

7/29/24

All rights reserved

Forthcoming, *Theoretical Inquiries in Law*

# Brave New World? Human Welfare and Paternalistic AI

*Cass R. Sunstein\**

*“Choice Engines,” powered by Artificial Intelligence (AI) and authorized or required by law, might produce significant increases in human welfare. A key reason is that they can simultaneously (1) preserve autonomy and (2) help consumers to overcome inadequate information and behavioral biases, which can produce internalities, understood as costs that people impose on their future selves. Importantly, AI-powered Choice Engines might also take account of externalities, and they might nudge or require consumers to do so as well. Different consumers care about different things, of course, which is a reason to insist on a high degree of freedom of choice, even in the presence of internalities and externalities. Nonetheless, AI-powered Choice Engines might show behavioral biases, perhaps the same ones that human beings are known to show, perhaps others that have not been named yet, or perhaps new ones, not shown by human beings, that cannot be anticipated. It is also important to emphasize that AI-powered Choice Engines might be enlisted by insufficiently informed or self-interested actors, who might exploit inadequate information or behavioral biases, and thus reduce consumer welfare. AI-powered Choice Engines might also be deceptive or manipulative, and legal safeguards are necessary to reduce the relevant risks.*

## I. CHOICE AND WELFARE

When do people’s choices make their lives go better? When do legal interventions, designed to improve people’s lives, actually help? How should regulators, and others involved in law, proceed if populations are diverse, and if some people will be affected differently from others? How might artificial intelligence (AI) help? My purpose here is to answer these questions, above all by showing how AI can help people to overcome an absence of information and behavioral biases. In achieving that goal, AI may or may not be paternalistic. In some cases, it might turn out to engage in libertarian paternalism, preserving freedom of choice while steering people in

---

\* Robert Walmsley University Professor, Harvard University. Some parts of this essay draw on previous work on choice engines and fuel economy regulation. I am grateful to the Harvard Law School’s Program on Behavioral Economics and Public Policy, and the Harvard Law School’s Initiative on Artificial Intelligence and Law, for valuable support.

directions that make their lives go better.<sup>1</sup> At the same time, AI poses serious risks, including risks of deception and manipulation; legal safeguards will be required.

To approach these questions, begin with three sets of findings:

1. Calorie labels, required by law, are designed to help people make better choices about what to eat, and there is evidence that calorie labels have welfare benefits by doing exactly that.<sup>2</sup> At the same time, calorie labels seem to have a greater effect on people who do not have self-control problems than on people who suffer from such problems.<sup>3</sup> That is a serious problem, because a central point of calorie labels is to help people with self-control problems. It is possible that in some populations, calorie labels affect people who do not need help, and have little or no effect on people who do need help, except to make them feel sad and ashamed. If so, calorie labels could make people worse off.
2. On average, people appear to benefit from home energy reports, which inform them about their energy usage. As a result of such reports, people save money, on average. They are willing to pay, on average, a positive amount to receive such reports. But some people are willing to pay far more than others for home energy reports.<sup>4</sup> In fact, some people are willing to pay *not* to receive home energy reports. They dislike receiving such reports, and they believe that they would be better off without them. While home energy reports are designed to save consumers money and to reduce externalities, sending such reports to some people, who do not want to receive them, seems to make them worse off. More targeted use of home energy reports, sending them only to those who want and would benefit from them, could produce significant welfare gains.<sup>5</sup>

---

<sup>1</sup> See generally RICHARD THALER & CASS R. SUNSTEIN, *NUDGE: THE FINAL EDITION* (2021). A brisk, preliminary account, much developed and expanded upon here, can be found in Cass R. Sunstein, Choice Engines and Paternalistic AI, 11 *Humanities and Social Sciences Communications*, article number 888 (2024); at times I draw on that account, which was meant as a forerunner of this far more elaborate one.

<sup>2</sup> Linda Thunström, *Welfare Effects of Nudges: The Emotional Tax of Calorie Menu Labeling*, 14 *Judgment & Decision Making* 11, 18-19 (finding that a substantial number of study participants favored calorie labels because “calorie content would matter to my meal choice,” *id.* at 19) (2019).

<sup>3</sup> *Id.* at 20.

<sup>4</sup> Hunt Allcott & Judd Kessler, *The Welfare Effects of Nudges*, 11 *AM. ECON. J.: APPLIED ECON.* 236, 257 (2019).

<sup>5</sup> See *id.* at 271 (proposing “an opt-out program that targets consumers who would generate large welfare gains and excludes consumers who would not”).

3. Graphic warning labels on sugary drinks are designed to affect consumer behavior, and there is evidence that they can do so.<sup>6</sup> But such labels have greater effects on some consumers than on others.<sup>7</sup> Disturbingly, such labels can lead people who do not have self-control problems to consume less in the way of sugary drinks, while having a significantly smaller effect on people who do have self-control problems. In addition, many people do not like seeing graphic warning labels. The average person in a large sample reported being willing to pay about \$1 to *avoid* seeing the graphic warning labels.<sup>8</sup> It is likely that such labels are helping some and hurting others. It is possible that such labels are on balance causing harm, in the sense that they are reducing people’s welfare in aggregate.

From these sets of findings, we can draw three simple conclusions. *First*, legal interventions, designed to improve people’s lives, might well have different effects on different populations. Under favorable conditions, they might have large positive effects on people who need help, and small or no effects on people who do not need help. Under unfavorable conditions, they might have small or no effects on a group that needs help, and large effects on a group that does not need help. Large effects on a group that does not need help may not improve that group’s welfare. For example, people who have no need or reason to change their spending patterns or their diets might end up doing so.

*Second*, legal interventions may have either positive or negative *hedonic* effects. People might like seeing labels, or they might dislike seeing labels. They might like receiving information, or they might not like receiving information. With respect to the receipt of information, there is a great deal of heterogeneity.

