers reducing their consumption from $c(\alpha_L)$ to $c(\alpha_H) < c(\alpha_H)$. Each unit that they reduce their consumption goes either towards a net reduction of total consumption or towards allowing a switcher to consume more. The marginal cost to the naive consumers is at most $u'(c(\alpha_H))$, while the marginal benefit to a switcher is at least $u'(c(\alpha_H))$, since they used to consume strictly less. And the marginal social benefit of reducing total consumption is $P^d + K$, but we know that $P^d + K = u'(\underline{c}) > u'(c(\alpha_H))$, hence this change leads to a strict net benefit.

This proves part II.

Finally, consider $\alpha > \overline{\alpha}$. Then the price of the clean good must be $P^d + k$. We showed already that it cannot be larger. If it was smaller, then all the naive consumers would strictly prefer the clean good, which would lead to overdemand. When $p = P^d + k$, the naive consumers are indifferent between the two goods, while the rational consumers still prefer the dirty good. The naive consumers now choose consumption to solve $u'(c) = p = P^d + k$, which is the same as the rational consumers, hence everyone consumes $\underline{c}$. Thus social welfare is the same as when $\alpha = 0$.

This proves part III, and thereby Proposition 8.

Proof of Proposition 9. Let P^d denote the price at which the dirty resource 2 is available at perfectly elastic supply. Since $P^d = 1$, this merely serves the purpose of transparently highlighting the role this price plays in the derivations below.

First, let us show that there is always some use of the dirty resource in equilibrium. We have that $r_2^d = q_2 - r_2^c = q_2 + r_1^c - 1 = q_2 + q_1 - 1$, where q_i is the quantity of good i consumed, and r_2^j is the amount of resource $j \in \{\text{clean}, \text{dirty}\}$ used to produce good 2, while r_1^c is the amount of the clean resource used to produce good 1. Then, since $u_2'(1) > P^d + k$, this means that every consumer buys strictly more than 1 unit of good 2, even when they think that this leads to an externality of 1 – i.e. $q_2 > 1$. Hence $r_2^d = q_1 + q_2 - 1 > q_1$, and since there is some consumption of good 1 (due to $u_1'(0) > P^d + k$), we have that $r_2^d > 0$.

Since $\lim_{c \to \infty} u_2'(c) < 1 = P^d$, the quantity of good 2 demanded at any price equal or larger to P^d is finite even for selfish people, hence for everyone. And since the supply of the good at prices strictly below $P^d = 1$ is 0, the only price consistent with market clearing is P^d.

Next we show that the price of land has to equal P^d in any equilibrium. If it was strictly lower, then the producers of the dirty good would stop using the dirty resource in favor of land, which contradicts the fact that in every equilibrium we have some amount of the dirty resource used. If it was strictly higher, no land would be used in the production of the dirty good. Thus the total use of land would be given by the total consumption of good 1. However, selfish consumers would solve $u_1'(c_1) = p$, and since $u_1'(1) < 1$ and by assumption $p > 1$, we have that $u_1'(1) < p$, so $c_1 < 1$. Thus even selfish consumers wouldn't consume enough to use up all the land, thus at a price of p, demand would be strictly below supply. Therefore we have that the price of land equals P^d in any equilibrium.

Naive consumers maximize the following utility:

$$u_1(c_1^n) + u_2(c_2^n) - P^d c_1^n - P^d c_2^n - k c_2^n,$$

because they perceive the externality cost of 1 for the dirty good (good 2), and of 0 for the clean good (good 1). Thus the FOCs are:

$$u_1'(c_1^n) = P^d$$
$$u_2'(c_2^n) = P^d + k$$

Rational consumers realize instead that the total production of the dirty resource is given by $r_2^d = q_2 + q_1 - 1$ (see above) and that this quantity is strictly positive in any equilibrium. Moreover, they realize that any equilibrium has the same prices, so there is no dampening in terms of quantity consumed by others. Therefore their impact on the total consumption of good i is equal to 1 for both goods, and the impact of total consumption of good i on the externality is 1. Therefore they realize that every unit of consumption of any good causes an increase in the externality of 1.

So the rational consumer solves instead:

$$u_1(c_1^s) + u_2(c_2^s) - P^d(c_1^s + c_2^s) - k(c_1^s + c_2^s)$$

and the FOCs are

$$u_1'(c_1^s) = P^d + k$$
$$u_2'(c_1^s) = P^d + k$$

When $k = K$ and $\alpha = 0$, the rational consumers act optimally, realizing the externality they impose. As we increase α, we replace some rational consumers who consume less of good 1 by naive consumers who consume strictly more of it. Since the marginal benefit to the naive consumers is strictly less than the social cost of the additional externality $P^d + k$, this reduces total welfare.

This completes the proof of Proposition 9.

Proof of Proposition 10. Let us solve for consumer i's best response $c_i(p)$ (a schedule) given given other consumers $j \neq i$ are adopting schedule $c_j(p) = a - bp$. Let us solve for the hypothetical best response if the consumer knew the actual ex post realization s_0. It will turn out that, while the consumer does not know it ex ante, in equilibrium there is a one-to-one mapping between the equilibrium price p and s_0. Thus, since the consumer can condition their consumption on p via their schedule, they effectively can condition their consumption on s_0, hence they can achieve the outcome as if they knew s_0.

For now, suppose the consumer knows what s_0 will be. Then if they submit $c_i(p)$, they know that this will determine via market clearing their own consumption and the equilibrium price:

$$c_i(p(s_0)) + I(a - bp(s_0)) = s_0 + Isp(s_0) \implies c_i(p(s_0)) = s_0 + Ip(s_0)(s + b) - Ia \tag{23}$$

Since we assume that the consumer knows what s_0 will be, they would obtain exactly the same outcome if they submitted a constant schedule equal to $c_i(p(s_0))$ everywhere. This leads to exactly the same consumption for consumer i, the same equilibrium price $p(s_0)$ and hence the same total consumption. Therefore it has the same utility for consumer i. By conditioning on s_0, we can therefore reframe the game as one in which the consumer submits their consumption level x_i.

Let us define $x_i \equiv c_i(p(s_0))$. Then the utility from submitting x_i is

$$U(x_i) = Ax_i - \frac{1}{2}Bx_i^2 - p(s_0)x_i - k\left(x_i + \sum_{j\neq i} c_j\right) \tag{24}$$

Taking derivatives with respect to x_i of equation (23) (noting that $c_i(p(s_0)) = x_i$) and rearranging, we get the price impact of x_i:

$$\frac{dp(s_0)}{dx_i} = \frac{1}{I(b+s)}. \tag{25}$$

Next, we can take the derivative of total consumption with respect to x_i, and use the price impact of x_i to obtain the quantity impact of x_i:

$$\frac{d}{dx_i}\left(x_i + \sum_{j\neq i} c_j\right) = 1 - \frac{d}{dx_i}I(a - bp) = 1 - \frac{Ib}{I(b+s)} = \frac{s}{b+s}.$$

Finally, using these impacts, we can compute the first order condition for the utility maximization:

$$A - Bx_i - p(s_0) - x_i\frac{1}{I(b+s)} - k\frac{s}{b+s} = 0$$
$$\implies x_i = \left(A - \frac{ks}{b+s}\right) \cdot \frac{1}{B + \frac{1}{I(b+s)}} - \frac{1}{B + \frac{1}{I(b+s)}}p(s_0)$$

Thus we have $x_i = \alpha - \beta p(s_0)$ for some constants α and β, with $\beta > 0$ that do not depend on s_0. This provides the best possible utility if the consumer could know s_0 ex ante. It is easy to check — by plugging the value of $x_i = \alpha - \beta p(s_0)$ into the market clearing condition — that $p(s_0)$ is a strictly decreasing linear function of s_0. So $p(s_0)$ has a strictly decreasing (and linear) inverse, which we can write $s_0(p)$.

Now we can show that the consumer can achieve this outcome with the linear schedule $c_i(p) = \alpha - \beta p$. When the consumer ends up paying the equilibrium price p, then it must be the case that the realized s_0 is equal to $s_0(p)$. The realized s_0 must satisfy the following two conditions:

$$c_i(p) = s_0 + Ip(s+b) - Ia$$
$$c_i(p) = \alpha - \beta p$$

Since both equations are linear in p, this uniquely determines s_0 as a function of p, and we know that $s_0(p)$ satisfies both of these equations, so the realized s_0 equals $s_0(p)$. Hence this schedule achieves the same utility as a schedule when the consumer knows the ex post realization of s_0, where the consumer can maximize state by state, hence this schedule is the optimal schedule.

Imposing symmetry, the coefficient on p must equal the slope of other consumers' schedules, b:

$$b = \frac{1}{B + \frac{1}{I(b+s)}} \iff IBb^2 + (IBs - I + 1)b - Is = 0.$$

Notice that one root is positive and one root is negative, hence the positive one is chosen.

$$b = \frac{-(IBs - I + 1) + \sqrt{\Delta}}{2IB}, \text{ where } \Delta = (IBs - I + 1)^2 + 4BI^2 s.$$

This completes the proof.

Proof of Proposition 11. We prove $I \Rightarrow II$ and $II \Rightarrow I$ separately.

Step 1: $I \Rightarrow II$. Assume that the tuple p^*, q^*, q_c^*, q_p^* constitutes a competitive equilibrium.

Since p^*, q^* are part of a competitive equilibrium, condition #2 implies $u'(q^*) = p^* + k \cdot q_c^*$. We quadratically approximate consumer utility around q^*, which implies $B = -u''(q^*)$ and $A - Bq^* = u'(q^*) = p^* + k \cdot q_c^*$.

Condition #3 and condition #4 imply that $q_c^* = \frac{s}{s - q_p^*} = \frac{s}{s - \frac{1}{u''(q^*}}$, so that $A - Bq^* = p^* + k \cdot \frac{s}{\frac{1}{B} + s}$. This is equivalent to

$$q^* = (A - \frac{ks}{\frac{1}{B} + s})\frac{1}{B} - \frac{1}{B}p^*.$$

From proposition 10, we know that optimal demand schedule for these values of A and B is:

$$c_i(p) = a(I) - b(I)p, \text{ where}$$
$$a(I) = A - \frac{ks}{b(I)+s} \cdot \frac{1}{B + \frac{1}{I(b(I)+s)}}$$
$$b(I) = \frac{-(IBs - I + 1) + \sqrt{\Delta}}{2IB} p$$
$$\Delta = (IBs - I + 1)^2 + 4BI^2 s$$

Computing limits as $I \to \infty$, we find that $b(I) \to b_\infty = \frac{1}{B}$, and hence $a(I) \to \left(A - \frac{ks}{\frac{1}{B}+s}\right) \cdot \frac{1}{B}$. Therefore in the limiting Kyle equilibrium the schedule is given by

$$c_i(p) = \left(A - \frac{ks}{\frac{1}{B} + s}\right) \frac{1}{B} - \frac{1}{B} p. \tag{26}$$

Hence the pair (p^*, q^*) is on the demand schedule for the limiting Kyle schedule, i.e. $q^* = c_i(p^*)$. Finally, since market clearing holds in the competitive equilibrium, we have $q^* = S(p^*)$, hence $c_i(p^*) = S(p^*)$, which is the market clearing condition of the limiting Kyle equilibrium when the shock $s_0 = 0$ is realized.

In this limiting Kyle equilibrium, the consumers' price responsiveness is $c_i'(p) = -\frac{1}{B} = \frac{1}{u''(q^*)}$ by the quadratic approximation.

Step 2: $II \Rightarrow I$**.** Suppose the pair (p^*, q^*) constitutes the realized outcome in the limiting Kyle equilibrium of the economy in which consumer utility is quadratically approximated around q^*, and when $s_0 = 0$. Due to the quadratic approximation, we know that $u'(q^*) = A - Bc_i(p^*)$ and $B = -u''(q^*)$. Moreover, the consumers' price responsiveness is $c_i'(p^*) = -b_\infty = -\frac{1}{B} = \frac{1}{u''(q^*}$. Using these equalities, we check that the pair (p^*, q^*) satisfies condition #1 – #4 one by one.

