

Harry Potter and the Welfare of the Willfully Blinded

Felix Bierbrauer

Abstract

This chapter presents a selective review of welfare economics. It argues that welfare analysts need to turn a blind eye to various aspects of individual preferences, otherwise applications of welfare economics yield repugnant conclusions. This situation is illustrated with characters from Hogwarts and then related to the theory of optimal taxation. Individual decisions to ignore relevant information, and the welfare implications that result, are then examined, as is the suppression of information that may affect the behavior of others. Such acts may conflict with liberal values. In the presence of behavioral biases, however, they may still positively affect welfare, in line with Lipsey and Lancaster's (1956) theory of the second best.

This reflection on the limits of welfare economics is not specific to the theme of deliberate ignorance. However, these limitations need to be at the center of any debate concerned with applications. Looking at the welfare implications of deliberate ignorance is not a straightforward application of the concept of externalities. It necessitates a reflection on the welfare implications of behavior that is potentially self-damaging. Moreover, it may conflict with liberal values and lead to repugnant conclusions unless there is a systematic reflection of what preferences to feed into welfare analysis.

Introduction

Let us consider a decision that affects the well-being of several individuals: Decision-relevant information is available. However, before the decision problem of interest can be addressed, another decision must be taken; namely, whether to use this information or to take the decision under ignorance.¹

¹ To be clear, in our discussion of this problem, we are not interested in a trade-off of the following sort: The information, if available, would improve the collective decision. Acquiring it, however, is costly. A cost-benefit analysis, therefore, must strike a balance between improved decision quality and the costs of information acquisition. Thus, we assume throughout that information is available at no cost.

To fix ideas, consider the following problem: At Hogwarts, a cake of given size has to be divided between Harry and Draco. Albus, a benevolent planner, chooses the division. He contemplates an application of utilitarian principles. If both Harry and Draco were selfish, representing their preferences by the same concave utility function and maximizing a sum of utilities would give rise to an equal split of the cake. At first glance, this seems to be an appropriate outcome. Harry, however, has altruistic feelings and derives utility from every piece received by Draco. Taking these feelings into account implies that Albus should assign a larger share to Draco. Deviating from a fifty-fifty split in this way has an opportunity cost, Harry's forgone utility, as he is eating less, and a welfare gain, Draco's extra utility from eating more, plus Harry's extra utility from Draco's extra utility. The latter implies that the welfare gain dominates. The conclusion that Draco should receive a larger share is, moreover, reinforced by Draco's sensations of envy which imply that every piece assigned to Harry reduces Draco's utility by more than just his forgone consumption utility. Thus, a consequence of utilitarianism seems to be that Harry is punished for his altruism and Draco is rewarded for his envy. Albus thinks twice. What information about preferences and utilities should be considered? What information should be ignored?

This example illustrates the possibility that taking account of information on preferences makes it possible to achieve higher welfare levels, even when the consequences seem repugnant. More generally, the question is: What types of information should welfare analysis be responsive to and what types of information should welfare analysis disregard? We begin by examining this problem at a broad conceptual level, and then discuss more specifically the welfare implications of information acquisition and information avoidance by individuals. To what extent are individual choices in this regard aligned with social welfare? To what extent would a welfare maximizer want to interfere with individual choice?

Blinding the Welfare Analyst

What Should Be the Domain of Welfare Analysis?

Let us start with an example from Coase (1960), in which an application of welfare economics seems uncontroversial. A fishery and a chemical plant reside along the same river. The chemical plant, which pollutes the water of the river, is upstream from the fishery. This negatively impacts the fish population and reduces the return to the fishery. This classic example is discussed in textbook treatments of the market failures that arise in the presence of externalities. Under *laissez-faire*, the chemical plant does not consider that its activities have negative consequences elsewhere. Emissions are then too large from a welfare perspective: Less chemical production or an investment in cleaner technology

combined with monetary compensation for missed profits would make both the chemical plant and the fishery better off.

An alternative example by Sen (1970) involves a person's decision whether or not to read *Lady Chatterley's Lover*, a novel with explicit accounts of sexual actions, and another person who is a prude and feels that no one should read such a book. If the first person reads the book, this has adverse consequences for the second. Does such a negative externality warrant the same treatment as the example with the chemical plant and the fishery? The logic of the latter suggests that censorship by the second person, in combination with compensation to the first person for being censored, would make both better off.