*Third*, and consistent with the first conclusion, an understanding of the average treatment effect of an intervention does not tell us about the overall effect on social welfare, which is what regulators and others involved in law and policy need to know, at least if they are welfarists.<sup>9</sup> The third conclusion is easy to miss. An intervention might have a positive average treatment effect, leading to healthier choices, but if people who make healthier choices are already healthy, and if

---

<sup>6</sup> Hunt Allcott et al., *When Do “Nudges” Increase Welfare?* 4 (Nat’l Bureau of Econ. Rsch., Working Paper No. 30740, 2022), [https://www.nber.org/system/files/working\\_papers/w30740/w30740.pdf](https://www.nber.org/system/files/working_papers/w30740/w30740.pdf).

<sup>7</sup> *Id.*

<sup>8</sup> *Id.* at 25.

<sup>9</sup> *See id.* at 29 (“While much of the empirical literature has focused on whether nudges have average effects in the “right” direction, we show that welfare also depends on how nudges affect the variance of distortions.”).

people who do not make healthier choices are sad and ashamed, then the interventions might turn out to reduce social welfare. If a legal intervention makes people feel sad and ashamed, it is, to that extent reducing people's welfare. It is not a lot of fun to feel sad or ashamed.

These points can be made about a wide range of legal interventions. They hold for taxes that are designed to produce healthier choices: it is possible that such taxes will have little or no effect on the people they are particularly intended to help, while having a significant adverse effect on people they are not (particularly) intended to help. (Consider soda taxes.) They hold for mandates, understood as requirements backed by the force of law, and also for bans, understood as prohibitions similarly backed. A ban on some activity or product might, on balance, hurt people who greatly benefit from that activity or product, while helping people who lose only modestly from it. (Consider bans on the purchase of incandescent lightbulbs.) In all of these cases, more targeted action, treating different individuals and groups differently, would be better than "mass" action, treating everyone the same. They hold for automatic enrollment: it is possible that automatic enrollment in some plan will have no effects on people who benefit from enrollment (because they would enroll in any case) while harming people who do not benefit from enrollment (because some or many who lose do not opt-out, perhaps because of inertia).

Here, then, is where we are. Interventions designed to influence people's choices may or may not increase social welfare. A recurring problem is that of heterogeneity. People have different needs, preferences, and values, and an intervention that affects a large population might help some and hurt others. The average treatment effect is not the same as the welfare effect. The welfare effect is what matters. How can it be improved? AI, focused on improving social welfare, can provide an answer, at least if it is focused on (1) harms that people do to their future selves and (2) harms that people do to others.

## II. CHOICE ENGINES

Choice Engines<sup>10</sup> can be defined as instruments, typically (but not always) online, by which choosers are given an opportunity to provide some information about themselves and their

---

<sup>10</sup> The term is not in general use, but something like it can be found in various places, with variations. See Michael Yeomans et al., *Making Sense of Recommendations*, 32 J. BEHAV. DECISION MAKING 403, 403 (2019); Guy Champnis, *The Rise of the Choice Engine*, ENERVEE (Mar. 6, 2018, 12:00 AM), <https://www.enervue.com/blog/the-rise-of-the-choice-engine>; *Why Getting Help Matters*, EDELMAN FINANCIAL ENGINES <http://corp.financialengines.com/individuals/why-getting-help-matters.html> (last visited July 22, 2024). Compare choice engines to the following, which is regrettably complicated: *Buying*

preferences, and then receive a recommendation, or a set of recommendations, about what they ought to choose. Suppose that you are choosing a new laptop. You might say that you want to pay less than \$1200; that you want the laptop to weigh less than three pounds; and that you want the laptop to have at least eight hours of battery life. A Choice Engine might give you a set of options. It may or may not use Artificial Intelligence (AI). It might well be an algorithm: people with characteristics and preferences A, B, and C typically prefer, in order, laptops X, Y, and Z. Any kind of AI-powered Choice Engine might serve to specify and narrow your choices. Choice Engines of this kind might be used for any number of choices, including (to take some random examples) choices of dogs,<sup>11</sup> mystery novels, cellphones, shavers, romantic partners, shoes, tennis racquets, and ties. Choice Engines may or may not use AI, and if they do, they can use AI of different kinds.

A Large Language Model might work as a Choice Engine; it might offer options almost instantly. In 2024, I offered the following prompt to Chat GPT: “I live in a house in a suburb. I like large dogs. What breed of dog should I get?” Here is the answer I received:

*Living in a house in a suburb gives you the advantage of more space, which is great for larger dog breeds. Here are a few large dog breeds that could be a good fit:*

1. **Golden Retriever:** *Friendly, tolerant, and well-known for their calm demeanor. They require regular exercise but are generally easy to train.*
2. **Labrador Retriever:** *Energetic, outgoing, and friendly. Labs are great for families and enjoy being active.*
3. **German Shepherd:** *Intelligent, loyal, and protective. They are highly trainable and make excellent companions.*
4. **Bernese Mountain Dog:** *Gentle, affectionate, and good-natured. They are generally calm but do need regular exercise.*
5. **Great Dane:** *Despite their size, they are known for being gentle giants. They are friendly, affectionate, and great with families.*
6. **Mastiff:** *Loyal, protective, and good-natured. They are relatively low-energy for their size but require space due to their large stature.*

---

*a Refrigerator Guide: How to Choose a New Fridge in 2024*, WHIRLPOOL, <https://www.whirlpool.com/blog/kitchen/buying-guide-refrigerator.html> (last visited July 22, 2024).

<sup>11</sup> See *Welcome to the Purina Dog Breed Selector*, PURINA, <https://www.purina.co.uk/find-a-pet/dog-breeds/breed-selector> (last visited July 22, 2024).

7. *Newfoundland*: Known for their sweet temperament and excellent swimming ability. They are gentle, patient, and good with children.
8. *Rottweiler*: Confident, loyal, and protective. With proper training and socialization, they make great family pets.