1. condition #1: the market clearing for the Kyle equilibrium with realization $s_0 = 0$ yields $c_i(p^*) = S(p^*) + s_0 = S(p^*)$, so market clearing in the competitive equilibrium, $q^* = S(p^*)$, holds with $q^* = c_i(p^*)$.

2. condition #2: Using the quadratic approximations and the schedule of the limiting Kyle equilibrium from equation (26), we get

$$u'(q^*) = A - Bc_i(p^*) = A - (A - \frac{ks}{\frac{1}{B} + s}) + p = p - \frac{ks}{s + \frac{1}{B}} = p - \frac{ks}{s - \frac{1}{u''(q^*)}} = p - k\frac{s}{s - q_q^*}$$

3. condition #4: the consumers' price responsiveness q_p^* is $c_i'(p) = b_\infty = -\frac{1}{B} = \frac{1}{u''(q^*)}$.

4. condition #3: note that setting $q_c^* = \frac{s}{s - q_p^*}$, this is consistent with condition #2 and #4.

This completes the proof.

Proof of Proposition 12. We consider consumer i's strategic situation when all other consumers are choosing the pair a, b. Given all consumers' strategies, the realized price is a function of s_0, which we denote by $p(s_0)$. As a reminder, $s_0 = \varepsilon \bar{s}_0$, where $\bar{s}_0$ is continuously distributed with support $[-1, 1]$ and with mean 0. Similarly, consumer i's consumption is a function $c_i(s_0)$. With some abuse of notation, we redefine a_i as consumer i's consumption level when $s_0 = 0$. (We cannot, and do not, at the same time redefine the a in other consumers' strategies.) This means that $c_i(s_0) = a_i - b_i \Delta p(s_0)$, where $\Delta p(s_0) = p(s_0) - p(0)$. Market clearing has the following implications:

$$p(0) = \frac{a_i + Ia}{I(b + s)}$$
$$\Delta p(s_0) = \frac{-s_0}{b_i + I(b + s)}$$

Furthermore, the total quantity in the market is $s_0 + Isp(s_0) = s_0 + Is(p(0) + \Delta p(s_0))$. For future reference,

$$-b_i \Delta p(s_0) = \frac{b_i s_0}{b_i + I(b + s)} = s_0 - \frac{s_0 I(b + s)}{b_i + I(b + s)}$$

We only consider $a_i, b_i \geq 0$. Define $A(a_i) \equiv u'(a_i) - \frac{2a_i + Ia}{I(b+s)} - k\frac{s}{b+s} = u'(a_i) - p(0) - \frac{1}{I(b+s)} a_i - k\frac{s}{b+s}$. Let $\bar{a}$ be the unique solution to $A(a_i) = 0$:

$$u'(\bar{a}) - \frac{2\bar{a} + Ia}{I(b + s)} - k\frac{s}{b + s} = 0 \tag{27}$$

This is unique since $A(a_i)$ is strictly decreasing at a rate of more than $2/(I(b+s))$, with $\bar{a} > 0$. Note that $A(a_i) < 0$ if $a_i > \bar{a}$ and $A(a_i) > 0$ if $a_i < \bar{a}$.

We want to show that the consumer's expected utility is maximized in some compact and strictly positive region $O = [a_0, a_1] \times [b_0, b_1]$, where $a_0,\ b_0 > 0$ and $a_1,\ b_1 < \infty$. For this, we replace the consumption utility $u(\cdot)$ by $w(\cdot)$, where $w(x) = u(x)$ for all $x \geq \delta$. But for $x < \delta$, we set $w''(x) = \max\{u''(\delta), u''(x)\}$, so that together with $w(\delta) = u(\delta)$, $w'(\delta) = u'(\delta)$, and $u''(\delta) = w''(\delta)$ we have a well-defined function w, s.t. w is twice continuously differentiable, and since $w''(x) \geq u''(x)$ (alongside the boundary conditions), we have that $w(x) \geq u(x)$.[31] Let EU and EW denote the expected utility under u and w respectively.

Lemma 5. *Let $\delta < \frac{1}{2}a_0$ and $\varepsilon < \frac{1}{2}a_0$, then if the expected utility with w achieves any global maximum in $O = [a_0, a_1] \times [b_0, b_1]$ with $a_0,\ b_0 > 0$ and $a_1,\ b_1 < \infty$, then the expected utility under u also achieves any global maximum in O.*

Proof. Suppose that w achieves its maximum on O. Since the optimal $(a_i, b_i) \in O$, we have that the utility is given by

$$EW = E_{s_0}[w(c_i(s_0)) - p(s_0)c_i(s_0) - kIsp(s_0)]$$

We have that $c_i(s_0) > \delta$ if and only if $a_i + \frac{b_i s_0}{b_i + I(b+s)} > \delta$, that is if and only if

$$s_0 > -(a_i - \delta)(\frac{b_i + I(b+s)}{b_i}) \iff |s_0| < |a_i - \delta|(\frac{b_i + I(b+s)}{b_i})$$

The right-hand side is strictly larger than $|a_i - \delta|$, which is strictly larger than $(a_0 - \delta) > \frac{1}{2}a_0$, while $|s_0| \leq \varepsilon$. Thus a sufficient condition for $c_i(s_0) > \delta$ for any schedule chosen from O is that $\varepsilon < \frac{1}{2}a_0$.

Therefore $EW = EU$ for any $(a_i, b_i) \in O$. Hence if some $(a_i, b_i) \in O$, maximizes EW, then it clearly maximizes also EU. Suppose this is not the case, so we have some $(a_i, b_i) \in O$ leading to schedule $c_i(\cdot)$ maximizing EW and some $(a_j, b_j) \notin O$ leading to schedule $c_j(\cdot)$ maximizing EU. But since $w(c) \geq u(c)$, we have $EW(c_j) \geq EU(c_j)$, which since it maximizes EU implies $EU(c_j) > EU(c_i)$, yet this latter equals $EW(c_i)$ since we just showed that these schedules have the

[31] When $\lim_{x \to x_0} u''(x) \to -\infty$ for some $x_0 \leq 0$ — such as for $u(x) = \log(x)$ — we treat $u''(x) = -\infty$ and thus $w''(x) = u''(\delta)$ for all $x \leq x_0$.

same expected utility. But then $EW(c_j) > EW(c_i)$, contradicting that c_i maximized EW. $\square$

Anticipating that we will use Lemma 5 later on, we now only consider utility functions that are twice differentiable with bounded first- and second-order derivatives. We will show that these have their global maxima inside of some region O, thus we can apply the Lemma to w defined above to show that the original utility function (which might have unbounded derivatives) also achieves its global maxima in that same region. Therefore the utility is well-defined also for negative consumption, which can in principle occur for some consumption schedules.

For steps 1 through 4, we write u for such a utility function with bounded derivatives.

For the remaining steps, assume that $\delta < \frac{1}{3}\bar{a}$.

Step 1: For any b_i, the a_i that maximizes EU lies in $[\bar{a} - \delta, \bar{a} + \delta]$ for $\varepsilon < \delta C$ for some constant C, with δ and C independent of b_i. The consumer's expected utility is

$$\begin{aligned} &E_{s_0}\left[u(c_i(s_0)) - p(s_0)c_i(s_0) - kIsp(s_0)\right] \\ =&E_{s_0}\left[u(a_i - b_i\Delta p(s_0)) - (p(0) + \Delta p(s_0))(a_i - b_i\Delta p(s_0)) - kIs(p(0) + \Delta p(s_0))\right], \end{aligned}$$

where E_{s_0} denotes expectation taken over the distribution of s_0.

The derivative of the consumer's utility with respect to a_i is

$$\frac{\partial EU}{\partial a_i} = E_{s_0}\left[u'(c_i(s_0)) - (p(0) + \Delta p(s_0)) - \frac{1}{I(b+s)}c_i(s_0) - k\frac{s}{b+s}\right].$$

Applying the intermediate value theorem, there is some $\delta(s_0) \in (0, s_0)$ s.t. this equals

$$\begin{aligned} &u'(c_i(0)) - p(0) - \frac{1}{I(b+s)}c_i(0) - k\frac{s}{b+s} + \frac{1}{b_i + I(s+b)}E_{s_0}\left[s_0\left(u''\Big(c_i\big(\delta(s_0)\big)\Big)b_i + 1 - \frac{b_i}{I(b+s)}\right)\right] \\ =&u'(a_i) - p(0) - \frac{a_i}{I(b+s)} - k\frac{s}{b+s} + \frac{1}{b_i + I(s+b)}E_{s_0}\left[s_0\left(u''\Big(c_i\big(\delta(s_0)\big)\Big)b_i + 1 - \frac{b_i}{I(b+s)}\right)\right] \end{aligned} \tag{28}$$

Let $K = \sup_{x \geq 0} |u''(x)|$. Then:

$$\begin{aligned}
|\frac{\partial EU(a_i, b_i)}{\partial a_i} - A(a_i)| &= \frac{1}{b_i + I(s+b)} |E_{s_0}\left[s_0\left(u''\Big(c_i\big(\delta(s_0)\big)\Big)b_i + 1 - \frac{b_i}{I(b+s)}\right)\right]| \\
&\leq \frac{1}{b_i + I(s+b)} E_{s_0}[|s_0|](Kb_i + 1 + \frac{b_i}{I(b+s)}) \\
&\leq \frac{b_i}{b_i + I(s+b)} E_{s_0}[|s_0|](K + \frac{1}{I(b+s)}) + \frac{1}{b_i + I(s+b)} E_{s_0}[|s_0|] \\
&\leq \varepsilon(K + \frac{2}{I(s+b)})
\end{aligned} \tag{29}$$

Now, remembering that $A(a_i)$ is decreasing with slope at least $2/(I(b+s))$, we have that $A(\bar{a}-\delta) \geq A(\bar{a}) + 2/(I(b+s))\delta = 2\delta/(I(b+s))$, and similarly $A(\bar{a}+\delta) \leq -2\delta/(I(b+s))$. Now pick ε_1 s.t.

$$\varepsilon(K + 1 + \frac{1}{I(b+s)}) < \frac{2\delta}{I(b+s)}.$$

Then the derivative of EU with respect to a_i is strictly positive for all $a_i < \bar{a} - \delta$, and it is strictly negative for all $a_i > \bar{a} + \delta$. Hence EU is maximized for some $a_i \in [\bar{a} - \delta, \bar{a} + \delta]$.

Step 2: Let $\varepsilon < \frac{1}{2}\bar{a} - \delta$. For any $a_i \in [\bar{a} - \delta, \bar{a} + \delta]$, the b_i that maximizes EU is achieved for some $b_i \leq \overline{B}$ independent of δ, ε, or a_i. Fix some $a_i \in I_A = [\bar{a} - \delta, \bar{a} + \delta]$. Consider only $\varepsilon < \frac{1}{2}\bar{a} - \delta$. Then $c_i(s_0)$ is strictly bounded away from 0:

$$c_i(s_0) = a_i + \frac{b_i s_0}{b_i + I(b+s)} > \bar{a} - \delta - \varepsilon > \frac{1}{2}\bar{a}$$

Therefore $\bar{u}'' \equiv \max_{[\frac{1}{2}\bar{a}, \frac{3}{2}\bar{a}]} u''(x)$ and $\underline{u}'' \equiv \min_{[\frac{1}{2}\bar{a}, \frac{3}{2}\bar{a}]} u''(x)$ are both well-defined strictly negative numbers. They satisfy $\underline{u}'' \leq u''(c_i(s_0)) \leq \overline{u}''$ for all consumption levels that are possible given $a_i \in I_A$ and $\varepsilon < \frac{1}{2}\bar{a} - \delta$.

Now define $\overline{v}(c) \equiv u(a_i) + u'(a_i)(c - a_i) + \overline{u}'' \frac{(c-a_i)^2}{2}$. Note that $\overline{v}(a_i) = u(a_i)$, $\overline{v}'(a_i) = u'(a_i)$, and $\overline{v}''(c) = \overline{u}'' \geq u''(c)$, therefore:

$$\overline{v}(c) - u(c) = \int_{a_i}^{c} (\overline{v}'(x) - u'(x))dx = \int_{a_i}^{c} \int_{a_i}^{x} (\overline{v}''(y) - u''(y))dydx \geq 0$$

and thus $\overline{v}(c) \geq u(c)$, with $\overline{v}(a_i) = u(a_i)$.