Sen uses this example to illustrate a conflict between the principles of welfare economics and liberal values that arises as soon as individuals have preferences over the private choices of others (e.g., what books to read, how to dress, whom to meet, and what opinions to express). Liberal principles require that such choices are respected. A stubborn application of welfare economics, by contrast, suggests that such choices should be corrected or moderated in return for compensation.

Goodin (1995) argues that the preferences used in welfare economics should go through a process of "laundering." Goodin is concerned with perverse or sadistic preferences, such as Draco's sensations of envy in the Hogwarts example. Welfare economics would be misguided if it took such preferences seriously. Draco would then be rewarded for his envy by receiving a bigger chunk of the cake than Harry. In Goodin's view, the censoring of preferences fed into welfare analysis does not require a paternalistic justification. It can often be justified by distinguishing what people really want, their true preferences, and the preferences that seem to be revealed by their choices: their revealed preferences. Revealed preferences may be shaped by temptations, short-term desires, or other sensations. Individuals might not want value judgments to reflect these sensations. In this case, there is a discrepancy between the normatively relevant true preferences and the revealed preferences.²

Note that the laundering of preferences can also be applied to address Sen's conflict of liberal values and welfare economics. In this case, laundering would have to remove preferences over the private choices of others. A liberal prude would admit that he feels annoyed by a fellow's reading of *Lady Chatterley's Lover* while not wanting this sensation to be used to justify the interference with private choices. An illiberal prude might disagree. There is, however, no conflict between liberal values and welfare economics provided that, for the purposes of welfare analysis, all liberals agree that their preferences should be laundered from attitudes toward the private choices of others.³

² For the conceptual distinction between revealed and true preferences, see Kahneman et al. (1997); for empirical applications, see Gruber and Köszegi (2001) and Allcott et al. (2019).

³ For the purposes of this chapter a formal definition of what constitutes a private choice is not needed. This will be intuitively clear from the context.

The preceding discussion points to the question: What types of preferences should one consider in welfare analysis? This is a normative question that cannot be answered by an application of welfare analysis itself. Laundering preferences from sensations of altruism or envy may be a way to avoid repugnant conclusions, such as punishing Harry for his altruism and rewarding Draco for his envy. Dismissing preferences over the private choices of others is a way to avoid illiberal conclusions, such as the censoring of books.

From the perspective of applied welfare economics, the deliberate ignorance of perverse or illiberal preferences may come with a cost. The welfare measure used in applied analysis would possibly take a higher value if dirty preferences were considered. The following examples from the welfare analysis of tax policy illustrate this point.

Welfare Economics of Taxation

In applied welfare analysis, there is a set of outcomes, and individuals have preferences over these outcomes. The outcomes and the preferences comprise the primitives of the problem. The problem then is to find the “right” outcome. Often this is taken to be the outcome that maximizes a utilitarian welfare measure. A more cautious approach—one that avoids interpersonal comparisons of utility—identifies a whole set of “right” outcomes, typically a set of Pareto optima.⁴ A field that makes heavy use of this framework is the analysis of tax policy in public finance. To give a feel for the relevance of the preceding discussion in applied work, let us explore various examples from this line of work.

The theory of optimal taxation applies welfare economics to the study of tax policy. The basic ingredients involve a government that uses tax policy to generate revenues and consumers who choose how much to consume, how much labor to supply, or how much to save. The choices of consumers are affected by the tax policy. A labor income tax affects the incentives to supply labor; a tax on capital income affects the return on savings. An optimal policy maximizes a utilitarian welfare objective by taking these behavioral responses of consumers into account.⁵ A well-known result is that taxes should follow an inverse elasticity rule: taxes should be high when behavioral responses, usually measured by the price elasticity of supply and demand, are low and vice versa.

It is an intuitive finding. If capital income is shifted abroad in response to taxation but labor income is not, then the tax on capital income should not be as high as the tax on labor income. If demand for necessities such as bread or gas is less price sensitive than the demand for luxuries, then the tax on bread

⁴ The defining property of a Pareto optimal outcome is that moving away from it necessarily makes some people worse off.