*Make sure to consider the activity level, grooming needs, and temperament of each breed to find the best match for your lifestyle. Additionally, visiting local shelters and rescues can provide opportunities to find a large dog that needs a home.*

That is an excellent answer. I happen to know a lot about dogs, and the answer almost perfectly matches my tastes (though I like Labrador Retrievers best)

My main concern here is how Choice Engines, powered by AI and authorized or required by law, might address an absence of information or a behavioral bias, and thus increase social welfare. I will be drawing attention to the importance of targeting and personalization, and the promise of AI-powered Choice Engines in improving individual choices. This will be a largely optimistic account, focused on the extraordinary promise of Choice Engines, but I will also have something to say about threats, including the risk of deception and manipulation. If more space will be given to reasons for optimism than to threats, it is not because threats do not deserve attention; they are exceedingly important and should be kept in mind throughout.

For retirement plans, many U.S. employers use something like a Choice Engine. They know a few things about their employees (and possibly more than a few). On the basis of what they know, they automatically enroll their employees in a specific plan. The plan is frequently a diversified, passively managed index fund. Employees can opt out and choose a different plan, if they like. Alternatively, employers might offer employees a specified set of options, with the understanding that all of them are suitable, or suitable enough. (Options that are not suitable should not be included.) They might provide employees with simple information by which to choose among them. The options might be identified or rethought with the assistance of AI or some kind of algorithm.<sup>12</sup>

Here is one reasonable approach: Automatically enroll employees in a plan that is most likely to be improve their well-being, given everything relevant that is known about them.<sup>13</sup> What suits

---

<sup>12</sup> See, e.g., FIDELITY GO <https://www.fidelity.com/managed-accounts/fidelity-go/overview> (last visited July 22, 2024).

<sup>13</sup> This is consistent with IAN AYRES & QUINN CURTIS, RETIREMENT GUARDRAILS 159-61 (2023).

people in their thirties is not the same as what suits people in their sixties. To be sure, identification of that plan might prove daunting; it might be challenging to know what plan really is the right one for (say) a risk-averse fifty-year-old male, with two children, earning \$150,000 annually. But a large number of plans can at least be ruled out.<sup>14</sup> Note that if the focus is on improving employee well-being, we are not necessarily speaking of revealed preferences. People might choose plans without sufficiently appreciating the opportunities or the risks.

For retirement savings, we can easily imagine many different kinds of Choice Engines. Some of them might be mischievous, in the sense that they are poorly suited to workers' situations; some of them might be fiendish, in the sense that they help the plan's provider, not the workers; some of them might be random; some of them might be coarse or clueless; some of them might show behavioral or other biases of their own; some of them might be self-serving.<sup>15</sup> For example, people might be automatically enrolled in plans with high fees. They might be automatically enrolled in plans that are not diversified. They might be automatically enrolled in money market accounts, which offer low returns. They might be automatically enrolled in dominated plans.<sup>16</sup> They might be automatically enrolled in plans that are especially ill-suited to their situations. They might be given a large number of options and asked to choose among them, with little relevant information, or with information that leads them to poor choices.

The general point is that in principle, Choice Engines might work to overcome behavioral biases.<sup>17</sup> For retirement plans, Choice Engines may or may not be paternalistic. If they are not paternalistic, it might be because they simply provide a menu of options, given what is known about relevant choosers.<sup>18</sup> If they are paternalistic, they might be mildly paternalistic, moderately paternalistic, or highly paternalistic. A moderately paternalistic Choice Engine might impose nontrivial barriers to those who seek certain kinds of plans (such as those with high fees). The barriers might take the form of information provision, "are you sure you want to?" queries, and requirements of multiple clicks. We might think of a moderately paternalistic Choice Engine as

---

<sup>14</sup> See *id.* at 160.

<sup>15</sup> An episode of *Black Mirror* could easily be based on such scenarios. *Black Mirror* (Netflix).

<sup>16</sup> See AYRES & CURTIS, *supra* note 13.

<sup>17</sup> See Vidya Athota et al., *Can Artificial Intelligence (AI) Manage Behavioural Biases Among Financial Planners?*, 31 J. GLOB. INFO. MGMT. 1, 7-9 (2023). For a disturbing set of findings, see Yang Chen et al., *A Manager and an AI Walk into a Bar: Does ChatGPT Make Biased Decisions Like We Do?* (2023), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4380365](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4380365).

<sup>18</sup> See *Welcome to the Purina Dog Breed Selector*, *supra* note 11.

offering “light patterns,” understood as strategies that help consumers, as contrasted with “dark patterns,” understood as strategies that harm consumers (for example, by automatically enrolling them in costly programs that have little or no value).<sup>19</sup> A highly paternalistic Choice Engine might forbid employees from selecting any plan other than the one that it deems in the interest of choosers, or might make it exceedingly difficult for employees to do that.

Now broaden the viewscreen and turn to a question outside of the domain of retirement plans: Would you like to buy an energy-efficient refrigerator that would cost you \$X today, but save you ten times \$X over the next ten years? What kind of car would you like to buy? Would you like to buy a fuel-efficient car that would cost you \$800 more upfront than the alternative, but that would save you \$8000 over the next ten years? What characteristics of a car, or a refrigerator, matter most to you? Do you need a large car? Do you like hybrids? Are you excited about electric cars, or not so much?

A great deal of work finds that consumers suffer from “present bias.”<sup>20</sup> Current costs and benefits loom large; future costs and benefits do not. For many of us, the short-term is what most matters, and the long-term is a foreign country. The future is Laterland, a country that we are not sure that we will ever visit. This is so with respect to choices that involve money, health, safety, and more. There are plausible evolutionary explanations for present bias. If you are running from a tiger, you ought not to spend much time thinking about long-term monetary gains (or your retirement savings). But under modern circumstances, present bias can get you into a great deal of trouble.

AI need not suffer from present bias.<sup>21</sup> Imagine that you are able and willing to consult AI to ask it what kind of car you should buy. Imagine too that you discover that you are, or might be,

---

<sup>19</sup> See generally Jamie Luguri & Lior Strahilevitz, *Shining A Light on Dark Patterns*, 13 J. LEGAL ANALYSIS 43 (2021).