We will next show that there is some $\overline{B}$ s.t. $\overline{v}(0) > \overline{v}(\overline{B})$. The consumer's expected utility $EV(b_i)$ when their consumption utility is $\overline{v}$ is:

$$E_{s_0}[\overline{v}(c_i(s_0)) - p(s_0)c_i(s_0) - kIsp(s_0)]$$

We expand this and, using $E_{s_0}[s_0] = 0$, all the terms linear in s_0 drop out, yielding:

$$\begin{aligned} EV(b_i) =& E_{s_0}\left[\overline{v}(c_i(s_0))\right] - p(0)c_i(0) + E_{s_0}[s_0^2]\frac{b_i}{(b_i + I(s+b))^2} - kIsp(0) \\ =& u(a_i) + \frac{\overline{u}''}{2}\frac{b_i^2}{(b_i + I(s+b))^2}E_{s_0}[s_0^2] - p(0)c_i(0) + E_{s_0}[s_0^2]\frac{b_i}{(b_i + I(s+b))^2} - kIsp(0) \\ =& u(a_i) + \left(\frac{\overline{u}''}{2}b_i + 1\right)\frac{b_i}{(b_i + I(s+b))^2}E_{s_0}[s_0^2] - p(0)c_i(0) - kIsp(0) \end{aligned}$$

Then, since the only terms depending on b_i are those multiplied by $E_{s_0}[s_0^2]$, $EV(b_i) - EV(0)$ equals

$$\left(\frac{\overline{u}''}{2}b_i + 1\right)\frac{b_i}{(b_i + I(s+b))^2}E_{s_0}[s_0^2]$$

Thus $EV(0) > EV(b_i)$ when $\frac{\overline{u}''}{2}b_i + 1 < 0$, or when $b_i > \frac{2}{-\overline{u}''}$. Thus letting $\overline{B} = \frac{2}{-\overline{u}''}$ is such that the global maximizer b_i for EV at this a_i has to satisfy $b_i \leq \overline{B}$. Since $EU(0) = EV(0) > EV(b_i) \geq EU(b_i)$, this shows that the maximum for EU (for a given a_i) is also achieved for $b_i \leq \overline{B}$. Since this bound is independent of $a_i \in I_A$, we have shown that for every $(a_i, b_i) \notin I_A \times [0, \overline{B}]$, there is some point in this region that achieves strictly higher expected utility.

Step 3: Let $\varepsilon < \frac{1}{2}\bar{a} - \delta$. For any $a_i \in [\bar{a} - \delta, \bar{a} + \delta]$, the b_i that maximizes EU is achieved for some $b_i \geq \underline{B}$ independent of δ, ε, or a_i. We now repeat the same argument as above, but this time bounding $u(\cdot)$ via $\underline{v}(c) \equiv u(a_i) + u'(a_i)(c - a_i) + \underline{u}''\frac{1}{2}(c - a_i)^2$. So now $\underline{v}''(c) \leq u''(c)$ for any consumption level that is possible for $a_i \in I_A$. We then find that $EV(b_i) - EV(0)$ is given by

$$\left(\frac{\underline{u}''}{2}b_i + 1\right)\frac{b_i}{(b_i + I(s+b))^2}E_{s_0}[s_0^2]$$

We want to show that the person prefers some strictly positive b_i, i.e. $EV(b_i) - EV(0) > 0$. This holds for every $b_i \in (0, \frac{2}{-\underline{u}''})$, with equality at the corners. Therefore EV achieves a maximum on the interior of $[0, \frac{2}{-\underline{u}''}]$. Denote this maximizer by M, which is independent of a_i and ε, so that $EV(M) > 0$. Notice that $\partial EU/\partial b_i \leq E_{s_0}[u'(c_i(s_0))s_0]\frac{I(b+s)}{(b_i+I(b+s))^2}$, since only the consumption utility term contributes positively (see equation (30) ahead). Since $c_i(s_0)$ is bounded, so is $u'(c_i(s_0))$, so there is some $\overline{u}' = \max_{c=c_i(s_0)} u'(c)$, which holds for all $a_i \in I_A$. So there is some $\lambda < \infty$ such that $\partial EU/\partial b_i \leq \lambda$. Since $EU(M) \geq EV(M) > EV(0) = EU(0)$, there is some minimal $x > 0$ s.t. $EU(x) = EV(M)$, and this x is at least $(EV(M) - EV(0))/\lambda \equiv \underline{B}$ given the function cannot increase too quickly. Hence for every $b_i < \underline{B}$, we have $EU(b_i) < EV(M) \leq EU(M)$, so there is some point $(a_i, b_i) \in I_A \times [\underline{B}, \overline{B}]$ that is larger.

Step 4: global maximum determined by FOCs. By steps 2 and 3, we have shown that EU has a global maximum, achieved on $I_A \times [\underline{B}, \overline{B}]$ for every ε sufficiently small. By continuity and differentiability of the functions involved, this global maximum satisfies the first order conditions.

Since w as defined in the beginning has bounded derivatives, this implies that it achieves its maximum in a region $O \in I_A \times [\underline{B}, \overline{B}]$, where $\overline{B}$ and $\underline{B}$ are both strictly positive and finite and independent of ε. Since EW achieves any global maximum on $O = I_A \times [\underline{B}, \overline{B}]$, by Lemma 5, so does EU. Moreover, any such maximum is characterized by the first order conditions, since u is continuous and differentiable on O

Step 5: taking limits From now on, having applied Lemma 5, we again write u for the original utility function, knowing that any maximum of EU is achieved inside of O. Since the global maximum is characterized by the first order conditions on the compact set $I_A \times [\underline{B}, \overline{B}]$, the best response in the limit is a limit of the best responses and inside this same region. Moreover, by continuity and differentiability, this limit best response satisfies the first order conditions in the limit, which we now compute.

As $\varepsilon \to 0$, we have $c_i(s_0) \to c_i(0) = a_i$, so we see from equation (28) that the first order condition converges to

$$u'(a_i) - p(0) - \frac{a_i}{I(b+s)} - k\frac{s}{b+s} = 0$$

which is the same as equation (27), so that $a_i \to \bar{a}$ as $\varepsilon \to 0$. (Notice that even if there are multiple solutions to the first order condition for $\varepsilon > 0$, they all converge to the same limit.)

Next let us compute the derivative of the consumer's utility with respect to b_i, when a_i satisfies the first order condition at ε:

$$E_{s_0}\left[u'(c_i(s_0))\frac{s_0 I(b+s)}{(b_i+I(b+s))^2} - \frac{s_0}{(b_i+I(b+s))^2}c_i(s_0) - p(s_0)\frac{s_0 I(b+s)}{(b_i+I(b+s))^2} - kIs\frac{s_0}{(b_i+I(b+s))^2}\right]. \tag{30}$$

Multiplying by $(b_i + I(b+s))^2/(I(b+s))$ gives

$$E_{s_0}\left[s_0\left(u'(c_i(s_0)) - \frac{1}{I(b+s)}c_i(s_0) - p(s_0) - k\frac{s}{b+s}\right)\right].$$

Subtracting the left-hand side of (27), which is zero, inside the expectation yields

$$E_{s_0}\left[s_0\left(\left(u'(c_i(s_0)) - u'(\bar{a})\right) - \frac{1}{I(b+s)}(-b_i\Delta p(s_0)) - \Delta p(s_0)\right)\right] + E_{s_0}\left[2s_0\frac{\bar{a}-a_i}{I(b+s)}\right]. \tag{31}$$

Let us focus on the first term in this expression and rewrite it. By the intermediate value theorem, there is some $\delta_1(s_0)$ lying between 0 and s_0 such that the first term in the previous equation equals

$$E_{s_0}\left[s_0\left(u''(c_i(\delta_1(s_0)))c_i'(\delta_1(s_0))s_0 - \frac{1}{I(b+s)}(-b_i\Delta p(s_0)) - \Delta p(s_0)\right)\right].$$

Using $c_i'(s) = \frac{b_i}{b_i+I(b+s)}$, so that $s_0 c_i'(s) = -b_i\Delta p(s_0)$, we can factor out $\Delta p(s_0)$ and substituting for it:

$$E_{s_0}\left[\frac{-s_0^2}{b_i+I(b+s)}\left(u''(c_i(\delta_1(s_0))(-b_i) - \frac{1}{I(b+s)}(-b_i) - 1\right)\right].$$

or $-\epsilon^2$ times

$$E_{\bar{s}_0}\left[\frac{\bar{s}_0^2}{b_i+I(b+s)}\left(u''(c_i(\delta(\epsilon\bar{s}_0)))(-b_i) - \frac{1}{I(b+s)}(-b_i) - 1\right)\right].$$

Note that $-\varepsilon^2$ times the expression in equation (31) equals zero at the optimum. Moreover, when $\epsilon \to 0$, we have $\delta_1(\epsilon\bar{s}_0) \to 0$ and that $a_i \to \bar{a}$. So for the consumer's first order condition with respect to b_i to be satisfied as $\varepsilon \to 0$, the whole expression must converge to 0. The second term converges to 0 since $a_i \to \bar{a}$, thus the first term must converge to 0 too. For this to hold, since the term inside the expectation is a product of a strictly positive number and a number that converges to a constant, this constant must be 0, yielding

$$b_i = \frac{1}{(-u''(a_i)) + 1/(I(b+s))}.$$

Notice that given (a, b), the solution of (a_i, b_i) that satisfies the limit first order conditions is unique. Since there is a global maximum satisfying the first order conditions, the unique solution to these first order conditions must be a local, and hence global, maximum.

This shows that we have a unique best response for given (a, b) characterized by these first order conditions. Finally we impose symmetry by requiring $a_i = a - bp(0)$ (since we transformed a_i to be the consumption amount consumed when the supply shock $s_0 = 0$, i.e. for the price $p(0)$) and $b_i = b$. Let us write q for the equilibrium consumption, which is linked to the equilibrium price p once via the schedules chosen by consumers, $q = a - bp$, and once via market clearing for shocks $s_0 = 0$, i.e. the market clearing condition for $p(0)$, $pI(b+s) = (I+1)a$. We thus obtain that any linear symmetric Nash equilibrium is characterized by $q = a - bp$, market clearing, and the following first order conditions:

$$\begin{aligned} 0 &= u'(q) - p - \frac{q}{I(b+s)} - k\frac{s}{b+s} \\ b &= \frac{1}{-u''(q) + \frac{1}{I(b+s)}} \end{aligned}$$

This proves the proposition.

Proof of Proposition 13.

I $\implies$ II Suppose we have a competitive equilibrium p^*, q^*, and q_p^*, which satisfy the following:

$$\begin{aligned} q^* &= sp^* \\ u'(q^*) &= p^* + k\frac{s}{s - q_p^*} \\ q_p^* &= \frac{1}{u''(q^*)} \end{aligned}$$

Then define $b(I) = -q_p^*$, and $p(I) = p^*$ independent of I. Further define the following

$$\begin{aligned}
q_I &\equiv \frac{I}{I+1} q^* \\
\lambda_I &\equiv \frac{u'(q^*)}{u'(q_I)} \\
\mu_I &\equiv \frac{u''(q^*) + \frac{1}{I(b(I)+s)}}{u''(q_I)} \\
u_I(x) &\equiv \mu_I u(x) + \left((\lambda_I - \mu_I) u'(q_I) + \frac{q_I}{I(b(I)+s)} \right) (x - q_I)
\end{aligned}$$

Then we can check that $a(I) = q_I + b(I)p(I)$, $b(I)$, $p(I)$ is a robust equilibrium. First, market clearing holds, since $(I+1)(a(I) - b(I)p(I)) = (I+1)q_I = Iq^*$, which by market clearing in the competitive equilibrium equals Isp^*. Thus market clearing holds.