⁵ Modern analysis of this problem dates back to Ramsey (1927). A rich body of literature has refined this approach in various ways, with seminal contributions by Mirrlees (1971), Atkinson and Stiglitz (1976), Diamond (1998), and Saez (2001).

should be higher than the tax on champagne. These examples raise distributive questions, yet even if these are taken into account, the logic of the inverse elasticity rule remains intact, all other things being equal: When two goods are consumed equally by the rich and the poor, the one with the lower elasticity of demand should be taxed at a higher rate (see, e.g., Diamond 1975).

In any case, such taxes interfere with the private choices of individuals. If demand for books such as *Lady Chatterley's Lover* or *The Satanic Verses* is less price sensitive than the demand for more "respectable" types of literature such as *Hamlet* or *The Koran*, the logic of optimal tax theory would suggest having higher taxes on the former and lower taxes on the latter. Hence, another round of laundering may be needed to avoid repugnant or illiberal conclusions from the application of optimal tax theory. The consequence of such laundering, however, is that the resulting tax system is not optimal from the perspective of a welfare measure that is based on the preferences that individuals reveal through their market behavior. Just as an insurance company loses profit when nondiscrimination requirements exclude different premia for men and women, a welfare maximizer has to live with the fact that laundering precludes reaching welfare levels that would otherwise be attainable.

From the perspective of practical tax policy, having different tax rates for different types of books is a contrived example. There are, however, more plausible implications of optimal tax theory that raise similar issues. There is a rich literature on the optimal taxation of couples (for seminal references, see Boskin and Sheshinski 1983; Kleven et al. 2009). These studies show that a differential taxation of primary and secondary earners in a couple is desirable from a welfare perspective. This finding combines the logic of the inverse elasticity rule with the empirical observation that the labor supply of females, who are more often in the role of the secondary earner, is more tax sensitive than the labor supply of males. Thus, an optimal tax system should apply different tax rates to the primary and secondary earners in a couple. In particular, income due to the secondary earner should be taxed at a lower rate.

Obviously, such differential taxation interferes with a private choice. It affects the assignment of roles in a couple, in particular the decision of who should contribute how much to the family's income. With a progressive income tax system, income splitting⁶ is the only way to have a couple's tax burden solely depend on their overall income, irrespective of who contributed how much. Hence, an attempt not to interfere with private choices implies that the inverse elasticity logic is not applied, with the consequence that potential welfare gains remain on the table and that the female labor supply is discouraged more than it would otherwise be.

The treatment of altruism plays a prominent role in the theory of capital and inheritance taxation. Results on the desirability of capital taxes crucially

⁶ Let y^p be the income of the primary earner and y^s the income of the secondary earner: under income splitting, the tax burden (T) of a couple, as a function of income, is $2T(y^p + y^s)/2$.

depend on assumptions about the altruism of parents toward their children. In the analysis of Farhi and Werning (2010), altruism implies that a bequest is a source of utility both for parents and children. A bequest subsidy is warranted to make sure that the positive externalities from leaving a bequest are taken into account. In Piketty and Saez (2013), by contrast, the degree of altruism varies from generation to generation in an unpredictable way. This is shown to imply that a redistributive taxation of bequests is desirable. The redistribution from lucky children with high bequests to unlucky ones is shown to be part of a welfare-maximizing policy. Needless to say, such an analysis only makes sense on the assumption that altruism is a legitimate ingredient of welfare analysis. The conclusions on the desirability of bequest taxes and subsidies would not survive a laundering of preferences from altruism.

These examples demonstrate the difficulty of addressing which preferences to feed into welfare analysis. A naive use of revealed preferences may give rise to repugnant or illiberal conclusions. Forcing the welfare maximizer to turn a blind eye (thereby exercising deliberate ignorance) to the dirty or private aspects of individual preferences may be an appropriate remedy. As seen in our discussion of tax policy, the extent to which this is done can drastically affect the policy implications of “applied work.”

Internalities

The Coasean bargaining example involving the chemical plant and fishery is one of externalities. One firm pursues its economic interests at the expense of another one. Welfare economics stipulates that such externalities must be considered and, moreover, that doing so in an appropriate way will make both firms better off.