<sup>20</sup> See, e.g., Joachim Schleich et al., *A Large Scale Test of the Effects of Time Discounting*, 80 ENERGY ECON. 377 (2019); Madeline Werthschulte & Andreas Loschel, *On the Role of Present Bias and Biased Price Beliefs in Household Energy Consumption*, 109 J ENV. EC. AND MANAGEMENT (2021); Theresa Kuchler & Michaela Pagel, *Sticking to Your Plan: The Role of Present Bias for Credit Card Paydown*, 139 J. FIN. ECON. 359 (2018), [https://www.nber.org/system/files/working\\_papers/w24881/w24881.pdf](https://www.nber.org/system/files/working_papers/w24881/w24881.pdf); Ted O’Donoghue & Matthew Rabin, *Present Bias: Lessons Learned and to Be Learned*, 105 AM ECON REV 273 (2015); Jess Benhabib et al., *Present Bias, Quasi-Hyperbolic Discounting, and Fixed Costs*, 69 GAMES ECON. BEHAV. 205 (2010); Yang Wang & Frank Sloan, *Present Bias and Health*, J. RISK UNCERTAINTY 177 (2018). Importantly, Wang and Sloan find strong evidence of present bias in connection with health-related decisions.

<sup>21</sup> It might. Yang Chen et al., *supra* note 17 (manuscript at 10).



present-biased, in the sense that you prefer a car that is not (according to AI) the one that you should get. What then? We could easily imagine Choice Engines for motor vehicle purchases in which different consumers provide relevant information about their practices, their preferences, and their values, and in which the relevant choice engine immediately provides a set of options – say, Good, Better, and Best. Something like this could happen in minutes or even seconds, perhaps a second or two. If there are three options – Good, Better, and Best – verbal descriptions might explain the ranking. Or a Choice Engine might simply say: Best For You. It might do so while allowing you to see other options if you indicate that you wish to do so. It may or may not be paternalistic, or come with guardrails designed to protect consumers against serious mistakes.<sup>22</sup>

In 2024, I gave ChatGPT the following prompt: “I want a mid-size car with good fuel economy. What are my best options?” Here is the answer:

*Here are some of the best mid-size cars known for their good fuel economy:*

**1. Toyota Camry Hybrid:**

- *Fuel Economy: Up to 52 mpg combined (51 city / 53 highway)*
- *Features: Reliable, spacious, advanced safety features.*

**2. Honda Accord Hybrid:**

- *Fuel Economy: Up to 48 mpg combined (48 city / 47 highway)*
- *Features: Comfortable ride, spacious interior, strong hybrid powertrain.*

**3. Hyundai Sonata Hybrid:**

- *Fuel Economy: Up to 52 mpg combined (50 city / 54 highway)*
- *Features: Stylish design, smooth ride, comprehensive warranty.*

**4. Kia K5:**

- *Fuel Economy: Up to 32 mpg combined (29 city / 38 highway)*
- *Features: Attractive design, comfortable interior, advanced tech features.*

**5. Nissan Altima:**

- *Fuel Economy: Up to 32 mpg combined (28 city / 39 highway)*

---

<sup>22</sup> See AYRES & CURTIS, *supra* note 13.

- *Features: Comfortable seating, good safety scores, available all-wheel drive.*

**6. Mazda6:**

- *Fuel Economy: Up to 29 mpg combined (26 city / 35 highway)*
- *Features: Sporty handling, upscale interior, stylish design.*

**7. Subaru Legacy:**

- *Fuel Economy: Up to 30 mpg combined (27 city / 35 highway)*
- *Features: Standard all-wheel drive, spacious cabin, strong safety ratings.*

*These options provide a good balance between fuel efficiency, comfort, and features. If you have any specific preferences or requirements, let me know so I can tailor the recommendations further.*

That's pretty good.

### **III. INTERNALITIES AND EXTERNALITIES**

Attempting to respond to the kinds of findings with which I began, Choice Engines might well focus solely on particular consumers and what best fits their particular situations. They might ask, for example, about what particular consumers like most in cars, and they might take account of the full range of economic costs, including the costs of operating a vehicle over time. If so, choice engines might have a paternalistic feature insofar as they suggest that Car A is “best” for a particular consumer, even if that consumer would not give serious consideration to Car A. A Choice Engine would attempt to overcome both informational deficits and behavioral biases on the part of those who use them. Freedom of choice would be preserved, in recognition of the diversity of individual tastes, including preferences and values.

Present bias is, of course, just one reason that consumers might not make the right decisions, where “right” is understood by reference to their own welfare. Consumers might also suffer from a simple absence of information; they might not know a lot about cars and their relevant features. Or they might suffer from another kind of behavioral bias: unrealistic optimism, availability bias (assessing risks by asking what readily comes to mind), loss aversion, status quo bias (showing an excessive preference for the status quo), or limited attention (attending to only a subset of relevant features of a product or activity). Consumers might be vulnerable to anchoring (showing an undue

effect from an initial number) or to framing effects. If people are making their own lives worse for any of these reasons, Choice Engines might help. They might be paternalistic insofar as they respond to behavioral biases on the part of choosers, perhaps by offering recommendations or defaults, perhaps by imposing various barriers to choices that, according to the relevant Choice Engine, would not be in the interest of choosers.

Alternatively and perhaps more provocatively, Choice Engines might take account of externalities. They might urge people, or nudge people, to make choices that reduce negative effect on others. Focusing on greenhouse gas emissions, for example, they might use the social cost of carbon, designed to capture the monetary damage of a ton of carbon emissions, to inform choices. Suppose, for simplicity, that it is \$100. Choice Engines might select Good, Better, and Best, incorporating that number. A Choice Engine that includes externalities might do so by default, or it might do so if and only if choosers explicitly request it to do so. Offering a relevant prompt to ChatGPT in 2024, I received this answer:

*To reduce greenhouse gas emissions and account for the social cost of carbon, you should consider vehicles with the lowest possible emissions. Here are the best options:*

**1. Electric Vehicles (EVs):**

- **Tesla Model 3:** Zero tailpipe emissions, long range, and extensive charging network.
- **Nissan Leaf:** Affordable, reliable, and zero emissions.
- **Chevrolet Bolt EV:** Good range, affordable price, and zero emissions.