Next, $u_I'(q_I) = (\mu_I + \lambda_I - \mu_I)u'(q_I) + \frac{q_I}{I(b(I)+s)} = \lambda_I u'(q_I) + \frac{q_I}{I(b(I)+s)} = u'(q^*) + \frac{q_I}{I(b(I)+s)}$. Using the first order condition of the competitive equilibrium, we can replace $u'(q^*)$ by $p^* + k\frac{s}{s - q_p^*}$, q_p^* by $-b(I)$, and p^* by $p(I)$:

$$u_I'(q_I) = p(I) + k\frac{s}{s + b(I)} + \frac{q_I}{I(b(I)+s)} \implies 0 = u_I'(q_I) - p(I) - \frac{q_I}{I(b(I)+s)} - k\frac{s}{s + b(I)}$$

which is exactly the first of the first order conditions.

Finally, we have that $u_I''(q_I) = \mu_I u''(q_I) = u''(q^*) + \frac{1}{I(b(I)+s)}$. Since $-b(I) = q_p^* = 1/u''(q^*)$, this implies:

$$u_I''(q_I) = -\frac{1}{b(I)} + \frac{1}{I(b(I)+s)} \implies b(I) = \frac{1}{-u_I''(q_I) + \frac{1}{I(b(I)+s)}}$$

which is the second of the first order conditions. Thus we have a robust equilibrium. Moreover, as $I \to \infty$, we have that $u_I(x) \to u(x)$ uniformly on any bounded interval, $p(I) = p^* \to p^*$, $a(I) - b(I)p(I) = q_I \to q^*$, and $-b(I) = q_p^* \to q_p^*$ as required.

II $\implies$ I Suppose that we have a sequence of robust equilibria as stated. By Proposition 12, we know that market clearing holds for each of the equilibria as well as the following equations:

$$\begin{aligned} q(I) &= a(I) - b(I)p(I) \\ (I+1)q(I) &= Isp(I) \\ 0 &= u_I'(q(I)) - p(I) - \frac{q(I)}{I(b(I)+s)} - k\frac{s}{b(I)+s} \\ b(I) &= \frac{1}{-u_I''(q(I)) + \frac{1}{I(b(I)+s)}} \end{aligned}$$

where $u_I \to u$ uniformly on bounded intervals. We have that $p(I) \to p^*$ and $-b(I) \to q_q^*$, and $q(I) \to q^*$. Hence market clearing holds, as $q^* = sp^*$ as $I \to \infty$. Similarly, by uniform convergence, we obtain the following limit first order conditions:

$$\begin{aligned} 0 &= u'(q^*) - p^* - \frac{s}{s - q_p^*} \\ q_p^* &= \frac{1}{u''(q(I))} \end{aligned}$$

which proves that we have a competitive equilibrium after defining $q_c^* = s/(s - q_p^*)$.

This proves the proposition.

□

C Additional material for the empirical study

C.1 Sample

Sampling We recruited respondents in October 2023 using the online survey company Prolific. We recruited respondents from different parts of the Prolific respondent pool in order to approximate the general US population in terms of gender, age, income, and region.

Final Sample Characteristics Table C.1 presents demographic summary statistics for our final sample and compares them to the demographic characteristics of the US adult population.

Exclusion Criteria All exclusion criteria are preregistered. The sample does not contain the following responses: incomplete responses, responses at both extreme 1% tails in the response duration, and duplicate respondents (very rare cases).

Attention Screener Only participants who pass an attention screener at the beginning of the survey can proceed to the main part of the survey.

Attrition A total of 2,358 respondents start the survey. One respondent does not confirm the consent form, and 93 respondents do not access the survey from a desktop computer, fail the attention screener, or do not complete the demographic questions. Hence, 2,264 respondents reach the main part of the survey. Of those, 2,092 (92%) complete the questions on beliefs about dampening, 2,043 (90%) complete the questions on the role of consequences, 2,042 (90%) complete the full survey, and 2,000 respondents (88%) are included in the final survey (see exclusion criteria). Conditional on reaching the main part of the survey and completing the first set of demographic questions, we find that high income and high education significantly predict completing the survey and being included in the final sample. However, both effects are small, our final sample is balanced in terms of income, and our results are robust to using post-stratification weights that correct for the imbalances in education (see sensitivity analyses in the next two subsections).

Survey Duration and Remuneration The survey is relatively short to avoid response fatigue and ensure that respondents are willing to respond carefully to the open questions. The median response duration is approximately 7.5 minutes and most respondents complete the survey within 5 and 12 min (20%-80% quantile range). The standard reward for survey completion is $1.75.

Table C.1: Comparison of the Sample to the American Community Survey (ACS)

Variable	**ACS (2022)**	**Sample**
Gender		
Female	50%	51%
Age		
18-34	29%	32%
35-54	32%	33%
55+	38%	35%
Household net income		
Below 50k	34%	34%
50k-100k	29%	31%
Above 100k	37%	35%
Education		
Bachelor's degree or more	33%	60%
Region		
Northeast	17%	18%
Midwest	21%	22%
South	39%	39%
West	24%	22%
Race and ethnicity		
White	73%	78%
Black or African American	13%	11%
Hispanic/Latino	17%	7%
Asian	7%	8%
Political affiliation*		
Democrat	31%	53%
Republican	29%	21%
Independent	39%	26%
Sample size	1,980,550	2,000

*Data on political affiliation is taken from Chinoy et al. (2023) and based on Gallup surveys from the year 2022 (https://news.gallup.com/poll/15370/party-affiliation.aspx).

Notes: This table presents summary statistics from our sample and compares them to the American Community Survey (ACS) 2022. Respondents can identify with multiple races or ethnicities. We report statistics for the adult US population (18 years and above).

Table C.2: Predictors of Attrition

	Respondent is part of final sample (binary dummy)
Female (binary dummy)	−0.006
	(0.013)
Age (continuous)	0.001
	(0.000)
Income: 50-100k (binary dummy)	−0.001
	(0.018)
Income: 100k+ (binary dummy)	0.029*
	(0.017)
At least Bachelor's degree (binary dummy)	0.048***
	(0.015)
Region: Midwest (binary dummy)	−0.004
	(0.021)
Region: South (binary dummy)	0.001
	(0.019)
Region: West (binary dummy)	0.001
	(0.021)
Constant	0.820***
	(0.029)
Observations	2,264
R^2	0.011

Notes: Results from an OLS regression, robust standard errors in parentheses. The sample includes all respondents who reached the main part of the survey. The outcome variable is a binary indicator that takes the value of 1 for respondents who are included in the final sample of the study. The regressors include various respondent characteristics. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

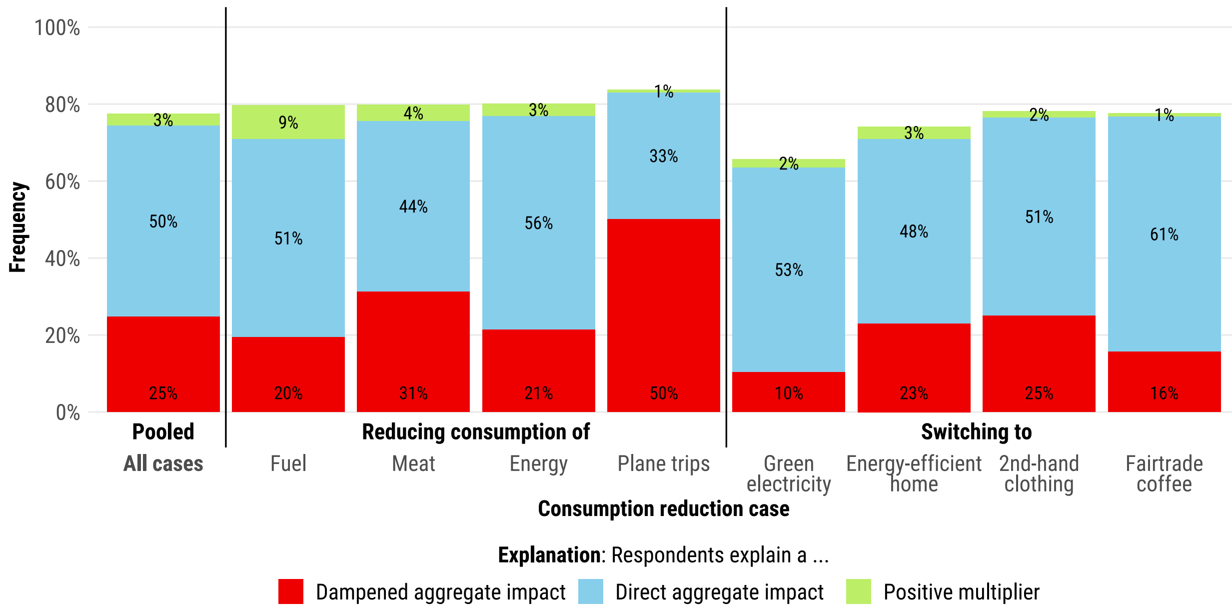
C.2 Additional Results: Beliefs in Dampening

Table C.3: Heterogeneity: Who Predicts Dampening?

	Belief in dampening (partial or full dampening; binary)	**Explanation of dampening** (binary)
	(1)	(2)
Female (binary)	−0.034	−0.008
	(0.022)	(0.019)
Age (continuous, in 10y)	-0.020^{***}	-0.023^{***}
	(0.007)	(0.006)
Income: 50-100k (binary)	0.022	0.035
	(0.027)	(0.024)
Income: 100k+ (binary)	−0.009	0.044^{*}
	(0.028)	(0.024)
At least Bachelor's degree (binary)	−0.001	−0.030
	(0.023)	(0.021)
Politics: Independent (binary)	0.025	0.017
	(0.026)	(0.023)
Politics: Republican (binary)	0.064^{**}	0.049^{**}
	(0.028)	(0.025)
Consumes good (binary)	0.018	−0.008
	(0.027)	(0.023)
Constant	0.328^{***}	0.290^{***}
	(0.056)	(0.050)
Region FE	✓	✓
Case FE	✓	✓
Observations	2,000	2,000
R^2	0.060	0.081

Notes: Results from OLS regressions, robust standard errors in parentheses. The outcome variable is a binary indicator that takes the value of 1 for respondents who predict that their own consumption reduction will lead to a partially or fully dampened reduction in aggregate consumption. The regressors include various respondent characteristics. The dummy "Consumes good" takes a value of 1 if the respondent reports that they regularly consume the good under consideration (fuel: drove 5,000 miles in last 12 months, meat: eat meat at least once per week, energy: annual energy bills of at least $500, flights: took at least one flight in last two years, electricity: annual electricity bill of at least $500, energy-efficient housing: annual energy bills of at least $500, clothing: purchase mostly new clothing, coffee: purchase mostly non-fairtrade coffee). The regressions contain census region and case (e.g., fuel, meat, flights ...) fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure C.1: Beliefs about Dampening: Explanations Found in the Qualitative Text Data



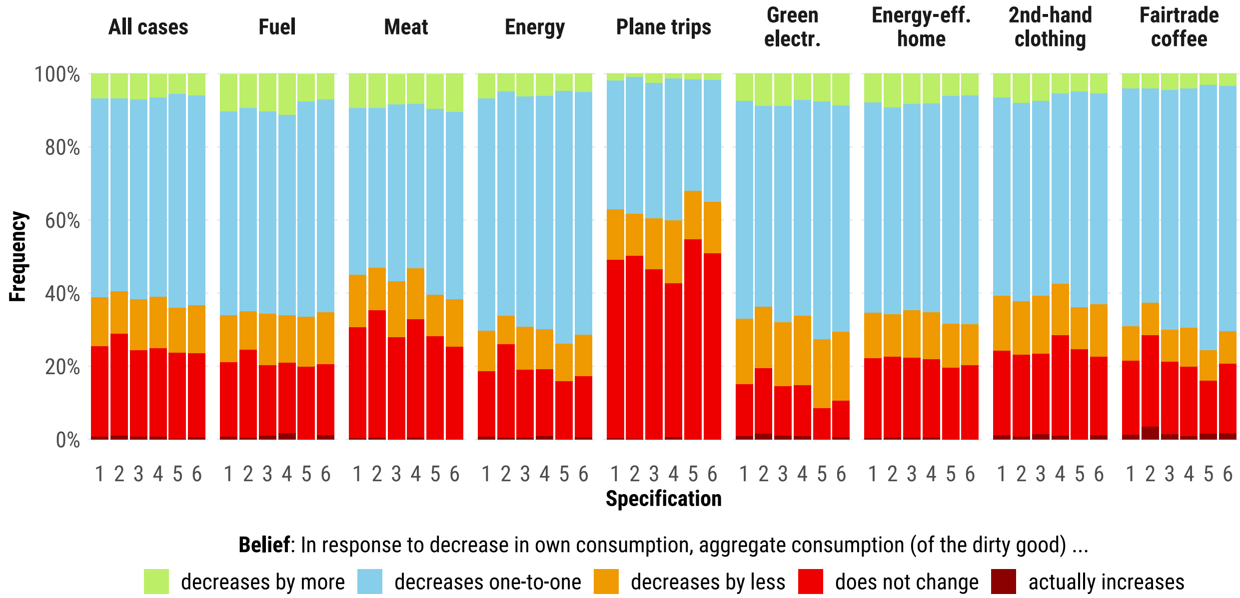
Notes: This figure displays the distributions of consumers' explanations for their beliefs about their own impact on aggregate consumption, as classified based on the open-ended text data. The first column displays results pooled across all eight cases, the other columns present the results for each of the eight cases. See Section C.4.1 for details on the coding scheme.