This logic has been extended to address internal conflicts that individuals may have. Self-control problems are a prominent example. An individual may have a long-term goal of leading a healthy life. In the short term, however, the individual is confronted with temptations such as alcohol, cigarettes, or unhealthy food. Giving in to such temptations damages the individual’s long-run goals. The literature often refers to such self-damaging behavior as creating internalities. Applied work in optimal tax policy has discussed corrective taxes that address such internalities. For instance, O’Donoghue and Rabin (2006) characterize “optimal sin taxes” that mitigate self-damaging behavior.⁷

Public policy that addresses internalities interferes with the private choices of individuals. This raises the question whether it provokes the type of conflict between liberal values and welfare economics illustrated by Sen’s Lady Chatterley example. In that example, a problem arises as one person has preferences over the private choices of another person. Here, the public policy

⁷ Taxes are not the only instrument that can be used to address internalities. A prominent alternative are nudges (Thaler and Sunstein 2008; Mariotti et al. 2018).

maker has preferences over the lifestyle of individuals (e.g., their drinking and smoking habits). Isn't this the same kind of problem?

This question has spurred controversies (see Gul and Pesendorfer 2008; Loewenstein and Haisley 2008). For the proponents of such policies, the answer is clearly “no,” provided that the policy maker does not pursue its own agenda but has preferences that are aligned with the individuals' long-term goals. The agenda on “soft” or “liberal” paternalism (for the best-known example, see Thaler and Sunstein 2008) focuses on situations where individuals can be enabled to behave in accordance with their long-run goals, without harming others who do not suffer from the same kind of self-control problem.

In line with this program, O'Donoghue and Rabin (2006) look at a population of smokers who differ in the intensity of their self-control problems. Some are heavy smokers and have pronounced self-control problems; others are occasional smokers and have self-control problems that are not as severe. O'Donoghue and Rabin show that a sin tax may, nevertheless, benefit all smokers. Heavy smokers like the tax as it helps with the self-control problem, and the revenue that is generated can be used to compensate light smokers, who would otherwise be harmed by the sin tax.

The analysis also points to the limits of liberal paternalism. A Pareto-improving sin tax is possible only if there is a one-to-one relation between the number of cigarettes smoked and the intensity of the self-control problem. If one introduces heavy smokers with no self-control problems to the system, the possibility of a Pareto-improving sin tax is gone. In this case, one has to make rational smokers worse off when one attempts to help the smokers with self-control problems. From the perspective of the rational smoker, this is akin to an illiberal interference with a private choice.

Blinding Oneself

Let us now turn from the issue of what information a welfare maximizer *should* ignore, to the information that individuals *do* ignore. A strand of behavioral research has investigated circumstances under which individuals take decisions while deliberately ignoring decision-relevant information (for a survey, see Golman et al. 2017; Hertwig and Engel, this volume, 2016). Here, our focus is on the welfare implications of such information avoidance. We will go through some prominent examples and discuss the criteria developed above:

- When a person engages in information avoidance, does this give rise to externalities (positive or negative effects on others) or internalities (positive or negative effects on that person)?
- If so, are mechanisms in place to ensure that these are considered?
- If not, is information avoidance a private choice that should not be the subject of welfare analysis?

- If not, are “dirty” preferences at play that should be removed from welfare analysis?

Let us consider climate change denial. Climate change is one of the most pressing problems currently facing humanity. It is also a prime example of an externalities problem. The CO₂ emissions from past and current generations have drastic consequences on younger and future generations. Classic welfare economics stipulates that such problems be addressed using corrective taxation or quantity controls. Yet, taking these measures, which would be effective in mitigating the course of climate change, is politically controversial. Some opponents of such policies even deny the existence of a problem that needs to be solved. In other words, they exercise deliberate ignorance of the overwhelming evidence on climate change.