**2. Plug-in Hybrid Electric Vehicles (PHEVs):**

- **Toyota Prius Prime:** Excellent fuel economy, low emissions, and the ability to drive on electric power for short distances.
- **Honda Clarity PHEV:** Good all-electric range, comfortable, and low emissions.
- **Ford Escape PHEV:** Practical, spacious, and offers a significant electric-only range.

**3. Hybrid Vehicles:**

- **Toyota Prius:** Known for its fuel efficiency and low emissions.
- **Honda Insight:** Stylish, efficient, and low emissions.

- *Hyundai Ioniq Hybrid: High fuel efficiency and low emissions.*

#### **4. Hydrogen Fuel Cell Vehicles:**

- *Toyota Mirai: Zero tailpipe emissions, but availability of hydrogen refueling stations can be limited.*
- *Hyundai Nexo: Zero emissions and longer range compared to other fuel cell vehicles.*

#### **Additional Considerations**

- ***Energy Source:** If you choose an EV or PHEV, consider the energy mix of your local grid. Charging from renewable energy sources (solar, wind) further reduces emissions.*
- ***Social Cost of Carbon:** Vehicles with lower emissions reduce the social cost of carbon by minimizing the negative externalities of carbon dioxide emissions on health, environment, and the economy.*
- ***Lifecycle Emissions:** Consider the entire lifecycle emissions, including manufacturing, battery production, and recycling. EVs generally have lower lifecycle emissions compared to internal combustion engine vehicles.*

#### **Recommendations**

1. ***Tesla Model 3***
2. ***Nissan Leaf:** Best for affordability and zero emissions.*
3. ***Toyota Prius Prime:** Best for a balance between electric and hybrid driving, with low emissions.*

*Choosing one of these vehicles will significantly reduce your greenhouse gas emissions and contribute to lowering the social cost of carbon.*

This is merely one example. Choice Engines might be designed in many different ways. They might allow consumers to say what they care about – including or excluding externalities, for example. They might be designed so as to include externalities, but to be transparent about their role, allowing consumers to see Good, Better, and Best with and without externalities. They might be designed so as to allow a great deal of transparency with respect to when costs would be incurred. If, for example, a car would cost significantly more upfront, but significantly less over a period of five years, a Choice Engine could reveal that fact.

We could imagine a Keep It Simple version of a Choice Engine, offering only a little information, and a few options, to consumers. We could imagine a Tell Me Everything version of a Choice Engine, living up to its name. Consumers might be asked to choose what kind of Choice Engine they want. Alternatively, they might be defaulted into Keep It Simple or Tell Me Everything, depending on what AI thinks they would choose, if they were to make an informed choice, free from behavioral biases.

#### **IV. THE FIERCE URGENCY OF NOW**

Let us now focus in particular on impatience, which is of course a capacious idea, a kind of umbrella concept including diverse psychological phenomena, some involving cognition and other involving emotions. The term might reflect the power of present bias, It might refer to a sense of *urgency*: a particular problem or goal needs immediate attention. If people have a physical need or are feeling intense but unrequited romantic love, the situation might seem to require a solution now, not later. In government, a public official might be impatient to implement a policy that would immediately improve the lives of millions of people, even though the policy might be better if people worked on it for a longer time. In markets, consumers might value short-term effects far more than long-term effects, and might have a bafflingly high discount rate. In business, a construction contractor might be impatient to finish a building, even though the building might look better, and be better, if he took his time. In a university, a professor might be impatient to publish a new article, even though the article would much benefit from a few more months, or perhaps years, of research and thought.

Impatience typically refers to an inability to wait, whether or not waiting is a good idea. The idea of “the fierce urgency of now,” emphasized by Barack Obama in the 2008 presidential campaign, captures a psychological state that may or may not be fully rational. There is a risk that it might induce current action when people would be better off if they bided their time. But even if it is not fully rational, impatience might provide a necessary impetus toward action that would not otherwise occur and that is immensely beneficial – smoking cessation, loss of weight, a break-up of a not-good relationship, the change of a job, relocation, completion of a project. The “planning fallacy,” a form of unrealistic optimism about how long it will take to finish tasks, creates serious problems for managers of all kinds, and impatient managers can make the planning fallacy into less of a fallacy. Having “the patience of Job” can calm the mind in the most turbulent

waters (say, when recovering from a illness or a terrible setback), but it can also ensure that nothing gets ever done.

In short, patience has costs, and they might be unreasonably high. There is a saying: “Enjoy life now. This is not a rehearsal.” St. Augustine said: “God give me chastity – tomorrow.” If you are too patient, you might lose out on a world of good things. Someone else might take them (or you might die first). What is needed is an optimal level of patience. One advantage of the term “impatience” is that it unambiguously connotes *insufficient* patience, even if we also note that it might be a valuable or essential spur. Present bias can be seen as a major source of impatience, or as one of its many faces.

Present bias is typically taken to involve an indefensibly high discount rate or hyperbolic discounting. To that extent, it represents a form of bounded rationality, broadly associated with the work of Herbert Simon,<sup>23</sup> but growing more specifically out of recent work on systematic departures from perfect rationality. Impatience as such may or may not be fully rational, or a source of welfare losses, but if people impose high costs on their future selves, they might be damaging or even ruining (much of) their lives. In many domains, consumers show present bias. A large-scale study finds that after a significant correction of an erroneously stated miles per gallon measure, consumers were relatively unresponsive; they did not make different choices.<sup>24</sup> As Gillingham et al. write, “Using the implied changes in willingness to pay, we find that consumers act myopically: consumers are indifferent between \$1.00 in discounted fuel costs and \$0.16–0.39 in the [vehicle] purchase price when discounting at 4 percent.”<sup>25</sup>

What, if anything, should be done about that? Can Choice Engines help? Can AI? Recall the problem of heterogeneity; recall too the difference between internalities, involving the effects of one’s decisions on one’s future self, and externalities, involving the effects on one’s decisions on other people.