Sensitivity Tests Figure C.2 replicates the results for a variety of different specifications:

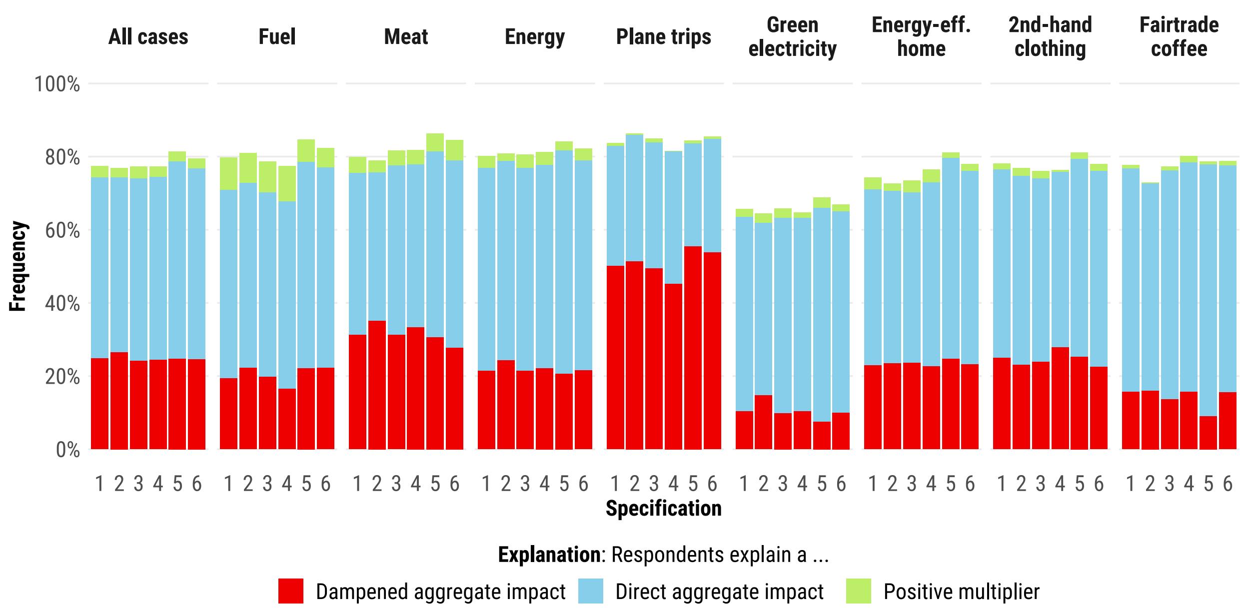
1. Main results.
2. Weighted sample. The weights correct for any imbalances in the characteristics reported in Table C.1. We follow the guidelines of the American National Election Study to calculate the survey weights using a raking procedure (Pasek et al., 2014).
3. We exclude the 20% of respondents with the shortest response duration, which potentially reflects that they paid comparatively less attention to the precise survey instructions.
4. We restrict the analysis to consumers who report regularly consuming the good under consideration.
5. We restrict the analysis to strict consequentialist consumers: valuation ratio $= 0$ (Figure 2).
6. We restrict the analysis to weak consequentialist consumers: valuation ratio < 1 (Figure 2).

Figure C.2: Robustness: Consumers' Understanding of Dampening

(a) **Beliefs about Dampening**



(b) **Explanations**



Notes: This figure displays the distributions of (a) consumers' beliefs and (b) explanations about their own impact on aggregate consumption for six different specifications. The specifications are described on the previous page. The first column displays results pooled across all eight cases, the other columns present the results for each of the eight cases.

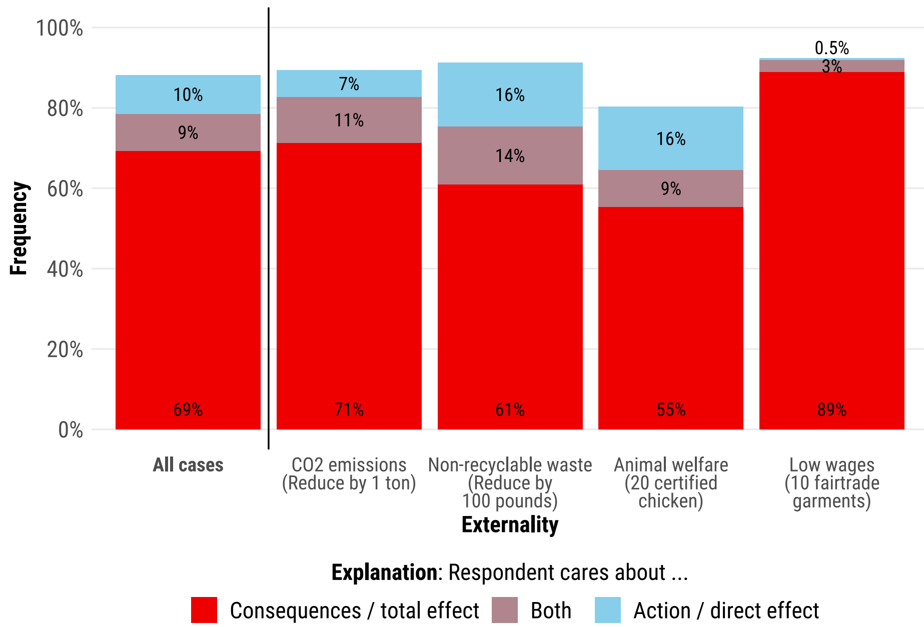
C.3 Additional Results: Nature of Social Concerns

Table C.4: Heterogeneity: Who Cares about Consequences?

	Cares at all?	Valuation data		Explanation data	
	Positive valuation	Strict conseq.	Weak conseq.	Strict conseq.	Weak conseq.
	(1)	(2)	(3)	(4)	(5)
Female	0.041***	0.026	0.011	0.015	0.035*
	(0.015)	(0.023)	(0.020)	(0.022)	(0.019)
Age (in 10y)	−0.014***	−0.028***	−0.036***	−0.020***	−0.024***
	(0.005)	(0.008)	(0.007)	(0.007)	(0.006)
Income: 50-100k	−0.001	0.000	0.005	0.054**	0.058**
	(0.019)	(0.029)	(0.025)	(0.028)	(0.025)
Income: 100k+	−0.004	0.003	0.010	0.064**	0.050**
	(0.020)	(0.030)	(0.025)	(0.028)	(0.025)
Bachelor's degree	0.013	0.015	0.022	0.004	0.010
	(0.017)	(0.025)	(0.021)	(0.024)	(0.021)
Politics: Independent	−0.113***	0.061**	0.030	0.070***	0.047**
	(0.020)	(0.028)	(0.024)	(0.026)	(0.023)
Politics: Republican	−0.132***	0.077**	0.031	0.016	−0.017
	(0.022)	(0.031)	(0.027)	(0.029)	(0.027)
Consumes good	−0.003	−0.015	0.002	−0.014	0.003
	(0.017)	(0.027)	(0.023)	(0.025)	(0.022)
Constant	1.012***	0.507***	0.750***	0.580***	0.693***
	(0.037)	(0.058)	(0.050)	(0.055)	(0.049)
Region FE	✓	✓	✓	✓	✓
Case FE	✓	✓	✓	✓	✓
Observations	2,000	1,714	1,714	1,702	1,702
R^2	0.057	0.088	0.089	0.091	0.081

Notes: Results from OLS regressions, robust standard errors in parentheses. The outcome variables are binary indicators. "Cares at all? Positive valuation" takes the value of 1 if the respondent positively values an effective reduction of the externality. The variables "Valuation data: Strict conseq." and "Weak conseq." take the value of 1 if the respondent has a valuation ratio of 0 or below 1, respectively (see Figure 2). "Explanation data: Strict conseq." and "Weak conseq." take the value of 1 if the respondent expresses (strict: only) consequentialist arguments in the open text data. The regressors include various respondent characteristics. The dummy "Consumes good" takes a value of 1 if the respondent reports that they regularly consume the good under consideration (CO_2: everyone, non-recyclable waste: everyone, animal welfare: eat meat at least once per week & eat mostly non-organic meat, low wages in textile industry: purchase mostly non-fairtrade garments). The regressions contain census region and case (e.g., CO_2, waste, ...) fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure C.3: Social Concerns: Explanations Found in the Qualitative Text Data



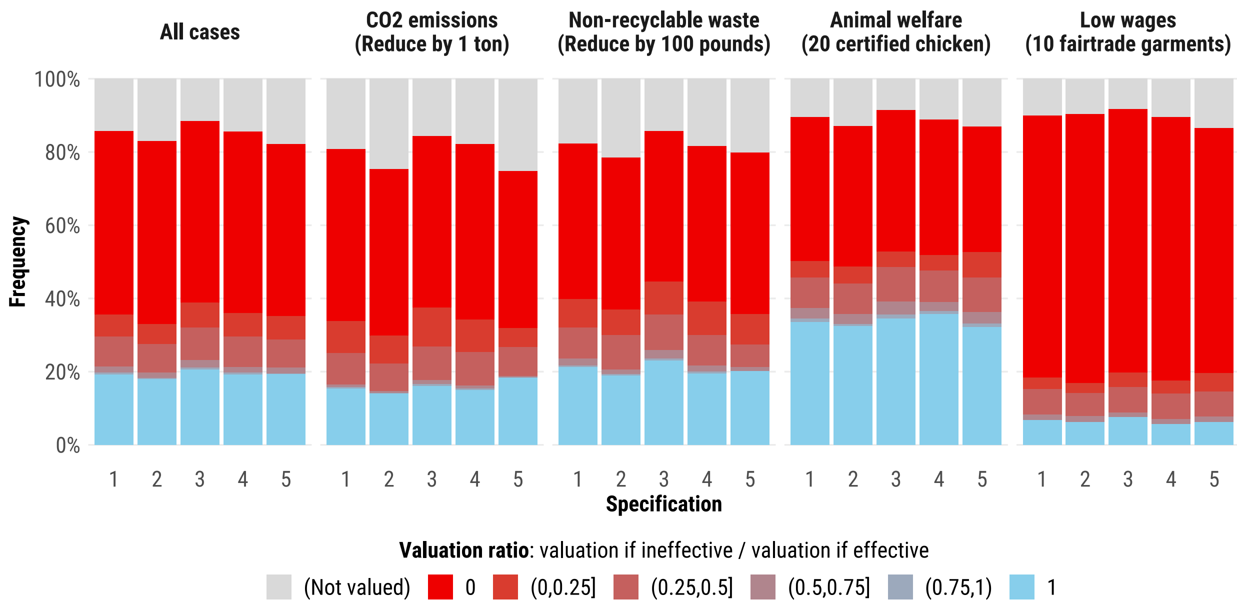
Notes: This figure displays the distributions of consumers' explanations for their valuations of effective and ineffective externality reductions, as classified based on the open-ended text data. Only consumers who positively value an effective reduction of the externality are asked to explain their responses. This figure presents results for these consumers. The first column displays results pooled across all four cases, the other columns present the results for each of the four cases. See Section C.4.2 for details on the coding scheme.