The Coasean example of the fishery and chemical plant rests on the premise that all involved accept the description of the problem. The calculus of welfare economics is applicable only if both parties agree that the emissions of the chemical plant are harmful to the fishery. If, for instance, the fishery denies this, there is no point in having a cost-benefit analysis determine the optimal level of emissions reduction, accompanied by compensation payments to ensure that both parties are better off. Thus, a deliberate ignorance of the harm caused by emissions is equivalent to denying that policies to address this problem can be justified with an appeal to welfare.⁸

This raises the question whether welfare economics can be applied in the face of such deliberate ignorance. From the perspective of those who deny climate change, public policies which seek to address it are unjustified and paternalistic. Should their welfare still be considered when such policies are formulated? If so, which preferences should enter the cost-benefit calculation: the preferences that are articulated in the political process or a laundered version that no longer contains traces of deliberate ignorance?

Ignorance of Performance Evaluations

Climate change denial has a political motivation. As another example, consider a teacher who does not want to look at teaching evaluations, for fear of a negative outcome. Here, the motivation is more personal: the desire to retain a positive self-image. Still, there are externalities: a mediocre teacher’s reluctance to explore ways of improving his teaching is harmful to his students.

Golman et al. (2017) argue that the hedonic consequences of information avoidance need to be considered. This concrete example raises the question whether the teacher’s hedonic utility from keeping a positive self-image should

⁸ According to the impossibility result by Myerson and Satterthwaite (1983), efficient Coasean bargaining is not possible if the intensity of the externalities problem cannot be objectively verified. If only the fishery knows how much harm is caused by emissions and only the chemical plant knows how costly it is to avoid them, efficiency is out of reach.

be weighed against the benefits afforded to students from improved teaching. The alternative perspective is that a desire to keep an unjustified positive self-image is a “dirty” preference that should be removed from a cost-benefit analysis of additional training.

Reluctance to Test for Diseases

Hertwig and Engel (this volume, 2016) report the case of James Watson, who had his genome sequenced but chose to remain ignorant about his predisposition for Alzheimer disease. This is an example of private choice, a choice that affects the welfare of one person and possibly his close relatives, or at least should be treated as such. Remember the lesson from Sen’s Lady Chatterley example: Preferences over the private choices of others have to be removed from welfare analysis, otherwise, liberal principles conflict with welfare analysis. If James Watson’s decision is not treated as a private choice, what else should be? If the disease were infectious and if the risk of infecting others could be reduced (e.g., by a vaccination), the conclusion would, of course, be different: Externalities would enter the picture. A welfare analysis that weighs the personal costs of acquiring unpleasant information against the health risks of others might appear quite reasonable.

Blinding Others

The previous examples involve individuals who blind themselves and, by doing so, potentially affect others negatively (e.g., opponents of climate change, teachers adverse to external evaluations). We now turn to the deliberate blinding of others.

Motivated Beliefs

Bénabou and Tirole (2006a) present a model of motivated beliefs in which individuals suppress unfavorable information to handle cognitive dissonances. Specifically, individuals have a desire to believe that the world is just, that those who work hard or invest in their human capital can reap the rewards and become financially better off. At the same time, individuals are confronted with the evidence that social mobility is imperfect, that economic inequality tends to persist over generations, and that hard work does not necessarily pay off. There is evidence that individuals bias their perceptions of social mobility against this evidence and instead remain overly optimistic. They stick to the American dream despite the facts that point to the contrary (Alesina et al. 2018).

In the model of Bénabou and Tirole (2006a), the suppression of this unfavorable evidence has a benefit: It keeps individuals going, such that they invest more in human capital than they otherwise would. This positive effect is due to

the assumption that individuals also suffer from a present bias. Educational effort, therefore, tends to be inefficiently low. Individuals give too much weight to the immediate costs of acquiring human capital and too little weight to the higher future income that results from the investment. A suppression of unfavorable information on the returns to education can thus mitigate an individual's tendency to procrastinate.

In their preferred interpretation of the model, Bénabou and Tirole take an intergenerational perspective. Parents tell their children about the returns to effort. The children, in turn, choose how much effort to exert when going to school. Thus, parents shield their children from unfavorable information on the returns to effort in an attempt to overcome their laziness. What are the welfare implications of these choices? Are parents doing harm to their children? The answer would be "yes" if there was no present bias; that is, if the children did not place too much weight on immediate gratification and too little on the long-term returns of educational effort. In this case, children who become victims of their parents' propaganda would invest more than is in their own interest. With the present bias, however, the parents' indoctrination may be regarded as second-best alternative, so that the children are better off with it.