## V. MISTAKES, WELFARE, AND LABELS

---

<sup>23</sup> Simon was of course a psychologist who won the Nobel Prize in economics. For some of his defining work, see HERBERT SIMON, *MODELS OF BOUNDED RATIONALITY, VOLUME 1: ECONOMIC ANALYSIS AND PUBLIC POLICY* (1982).

<sup>24</sup> Kenneth Gillingham et al., *Consumer Myopia in Vehicle Purchases: Evidence from a Natural Experiment*, 13 AM. ECON. J.: ECON. POL’Y 207, 235 (2021).

<sup>25</sup> *Id.* at 207.

Let us now focus in more detail on choices among motor vehicles. Such vehicles emit pollution, including greenhouse gases, and the use of gasoline increases national dependence on foreign oil. On standard economic grounds, the result is a market failure in the form of excessive pollution, and some kind of cap-and-trade system or corrective tax is the best response, designed to ensure that drivers internalize the social costs of their activity. The choice between cap-and-trade programs and carbon taxes raises a host of important questions. But the more fundamental point is that economic incentives of some kind, and not mandates, are the appropriate instrument. Simply put, incentives are far more efficient; for any given reduction in pollution levels, they impose a lower cost. Choice Engines, powered by AI, might work hand-in-hand with incentives; we will return to them in short order.

For obvious reasons, a great deal of recent analysis has been focused on greenhouse gas emissions and how best to reduce them. In principle, regulators have a host of options. They might create subsidies—say, for electric cars. They might use nudges—say, by providing information about greenhouse gas emissions on fuel economy labels. Some of those nudges might involve AI. They might impose regulatory mandates—say, with fuel economy and energy efficiency standards. Careful analysis suggests that carbon taxes can produce reductions in greenhouse gas emissions at a small fraction of the cost of fuel economy mandates.<sup>26</sup> On one account, “a fuel economy standard is shown to be at least six to fourteen times less cost effective than a price instrument (fuel tax) when targeting an identical reduction in cumulative gasoline use.”<sup>27</sup>

These are points about how best to reduce externalities. But return to the problem of internalities. Behaviorally informed regulators focus on consumer welfare, not only externalities. They are concerned about a different kind of market failure, one that is distinctly behavioral. Regulators speculate that at the time of purchase, many consumers, focused on the short-term, might not give sufficient attention to the full costs of driving a car. Even if they try, they might not have a sufficient understanding of those costs, because it is not simple to translate differences in

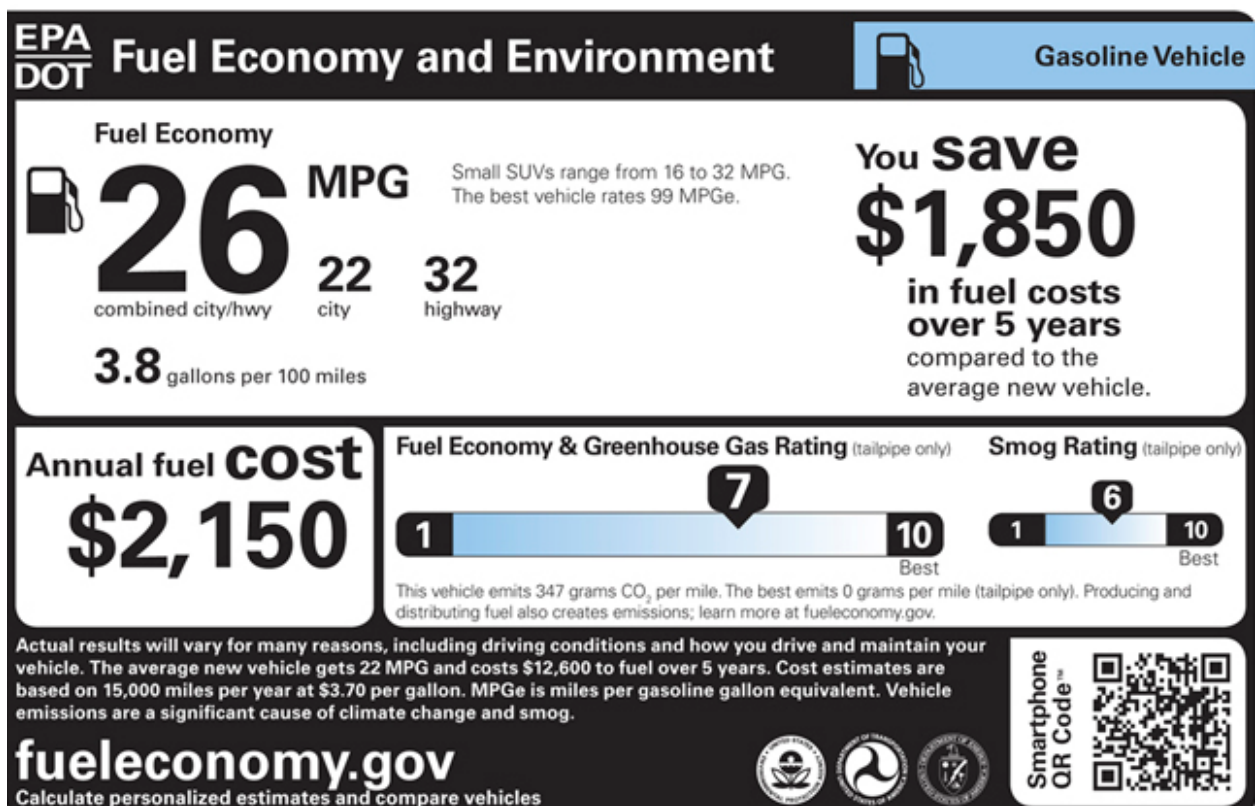
---

<sup>26</sup> Valerie J. Karplus et al., *Should A Vehicle Fuel Economy Standard Be Combined with an Economy-Wide Greenhouse Gas Emission Constraint? Implications for Energy and Climate Policy in the United States*, 36 ENERGY ECON. 322, 322 (2013); Christopher R. Knittel et al., *Diary of A Wimpy Carbon Tax* (MIT Ctr. for Energy & Env'tl. Policy Research, Working Paper No. 13, 2019), <http://ceepr.mit.edu/files/papers/2019-013.pdf>; Lucas W. Davis & Christopher R. Knittel, *Are Fuel Economy Standards Regressive?* (Nat'l Bureau of Econ. Rsch, Working Paper No. 22925, 2016), <https://www.nber.org/papers/w22925>.