Sensitivity Tests Figure C.4 replicates the results for a variety of different specifications:

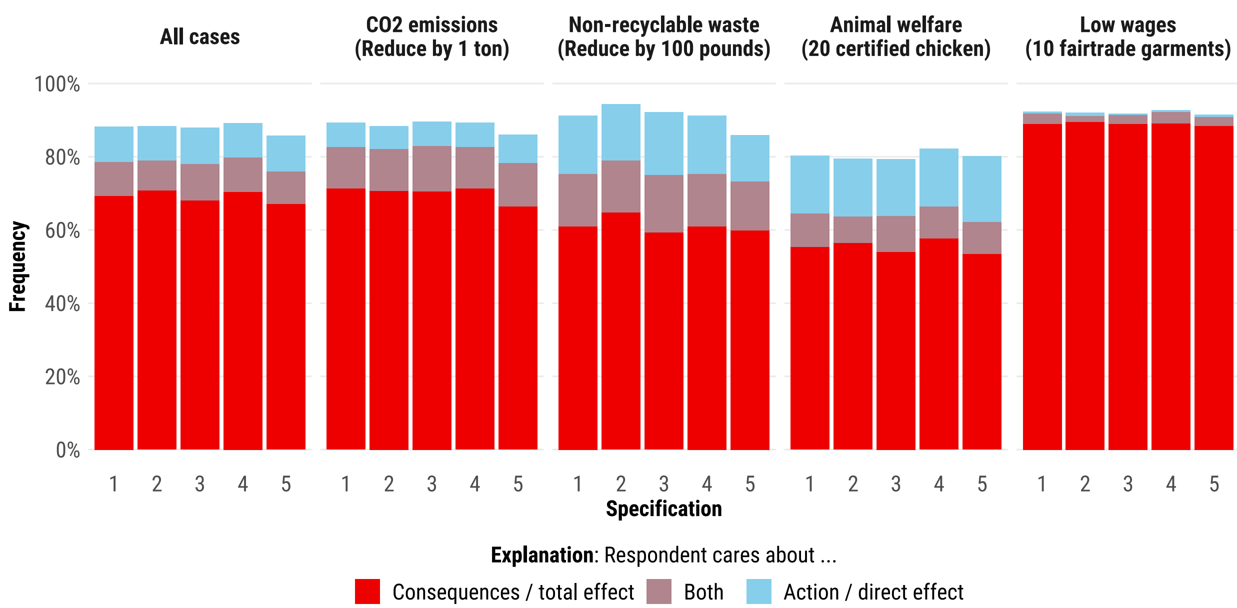
1. Main results.
2. Weighted sample. The weights correct for any imbalances in the characteristics reported in Table C.1. We follow the guidelines of the American National Election Study to calculate the survey weights using a raking procedure (Pasek et al., 2014).
3. We exclude the 20% of respondents with the shortest response duration, which potentially reflects that they paid comparatively less attention to the precise survey instructions.
4. We restrict the analysis to consumers who report regularly consuming the good under consideration.
5. We restrict the analysis the consumers who believe in full or partial dampening (Figure 1).

Figure C.4: Robustness: Consumers' Social Concerns

(a) **Valuation ratios**



(b) **Explanations**



Notes: This figure displays the distributions of (a) consumers' valuation ratios (consumers' valuation of the ineffective externality reduction divided by their valuation of the effective externality reduction) and (b) their explanations for their valuations for five different specifications. The different specifications are described on the previous page. The first column displays results pooled across all four cases, the other columns present the results for each of the four cases.

C.4 Categorization of the Qualitative Text-Data

C.4.1 Explanations of Beliefs in Dampening

Coding Scheme Each open-text response is coded according to a detailed coding scheme that categorizes the line of reasoning that is expressed by the respondent. A summary of the three primary categories within our coding scheme, along with extra examples, can be found in Table C.5. The coding scheme was designed prior to the main data collection and was influenced by initial pilot interviews and the theoretical analysis in the paper. Responses that do not distinctly fall into one of the defined categories are classified into a residual category.

Ancillary Categories The coding scheme includes a few additional ancillary codes for the following type of responses.

1. Responses that are assigned to the residual category and clearly reveal that the respondent did not even attempt to answer the question, e.g., copy-pasted their response from the scenario text or answered by writing about something completely unrelated to the question. This happens in less than 1% of all cases.

2. Responses that are assigned to the residual category and clearly reveal that the respondent misunderstood an important aspect of the question. For example, the response might reveal that the respondent believed that not only themselves but also others reduce their consumption or that the respondent predicted a general trend in global consumption rather than their own effect on global consumption. This applies to approximately 2% of all responses.

3. Among responses assigned to the dampening category, we distinguish between (i) respondents who refer to a consumption increase among other consumers and/or the price mechanism and (ii) respondents who argue that their individual impact is simply too small to shift aggregate quantities, without spelling out the precise dampening mechanism. 38% of all dampening explanations fall into the former category, and 62% fall into the latter category.

Coding Procedure Two research assistants trained in economics coded the text responses. They did not know the goals of our study. We use human coding because machine-based methods still struggle to detect the (often implicit) causal structure in human language.

To deal with the inherent subjectivity of human coding, we adopted two measures. First, we extensively trained the assistants. In the first training session, we taught the coding scheme and discussed many examples. Then, coders practiced on their own, and problematic cases were discussed, reviewed, and corrected in a second session — a process that we repeated one more time in a third and final training session. Second, each response was coded twice — independently by both reviewers. Whenever the two coders disagreed, we looked at the response and made a final decision. This approach ensures that close cases were reviewed a third time. It also allows us to assess the inter-rater reliability.

Inter-Rater Reliability We calculate how often the two independent reviewers assign the same code to a response. If we focus on our four key categories (dampening, direct effect, positive multiplier, residual category), the coders agree in 87% of all cases. If we check consistency for the full coding scheme, including the ancillary codes, they agree in 82% of all cases. These numbers show that our coding scheme has a high degree of reliability.

Table C.5: Overview of the Coding Scheme for Beliefs about Dampening

Explanation	Example
Code: Dampening explanation	
Respondent mentions that others might increase their demand, and/or respondent explicitly refers to the mediating role of prices.	"My own consumption could be offset elsewhere so that it is less than 10 pounds." "I am but one person. Someone may choose to eat more, thus making my reduction less of an impact." "If I didn't buy the conventional coffee, someone else would. Maybe at a lower price, but it would be destroyed. Someone would offer the seller an acceptable price and the buyer would enjoy his coffee."
Respondent argues that they are such a minuscule player on the global market that they have too little influence on aggregate consumption, and/or respondent argues that aggregate supply will not respond to their change in consumption.	"Just because I do not take a seat on the plane does not mean the plane will not fly." "Honestly, I don't think one person is going to impact global consumption of energy. The impact one person has for the entire global consumption is negligible."
Code: Direct effect explanation	
Respondent argues that their own consumption is part of the aggregate consumption so that a change in one variable implies a change in the other variable. Respondents might find this so obvious that they just refer to their own consumption reduction to explain the predicted global consumption reduction.	"It could only reduce the global consumption by the amount I saved." "If I reduce my own consumption it won't make others increase theirs. Therefore it will reduce the global consumption by the amount I reduced my own consumption." "Because my choosing to not buy 40 new garments would reduce the overall global consumption by 40 garments."
Code: Explanation for positive equilibrium multiplier	
Respondent explains why their own consumption reduction leads to an additional consumption reduction by others.	"Even though my reduced energy consumption would not be a significant reduction in the global consumption of energy my decrease would be an example for my family and friends." "I mean there's the 200 [gallons of fuel] we'd personally reduce and then less tankers and ships transporting it as well. So, I'd say more."

Notes: This table provides an overview of the three main categories in our coding scheme for respondents' explanation of their beliefs about dampening.

C.4.2 Explanations of Social Concerns

We follow an analogous procedure to classify the qualitative text data on consumers' reasoning about consequences.

Coding Scheme A summary of the two primary categories within our coding scheme, along with extra examples, can be found in Table C.6. Responses can also be assigned to both codes at the same time. Responses that do not distinctly fall into one of the defined categories are classified into a residual category.

Ancillary Categories The coding scheme includes a few additional ancillary codes for the following type of responses.

1. Responses that are assigned to the residual category and clearly reveal that the respondent did not even attempt to answer the question. This happens in less than 1% of all cases.

2. Responses that are assigned to the residual category and clearly reveal that the respondent misunderstood an important aspect of the question. This applies to less than 1% of all responses.

3. Responses that highlight potential positive consequences even in the scenario that states that respondents' contribution would be ineffective are an exception to this rule. These responses are common (11%) and signify strong consequentialist reasoning. In the main analysis, they are part of the consequentialist code.

4. Among responses assigned to the deontological / warm-glow code, we mark those who refer to the notion of personal responsibility (4% of all responses), choosing the morally right action (2%), or the desire to feel better about one's behavior (4%).

Inter-Rater Reliability We calculate how often the two independent reviewers assign the same code to a response. If we focus on our three key categories (consequentialist, deontological, residual category), the coders agree in 84% of all cases. If we check consistency for the full coding scheme, including the ancillary codes, they agree in 78% of all cases. These numbers show that our coding scheme has a high degree of reliability.

Table C.6: Overview of the Coding Scheme for Concerns for Consequences

Explanation	Example
Code: Consequentialist arguments	
The response reveals that consequences matter to the respondent.	"If it doesn't make a difference, then I don't see why I should pay anything to reduce my CO2 emissions." "I would pay extra to benefit workers only. I would not just pay extra and no one benefits."
Respondents explain a positive valuation for an ineffective reduction arguing that the action could eventually still lead to positive consequences.	"I would hope that the additional cost would maybe go to the organization's overhead and still benefit the workers in some way." "I still think it could have other positive effects. When people see you doing the right thing, they could be motivated to do the same."
Code: Deontological / warm-glow arguments	
The response reveals that the respondent cares about their action even if it has no net positive impact. For example, respondents might argue that they still want to do their own duty, follow a moral principle, or would feel better to at least try.	"It still seems like the right thing to do." "Because ethically it is the correct behavior. Just because the total impact is zero on the corporation's side, it still has impact personally since you are acting ethically. You do not get a pass to act unethically just because it has no effect on some other party." "I would be doing my bit and soothing my conscience. I would sleep better at night. This why I still recycle even though I've read creditable sources that my efforts are for naught [...]."

Notes: This table provides an overview of the two main categories in our coding scheme for respondents' explanation of their concerns about consequences.

C.5 Robustness Studies

In the design of our survey, we prioritize simplicity, a close relationship to relevant real-world settings, a clear mapping between evidence and theory, and we aim to give respondents a chance to share their reasoning. However, some of these design decisions also raise potential concerns that we address in this section.

All robustness studies are preregistered at www.doi.org/10.17605/osf.io/xw8mz.

C.5.1 Beliefs about Effects on Aggregate Production

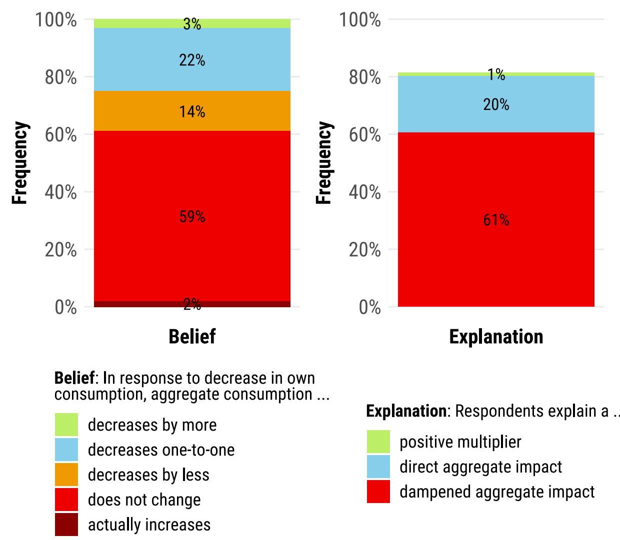
The main study elicits beliefs about aggregate consumption because the dampening mechanism in the model operates via the offsetting consumption increases of other agents. Of course, consumption and production are equivalent in the equilibrium of the model. Empirically, however, people's beliefs about dampening could be influenced by whether they think about the consumption or production side. How people reason about a problem often depends on which feature of the problem they pay attention to (Bordalo et al., 2023).