This example illustrates a more general lesson from what is known as the *theory of the second best* (Lipsey and Lancaster 1956). With distortions already in place, adding another distortion may have beneficial effects for welfare. A welfare analysis of deliberate ignorance might therefore be misguided if it focuses solely on one type of deliberate ignorance in isolation. Discovering that deliberate ignorance serves a useful purpose may require evaluating it against the background of the entire menu of individual biases relevant for the application at hand.

Manipulating the Salience of Taxes

Positive welfare effects of blinding others have also been documented in the context of tax policy. Chetty et al. (2009) report on a field experiment that involves a manipulation of price tags in U.S. supermarkets. The standard is a price tag that does not include sales taxes. The manipulated price tags highlighted, however, the tax inclusive price. Chetty et al. found that the manipulation triggers a behavioral response: fewer items are sold. Their work documents, however, that consumers are well-informed about sales taxes. Thus, the manipulated price tag did not provide new information; it only made available information more visible. This visibility had consequences: consumers bought more if the information on taxes was suppressed.

The conventional perspective in public finance is that any sales tax has an efficiency cost. Such a tax drives a wedge between the prices paid by consumers and the prices received by producers. As a consequence, gains from trade are not exhausted. A consumer who is willing to pay ten but faces a tax inclusive price of eleven will thus not purchase the product. If producers are willing to sell for nine,

there are gains from trade between the producer and the consumer. Those gains would be realized if there was no tax, but not with the tax. The forgone benefits of such transactions constitute the efficiency costs of taxes.

How is this logic affected by the behavioral responses to the salience of taxes? Chetty et al. (2009) assume that the demand with tax-inclusive prices reflects true preferences. Thus, individuals overconsume when taxes are not salient. This overconsumption, in turn, helps to mitigate the efficiency costs of taxation. This is another instance of a second-best logic, one that combines a behavioral bias with an inefficiency that also prevails with rational agents, the distortionary effects of taxation.⁹

Concluding Remarks

Applications of welfare economics require principled decisions on which type of preferences to use in the analysis. Taking account of sensations such as envy or altruism may give rise to repugnant conclusions. Incorporating preferences over the private choices of others may clash with liberal values. Thus, to be relevant, welfare analysis needs to turn a blind eye to certain aspects of individual preferences.

Welfare analysis must also account for decisions by individuals to ignore information that is readily available, or to suppress information that would otherwise be available to others. Analysis of such choices faces the difficulty of delineating the proper domain of welfare economics: Should welfare analysis take the preferences of those who deny climate change into account or ignore them? Should genetic tests for health risks be treated as a private affair that is not subjected to welfare analysis?

An interesting line of recent research looks at related questions from an empirical perspective, by trying to elicit the preferences that individuals want to be factored into welfare analysis. For instance, Weinzierl (2017) reports that individuals demand a laundering of preferences from sensations of envy. In this study, he asked individuals to assume the perspective of a policy maker and confronted them with two situations. In the second situation, incomes are, for everybody, higher than in the first. Inequality, though, is also higher and consequently the overall utilitarian welfare is lower. Even then, a majority of the respondents chose the second situation over the first.¹⁰ Pursuing this avenue further might prove useful for future research on the welfare implications of deliberate ignorance.

⁹ The analysis is, however, sensitive to the assumption that individuals overconsume when taxes are not salient. Consider the alternative assumption that true preferences correspond to the demand that is observed when consumers see the price tags they are used to. In this case, making taxes more salient will aggravate the tax distortions.

¹⁰ Relatedly, Charité et al. (2015) analyzed whether individuals support the use of welfare measures that respect the reference-point dependence of preferences, and Weinzierl (2014) investigated the support for alternatives to utilitarian welfare maximization.

Acknowledgments

Background paper prepared for the Ernst Strüngmann Forum on deliberate ignorance. I benefited from conversations with Carina Bierbrauer, Martin Hellwig, and François Maniquet. I am also grateful for comments by the participants of the Ernst Strüngmann Forum on deliberate ignorance. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2126/1–390838866.