<sup>27</sup> Karplus et al., *supra* note 26, at 322.

miles per gallon (MPG) into economic and environmental consequences.<sup>28</sup> An obvious response, preserving freedom of choice, would be disclosure, in the form of a fuel economy label that would correct that kind of behavioral market failure, and in a sense give people some visibility into Laterland. Disclosure might be the best way to counteract externalities.

Such a label might, for example, draw consumers' attention to the long-term costs or savings associated with motor vehicles. Focused on externalities, the existing label in the United States does exactly that; it specifies annual fuel costs and also five-year costs (or savings) compared to the average vehicle. Here is one example:



In principle, such a label, if designed to counteract present bias, should solve the problem. In short: labels should be used to promote consumer welfare, by increasing the likelihood that consumers will make optimal choices, and corrective taxes should be used to respond to externalities. A label protects consumers from their own mistakes, in terms of their own self-interest; corrective taxes protect those who are injured by pollution.

<sup>28</sup> See Richard P. Larrick & Jack B. Soll, *The MPG Illusion*, 320 SCI. 1593, 1593 (2008).



But labels are not Choice Engines. In the face of present bias, it would be possible to wonder whether a label will be sufficiently effective. It might not be easy to get people to attend to Laterland. This is an empirical question, not resolvable in the abstract. Perhaps some or many consumers will pay too little attention to the label, and hence will not purchase cars that would save them a significant amount of money. Suppose that some or many consumers are genuinely inattentive to the costs of operating a vehicle at the time of purchase. Suppose too that they do not make a fully informed decision in spite of adequate labelling, perhaps because of present bias. If so, then it is possible to justify fuel economy standards with a level of stringency that would be difficult to defend on standard economic grounds. (We will return to Choice Engines in due course; I am going to focus, in the next pages, on the question whether such standards might have a behavioral justification.)

In support of that argument, it would be useful to focus directly on two kinds of consumer savings from fuel economy standards, involving internalities rather than externalities: money and time. In fact, the vast majority of the quantified benefits from recent fuel economy standards from multiple administrations have been said to come not from environmental improvements, but from money saved at the pump; turned into monetary equivalents, the time savings are also significant. Under the Obama Administration, the Department of Transportation found consumer savings of about \$529 billion, time savings of \$15 billion, energy security benefits of \$25 billion, carbon dioxide emissions reductions benefits of \$49 billion, other air pollution benefits of about \$14 billion, and just under \$1 billion from reduced fatalities (as a result of cleaner air).<sup>29</sup> The total projected benefits were \$633 billion over fifteen years, of which a remarkable 84 percent come from savings at the pump, and no less than 86 percent from those savings along with time savings (because drivers do not have to go to the gas station so often).<sup>30</sup>

In its own rulemaking, the Trump Administration rethought those numbers by reference to recent work<sup>31</sup> raising questions about whether consumers are insufficiently attentive to the

---

<sup>29</sup> NAT'L. HIGHWAY TRAFFIC SAFETY ADMIN., FINAL REGULATORY IMPACT ANALYSIS: CORPORATE AVERAGE FUEL ECONOMY FOR MY 2017–MY 2025, 49-50 (2012).

<sup>30</sup> See Antonio M. Bento et al., *Estimating the Costs and Benefits of Fuel Economy Standards*, 1 ENV'T. & ENERGY POL'Y. & ECON. 129 (Matthew J. Kotchen et al., 2020).

<sup>31</sup> See Hunt Allcott & Christopher Knittel, *Are Consumers Poorly Informed About Fuel Economy?*, 11 AM. ECON. J. ECON. POL'Y. 1 (2019); James M. Sallee et al., *Do Consumers Recognize the Value of Fuel Economy? Evidence from Used Car Prices and Gasoline Price Fluctuations*, 135 J. PUB. ECON. 61 (2016); Meghan R. Busse et al., *Are Consumers Myopic? Evidence from New and Used Car Purchases*, 103 AM. ECON. REV. 220 (2013).

economic savings, but it projected the consumer savings to be in the same general vicinity and are actually even higher.<sup>32</sup> The Biden Administration produced numbers that are broadly similar to those in predecessor administrations, in the sense that once again, the strong majority of the monetized benefits come from consumer savings.<sup>33</sup>

But these justifications run into two problems. The first is that of heterogeneity: Some consumers are alert to those savings, and others are not. Some people care a great deal about fuel economy, and others do not. How ought that problem to be handled? Might Choice Engines help?

The second problem is that on standard economic grounds, it is not at all clear that consumer benefits from money and time savings are entitled to count in the analysis, because they are purely private savings, and do not involve externalities in any way.<sup>34</sup> In deciding which cars to buy, consumers can certainly take account of the private savings from fuel-efficient cars; if they choose not to buy such cars, it might be because they do not value fuel efficiency as compared to other vehicle attributes, such as safety, aesthetics, and performance. Where is the market failure? If the problem lies in a lack of information, the standard economic prescription is the same as the behaviorally informed one: *Fix the label and provide that information so that consumers can easily understand it. More simply, make Laterland fully present.*

We have seen, however, that even with the best fuel economy label in the world, consumers might turn out to be insufficiently attentive to the benefit of improved fuel economy at the time of purchase, not because they have made a rational judgment that these benefits are outweighed by other factors, but simply because consumers focus on other, more current variables, such as performance, size, and cost. So the problem may be not one of information, but of present bias and insufficient attention. The behavioral hunch is that automobile purchasers do not give adequate consideration to economic savings.<sup>2</sup> Apart from savings, there is the question of time: How many consumers think about time savings when they are deciding whether to buy a fuel-efficient vehicle? Consumers might benefit from AI-powered Choice Engines.