We explore this possibility in the **Production Robustness Study**. We conducted the study with the survey company Prolific and surveyed 259 US consumers in November 2023. Table C.7 describes the sample's demographic composition. The study builds on the "reducing your fuel consumption" scenario. Instead of asking respondents about their effect on the global total fuel *consumption*, we ask them about their effect on the global total fuel *production*.

The results are summarized in Figure C.5. When asked about their effect on production, even more consumers believe in dampening. The share of consumers who predict full or partial dampening increases to 73%, and the share of consumers who provide explanations consistent with dampening increases to 61%. Many consumers argue that they do not affect aggregate production because producers will not even take notice of their individual consumption decrease.

The fact that consumers reason differently about consumption and production could indicate that many consumers believe that production can exceed consumption with the residual being wasted. As most externalities are created on the production, not the consumption side, this suggests that our focus on consumption in the main study is conservative and tends to underestimate beliefs

Figure C.5: Beliefs and Explanations in the Production Robustness Study



Notes: This figure presents results from the Production Robustness Study. It displays the distributions of consumers' beliefs about their own impact on aggregate consumption (left column) and their corresponding explanations, as classified based on the open-ended text data. See Section C.4.1 for details on the coding scheme of the open-text data. Each response is classified by one research assistant.

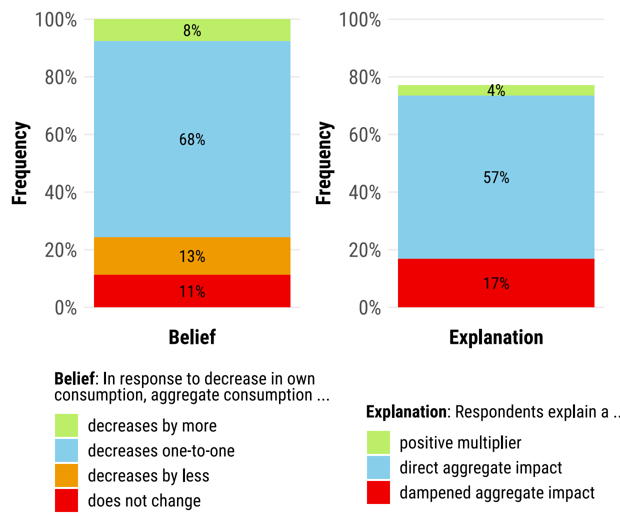
in a dampened impact.

C.5.2 Beliefs with Numeric Elicitation

The main study uses categorical response options. For example, in the scenario where respondents reduce their fuel consumption by 200 gallons, respondents predict whether 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. These categorical responses facilitate the subsequent measurement of open-ended explanations. We want respondents to explain why they think aggregate consumption falls by, say, less than 200 gallons (and not 200 gallons) rather than by 64 gallons (and not 89 gallons).

To ensure that this design choice does not have a strong effect on people's responses, we conduct the **Numeric Response Robustness Study**. We conducted the study with the survey company Prolific and surveyed 250 US consumers in November 2023. Table C.7 describes the sample's

Figure C.6: Beliefs and Explanations in the Numeric Response Robustness Study



Notes: This figure presents results from the Numeric Response Robustness Study. It displays the distributions of consumers' beliefs about their own impact on aggregate consumption (left column) and their corresponding explanations, as classified based on the open-ended text data. See Section C.4.1 for details on the coding scheme of the open-text data. Each response is classified by one research assistant.

demographic composition.

In this study, respondents predict the effect on aggregate consumption in an open numeric response box:

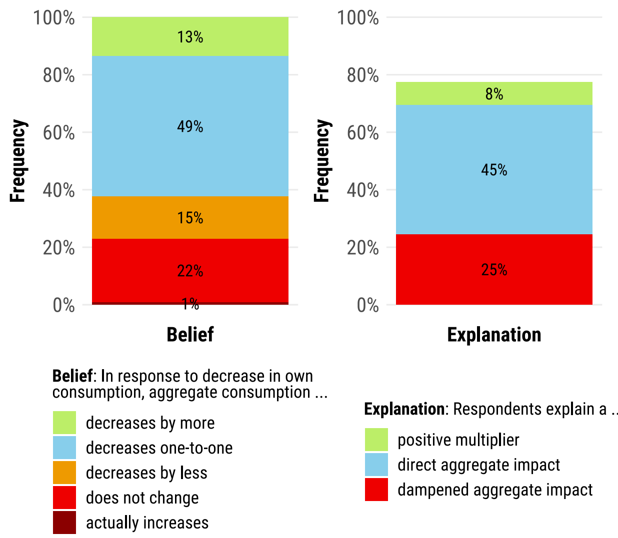
> You reduce your yearly fuel consumption by 200 gallons. This would reduce the yearly total global consumption of fuel by ...
>
> ___________ gallons

The results are summarized in Figure C.6 and qualitatively mirror the results of the main study.

C.5.3 Beliefs with Incentivization Scheme

Since we do not know the exact extent of dampening for the real-world markets and scenarios considered in the main study, we cannot incentivize beliefs in the main study. Fortunately, existing studies often find at most weak differences in the answers to incentivized and non-incentivized

Figure C.7: Beliefs and Explanations in the Incentivized Beliefs Robustness Study



Notes: This figure presents results from the Incentivized Beliefs Robustness Study. It displays the distributions of consumers' beliefs about their own impact on aggregate consumption (left column) and their corresponding explanations, as classified based on the open-ended text data. See Section C.4.1 for details on the coding scheme of the open-text data. Each response is classified by one research assistant.

questions (Stantcheva, 2023). Nevertheless, we design an additional robustness study, the **Incentivized Beliefs Robustness Study**. We conducted the study with the survey company Prolific and surveyed 252 US consumers in November 2023. Table C.7 describes the sample's demographic composition.

Respondents make two predictions. First, they face the standard "reducing your fuel consumption" scenario. Then, they face an additional scenario that describes an introduction of a carbon tax in the United States, leading to a 1 billion ton reduction in US-wide CO_2 emissions. Respondents predict how this change would affect the yearly global emissions of CO_2. The scenario revolves around the issue of "carbon leakage", the concern that climate policies implemented in one country merely shift emissions to other countries instead of reducing them. Carbon leakage has frequently been studied by economists and their common conclusion is that carbon leakage is positive but not full (Dechezleprêtre and Sato, 2017, Grubb et al., 2022). In other words, we know that researchers' "best estimate" would be that the policy reduced the yearly global emissions of CO_2 by *less than*

1 billion tons.

The study employs a probabilistic incentivization scheme. Respondents are informed that

> "You can earn an additional bonus of $2 if your predictions accord with recent research findings in economics. In particular, you will make two predictions. For one prediction, we reviewed the research literature in economics and determined a prediction that is plausible in light of recent research findings. If you make the same prediction, we will transfer a bonus of $2 (or £1.70) to your Prolific account. However, you will not be told which of your two predictions will be tested, so please take both predictions seriously."

This approach allows us to truthfully incentivize both predictions, even though we do not know the correct answer to the first prediction. The procedure is akin to the approach taken by Bardsley (2000).

The results are summarized in Figure C.7 and closely mirror the results of the main study.[32]

[32] We find qualitatively and quantitatively very similar beliefs about dampening for the effect of a US-wide carbon tax. 30% predict dampening of which, however, most respondents (77%) predict partial dampening. This suggests that beliefs in dampening could also have consequences for the political support for climate policies.

Table C.7: Demographic Characteristics of the Samples in the Robustness Studies

Variable	**ACS (2022)**	**Production**	**Numeric Response**	**Incentivized Beliefs**
Gender				
Female	50%	51%	45%	47%
Age				
18-34	29%	38%	40%	45%
35-54	32%	51%	49%	40%
55+	38%	11%	11%	14%
Household net income				
Below 50k	34%	32%	31%	38%
50k-100k	29%	44%	42%	40%
Above 100k	37%	23%	26%	22%
Education				
Bachelor's degree or more	33%	56%	64%	55%
Region				
Northeast	17%	17%	22%	21%
Midwest	21%	19%	19%	21%
South	39%	48%	39%	40%
West	24%	17%	20%	18%
Race and ethnicity				
White	73%	75%	68%	71%
Black or Afric. American	13%	12%	15%	16%
Hispanic/Latino	17%	10%	12%	7%
Asian	7%	10%	7%	10%
Political affiliation*				
Democrat	31%	43%	48%	52%
Republican	29%	26%	21%	19%
Independent	39%	31%	31%	28%
Sample size	1,980,550	259	250	252

*Data on political affiliation is taken from Chinoy et al. (2023) and based on Gallup surveys from the year 2022 (https://news.gallup.com/poll/15370/party-affiliation.aspx).

Notes: This table presents summary statistics from our sample in the Robustness Studies and compares them to the American Community Survey (ACS) 2022. Respondents can identify with multiple races or ethnicities. We report statistics for the adult US population (18 years and above).

C.6 Instructions of Main Study

The complete instructions are available online at https://osf.io/u67wp. The survey begins with a participation information and informed consent form. Respondents who participate on a mobile device are screened out. Next, respondents have to pass an attention check. Subsequently, respondents fill out a block of demographic questions. Then, the main part of the survey begins.

Belief elicitation for the case of fuel consumption

[Respondents are randomized to one out of eight cases. The other cases are displayed below.]

Reducing your fuel consumption

Fuel consumption has a substantial impact on the climate. Transportation fuels account for a significant portion of global greenhouse gas emissions.

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.

What do you think would happen to the yearly total global consumption of fuel?

To repeat, you reduce your yearly fuel consumption by 200 gallons.
How would this affect the yearly total global consumption of fuel?

- It would **decrease** the yearly global consumption of fuel **by more than 200 gallons**.
- It would **decrease** the yearly global consumption of fuel **by 200 gallons**.
- It would **decrease** the yearly global consumption of fuel **by less than 200 gallons**.
- It would **not affect** the yearly global consumption of fuel.
- It would **actually increase** the yearly global consumption of fuel.

Please explain why you chose this response.

[Note: The order of response options is randomly reversed across respondents. If respondents select "decreases (...) by more", "decreases (...) by less", or "actually increases", a follow-up question appears on the same page.]

Follow-up question for “decreases (...) by more”

Please provide your best estimate.
You reduce your yearly fuel consumption by 200 gallons.
This decreases the yearly total global consumption of fuel by ...

- between 201 gallons to 300 gallons
- between 301 gallons to 400 gallons
- between 401 gallons to 500 gallons
- between 501 gallons to 600 gallons
- more than 600 gallons

Follow-up question for “decreases (...) by less”

(Order of response options is randomly reversed.)

Please provide your best estimate.
You reduce your yearly fuel consumption by 200 gallons.
This decreases the yearly total global consumption of fuel by ...

- between 150 gallons to 199 gallons
- between 100 gallons to 149 gallons
- between 50 gallons to 99 gallons
- between 1 gallons to 49 gallons

Follow-up question for “actually increases”

Please provide your best estimate.
You reduce your yearly fuel consumption by 200 gallons.
This increases the yearly total global consumption of fuel by ...

- between 1 gallons to 100 gallons
- between 101 gallons to 200 gallons
- between 201 gallons to 300 gallons
- between 301 gallons to 400 gallons
- more than 400 gallons

Additional consumption cases: Reducing consumption

Reducing your meat consumption

Meat consumption has a significant impact on the climate. According to a recent study, almost 60% of greenhouse gas emissions from food production come from meat alone. Moreover, meat often comes from industrial farming systems that not only subject animals to cramped, stressful conditions but also exact a heavy toll on the environment. On average, an American consumes about 200 pounds of meat per year.

Your consumption of meat is part of the **total global consumption of meat**. We would like to know what you think would happen to the global consumption of meat if you reduced your own meat consumption. Would it make a difference to the total consumption of meat worldwide?

Consider these two scenarios:

> Scenario 1: You eat 200 pounds of meat every year.
>
> Scenario 2: You eat 100 pounds of meat every year.

In contrast to scenario 1, you would permanently reduce your yearly meat consumption by 100 pounds in scenario 2.

Reducing your number of plane trips

Plane trips are a substantial contributor to climate change. They are particularly harmful for the climate because they emit large quantities of greenhouse gases, and they do so at high altitudes, where greenhouse gases have a more potent effect on the world's climate.

Your number of plane trips is part of the **total global number of plane trips**. We would like to know what you think would happen to the global number of plane trips if you reduced your own number of plane trips. Would it make a difference to the total number of plane trips worldwide?

Consider these two scenarios:

> Scenario 1: You take 8 plane trips every year.
>
> Scenario 2: You take 0 plane trips every year.

In contrast to scenario 1, you would permanently reduce your yearly number of plane trips by 8 trips in scenario 2.

Reducing your energy consumption

Household energy consumption is a significant contributor to climate change. In American homes, a variety of appliances and gadgets require energy to operate. About 50% of a household's yearly energy consumption is typically allocated to two main functions: heating and cooling. Energy consumption is measured in kilowatt-hours (kWh). It is not uncommon for US households to consume about 30,000 kWh of energy each year (electricity and gas).

Your energy consumption is part of the **total global consumption of energy**. We would like to know what you think would happen to the global consumption of energy if you reduced your own consumption of energy. Would it make a difference to the total consumption of energy worldwide?

Consider these two scenarios:

> Scenario 1: You consume 30,000 kWh of energy every year.
>
> Scenario 2: You consume 20,000 kWh of energy every year.

In contrast to scenario 1, you would permanently reduce your yearly energy consumption by 10,000 kWh in scenario 2.

Additional consumption cases: Reallocating consumption between close substitutes

Switching from brown to green electricity

You can choose between two types of electricity for your home: green or brown. Green electricity comes from clean sources like solar, wind, and water. Brown electricity comes from fossil fuels like coal and oil, which have a much higher carbon footprint and contribute to climate change. Electricity consumption is measured in kilowatt-hours (kWh). On average, a home in the US uses about 10,000 kWh of electricity each year.

Your consumption of brown electricity is part of the **total global consumption of brown electricity**. We would like to know what you think would happen to the global consumption of brown electricity if you fully switched from brown to green electricity. Would it make a difference to the total consumption of brown electricity worldwide?

Consider these two scenarios:

> Scenario 1: You only use brown electricity. Each year, you use 10,000 kWh of it.
>
> Scenario 2: You fully switch to green electricity. Each year, you use 10,000 kWh of it.

In contrast to scenario 1, you would permanently switch to green electricity in scenario 2.

Moving from an energy-inefficient to an energy-efficient home

Homes vary in their energy efficiency levels. An energy-efficient home is designed to use less energy due to [...]. On the other hand, an energy-inefficient home lacks these features and often wastes energy [...]. Energy consumption is measured in kilowatt-hours (kWh). It is not uncommon for US households to consume about 30,000 kWh of energy each year (electricity and gas).

Your consumption of energy is part of the **total global consumption of energy**. We would like to know what you think would happen to the global consumption of energy if you moved from an energy-inefficient home to an energy-efficient home. Would it make a difference to the total consumption of energy worldwide?

Consider these two scenarios:

> Scenario 1: You live in an energy-inefficient home. Each year, your household consumes 35,000 kWh of energy.
>
> Scenario 2: You move to an energy-efficient home. As a result, your household consumes 25,000 kWh of energy each year.

In contrast to scenario 1, you would permanently reside in an energy-efficient home and consume 10,000 kWh less each year in scenario 2.

Switching from new to second-hand clothing

You have a choice between two types of clothing: new clothing and second-hand clothing. Second-hand clothing refers to clothes that have been previously owned by someone else. By contrast, new clothing is freshly produced and has not been worn by other people before. The production of clothing is highly resource-intensive and clothes are often produced under poor conditions with low wages for laborers. It is not uncommon in the US to buy 40 garments per year.

Your consumption of new clothing is part of the **total global consumption of new clothing**. We would like to know what you think would happen to the global consumption of clothing if you fully switched from new clothing to second-hand clothing. Would it make a difference to the total consumption of new clothing worldwide?

Consider these two scenarios:

> Scenario 1: You only buy new clothing. Each year, you purchase 40 new garments.
>
> Scenario 2: You fully switch to second-hand clothing. Each year, you purchase 40 second-hand garments.

In contrast to scenario 1, you would permanently switch to second-hand clothing in scenario 2.

Switching from conventional to fairtrade coffee

You can choose between two types of coffee: fairtrade or conventional. Fairtrade coffee ensures that farmers receive a fair wage and work in safe conditions. The fairtrade system also encourages sustainable farming practices that are better for the environment. By contrast, conventional coffee often comes from large-scale industrial farming systems that may underpay farmers and usually does not prioritize sustainable farming methods. On average, an American consumes about 10 pounds of coffee beans per year.

Your consumption of conventional coffee is part of the **total global consumption of conventional coffee**. We would like to know what you think would happen to the global consumption of conventional coffee if you fully switched from conventional to fairtrade coffee. Would it make a difference to the total consumption of conventional coffee worldwide?

Consider these two scenarios:

> Scenario 1: You only drink conventional coffee. Each year, you consume 10 pounds of conventional coffee.
>
> Scenario 2: You fully switch to fairtrade coffee. Each year, you consume 10 pounds of fairtrade coffee.

In contrast to scenario 1, you would permanently switch to fairtrade coffee in scenario 2.

Questions on the nature of social concerns

[Respondents are randomized to one out of four cases. The other cases are displayed below.]

Carbon emissions

Now, on a different topic …

The average US American emits about 15 tons of carbon dioxide (CO2) each year. Reducing CO2 emissions often comes at a financial cost. For example, the cost could arise because you buy eco-friendly products, invest in energy-saving solutions, transition to a renewable energy supplier, or acquire carbon offsets. We want to learn what the highest cost is that you would be willing to pay to reduce your personal CO2 emissions.

Please consider two different situations.

Situation 1: Your action has positive consequences

In situation 1, if you reduce your personal CO2 emissions by one ton, the **total global CO2 emissions also decrease by one ton**.

> **Your action**
> You reduce your personal CO2 emissions by one ton.
> **Consequence**
> The total global CO2 emissions decrease by one ton.

In situation 1, how much money would you be willing to pay to reduce your carbon emissions by 1 ton?

$__________

Situation 2: Your action does not have any consequence

Now, please assume that, if you reduce your personal CO2 emissions by one ton, the **total global CO2 emissions do not change**. (For example, this could happen because your reduced emissions make it cheaper for others to emit, and consequently others increase their CO2 emissions by one ton, exactly offsetting your reduction.)

> **Your action**
> You reduce your personal CO2 emissions by one ton.
> **Consequence**
> The total global CO2 emissions do not change.

In situation 2, how much money would you be willing to pay to reduce your carbon emissions by 1 ton?

$__________

[Note: Depending on the answers, open-ended text questions appear on the same page.]

Follow-up question if answers are identical

Please explain why you gave <u>the same answer</u> in the two situations.

Follow-up question if answers differ

Please explain why you gave <u>different answers</u> in the two situations.

Follow-up question if answer is positive in situation 2

Please explain why you would be willing to pay money in situation 2 where the total impact would be zero.

Additional cases (shortened)

Non-recyclable waste

Many households generate a significant amount of waste, a large proportion of which is not recyclable. Non-recyclable waste often results in more environmental pollution [...]

Please consider two different situations.

Situation 1: Your action has positive consequences

In this situation, if you reduce your non-recyclable waste by 100 pounds, the **total amount of non-recyclable waste in your community and all across the world also decreases by 100 pounds**.

Your action [...] / **Consequence** [...]

In situation 1, how much money would you be willing to pay to reduce your non-recyclable waste by 100 pounds?

Situation 2: Your action does not have any consequence

Now, please assume that, if you reduce your personal non-recyclable waste by 100 pounds, the **total non-recyclable waste in your community and all across the world does not change**. (For example, this could happen because your reduced waste makes it cheaper for others to dispose waste, and consequently others increase their non-recyclable waste by 100 pounds, exactly offsetting your reduction.)

Your action [...] / **Consequence** [...]

In situation 2, how much money would you be willing to pay to reduce your non-recyclable waste by 100 pounds?

Fairtrade wages

Workers in the textile industry in developing countries often earn wages that are below a decent living wage. Paying higher "fairtrade" wages (as defined by the Fairtrade Foundation) [...]

Please consider two different situations.

Situation 1: Your action has positive consequences

In this situation, for every additional dollar that you spend on fairtrade wages, **workers' total wages increase by that same dollar**.

Your action [...] / **Consequence** [...]

In situation 1, how much additional (beyond the standard price) would you be willing to spend on ten fairtrade garments?

Situation 2: Your action does not have any consequence

Now, please assume that if you choose to pay a premium to support fairtrade wages, **workers' total wages do not change**. (For example, this could happen because when you pay more for fairtrade products, you increase the demand for these products, making them more expensive for others, and they buy less of it.)

[...]

Animal welfare

On average, an American eats more than 20 chickens annually. Many of these chickens are raised in poor, cramped conditions [...]

Please consider two different situations.

Situation 1: Your action has positive consequences

In this situation, if you choose to pay a premium to support enhanced animal welfare standards for 20 chickens, the **overall number of chickens raised under better welfare standards increases by 20**.

Your action [...] / **Consequence** [...]

In situation 1, how much additional (beyond the standard price) would you be willing to spend on 20 animal-welfare-certified chickens?

Situation 2: Your action does not have any consequence

Now, please assume that if you choose to pay a premium to support enhanced animal welfare standards for 20 chickens, the **overall number of chickens raised under better welfare standards does not change**. (For example, this could happen because when you pay more for certified animal welfare products, you increase the demand for certified chicken, making it more expensive for others, and they buy less of it.)

Your action [...] / **Consequence** [...]

In situation 2, how much additional (beyond the standard price) would you be willing to spend on 20 animal-welfare-certified chickens?

Recent Issues

No. 410	Carmelo Latino	Surfing the Green Wave: What's in a "Green" Name Change?
No. 409	Kevin Bauer, Yan Chen, Florian Hett, Michael Kosfeld	Group Identity and Belief Formation: A Decomposition of Political Polarization
No. 408	Gianluca Anese, Marco Corazza, Michele Costola, Loriana Pelizzon	Impact of Public News Sentiment on Stock Market Index Return and Volatility
No. 407	Michele Costola, Matteo Iacopini, Casper Wichers	Bayesian SAR Model with Stochastic Volatility and Multiple Time-Varying Weights
No. 406	Peter Andre, Philipp Schirmer, Johannes Wohlfart	Mental Models of the Stock Market
No. 405	Peter Andre	Shallow Meritocracy
No. 404	Christian Alemán-Pericón, Christopher Busch, Alexander Ludwig, Raül Santaeulàlia-Llopis	Stage-Based Identification of Policy Effects
No. 403	Monica Billio, Roberto Casarin, Michele Costola	Learning from Experts: Energy Efficiency in Residential Buildings
No. 402	Julian Detemple, Michael Kosfeld	Fairness and Inequality in Institution Formation
No. 401	Kevin Bauer, Oliver Hinz, Michael Kosfeld, Lena Liebich	Decoding GPT's Hidden 'Rationality' of Cooperation
No. 400	Andreas Hackethal, Philip Schnorpfeil, Michael Weber	Households' Response to the Wealth Effects of Inflation
No. 399	Raimond Maurer, Sehrish Usman	Dynamics of Life Course Family Transitions in Germany: Exploring Patterns, Process and Relationships
No. 398	Pantelis Karapanagiotis, Marius Liebald	Entity Matching with Similarity Encoding: A Supervised Learning Recommendation Framework for Linking (Big) Data
No. 397	Matteo Bagnara, Milad Goodarzi	Clustering-Based Sector Investing
No. 396	Nils Grevenbrock, Alexander Ludwig, Nawid Siassi	Homeownership Rates, Housing Policies, and Co-Residence Decisions