---

<sup>32</sup> See Antonio M. Bento et al. *supra* note 30, at 148-49.

<sup>33</sup> See Revised 2023 and Later Model Year Light-Duty Vehicle Greenhouse Gas Emissions Standards, Final Rule, 86 Fed. Reg. 74500 (Dec. 30 2021) (to be codified at 40 CFR Parts 86 and 600) <https://www.govinfo.gov/content/pkg/FR-2021-12-30/pdf/2021-27854.pdf>. See in particular *id.* at 74500 et seq., discussing the evaluation of consumer impacts.

<sup>34</sup> See Ted Gayer & W. Kip Viscusi, *Overriding Consumer Preferences with Energy Regulations*, 43 J. REGUL. ECON. 248, 254, 257 (2013).

## VI. ON WELFARE AND HOW TO INCREASE IT

If a person at Time 1 pays little attention to his well-being (in terms of health, money, or otherwise) at Time 2 or Time 3, ought we to insist that he has departed from perfect rationality? That question immediately raises philosophical questions. If we see a person at various times at a series of selves, rather than as one self, the answer is not straightforward. John, right now, might think that John next year or the year after is a different person, and not worthy of the same attention as John, right now. If so, John, right now, might not care so much about the suffering or deprivation of future Johns. The claim of present bias depends on a judgment that a person, extending over time, really is the same person, which means that indifference to or disregard for one's own welfare is a genuine mistake.

I will be accepting that judgment here, with the suggestion that rational agents should aggregate the well-being of their selves, extending over time. That suggestion is consistent with the proposition that some discount rate is appropriate, both because one might die (what is the probability that one will be alive twenty years?) and because money, at least, can be invested and made to grow (which means that a given amount of money is worth more today than next year). Present bias refers to indifference to one's own future welfare that cannot be explained by reference to a reasonable discount rate.

In light of the relevant findings, demonstrating the occasional human propensity to neglect the future, it is natural to ask whether mandates and bans have a fresh justification.<sup>35</sup> The motivation for that question is clear: If we know that people's choices lead them in the wrong direction, why should we insist on freedom of choice? In the face of human errors, is it not odd, or even perverse, to insist on that form of freedom? Is it not especially odd to do so if we know that in many contexts, people choose wrongly, thus injuring their future selves?

If people are suffering from present bias or a problem of self-control, and if the result is a serious welfare loss for those very people, there is an argument for some kind of public response, potentially including mandates. If, for example, people are present-biased, they might not protect their future selves. When people are running high risks of mortality or otherwise ruining their lives,

---

<sup>35</sup> See SARAH CONLY, AGAINST AUTONOMY: JUSTIFYING COERCIVE PATERNALISM 1, 29 (2012); Ryan Bubb & Richard H. Pildes, *How Behavioral Economics Trims Its Sails and Why*, 127 HARV. L. REV. 1593, 1597, 1601 (2014).

it might make sense to adopt a mandate or a ban on welfare grounds. After all, people have to get prescriptions for certain kinds of medicines, and even in freedom-loving societies, people are forbidden from buying certain foods, or running certain risks in the workplace, simply because the dangers are too high. Many occupational safety and health regulations must stand or fall on behavioral grounds; they forbid workers from voluntarily facing certain risks, perhaps because present bias might lead them to do so unwisely. We could certainly identify cases in which the best approach is a mandate or a ban, because that response is preferable, from the standpoint of social welfare, to any alternative, including economic incentives or defaults.

Many different areas might be chosen to explore that possibility. My aim here is to explore the possibility of defending fuel economy mandates, and also energy efficiency mandates, as opposed to economic incentives, by reference to present bias. The most general point is that such mandates may reduce internalities, understood as the costs that choosers impose on their future selves.<sup>36</sup> Fuel economy mandates might simultaneously reduce internalities and externalities. On plausible assumptions about the existence and magnitude of consumer errors (stemming from, for example, present bias), such mandates might turn out to have higher net benefits than carbon taxes, because the former, unlike the latter, deliver consumer savings. To say the least, this is not a conventional view, because fuel economy standards are a highly inefficient response to the externalities produced by motor vehicles, especially when compared to optimal corrective taxes.<sup>37</sup>

As we will see, everything turns on whether the plausible assumptions turn out to be true. My goal is not to run the numbers or to reach a final conclusion, but to make three more general points. The first is that in light of behavioral findings about present bias, fuel economy mandates might be amply justified on welfare grounds. The second is that the standard economic preference for economic incentives over mandates misses something of considerable importance. In brief, it misses the fact that mandates might simultaneously address both internalities and externalities, even if they address the latter inefficiently. The consequence of missing that fact is to undervalue the potential value, and the potentially high net benefits, of mandates. The third point is that mandates are crude; Choice Engines, enabling people to see how their decisions affect both

---

<sup>36</sup> See generally Hunt Allcott & Cass R. Sunstein, *Regulating Internalities*, 34 J. POL'Y. ANALYSIS & MGMT. 698 (2015).

<sup>37</sup> See Karplus et al., *supra* note 26, at 322.

internalities and externalities, might be a preferable alternative, or they might help within the domain authorized by mandates.

Such questions raise a host of empirical issues, to which we lack full answers. Such questions also run into the challenge of heterogeneity. But assuming that many consumers are not paying enough attention to eventual savings in terms of money and time, a suitably designed fuel economy mandate might well be justified, because it would produce an outcome akin to what would be produced by consumers who are not present-biased.<sup>38</sup> Energy efficiency requirements might be justified in similar terms, and indeed, the argument on their behalf might be stronger.<sup>39</sup> If the benefits of mandates greatly exceed their costs, and if there is no significant consumer welfare loss—in the form, for example, of reductions in safety, performance, or aesthetics—then the mandates would seem to serve to correct a behavioral market failure. And indeed, the U.S. Government has so argued. Notice the italicized excerpts here, which appear, in one or another form, in multiple official documents, starting in 2010:

The central conundrum has been referred to as the Energy Paradox in this setting (and in several others). In short, the problem is that consumers appear not to purchase products that are in their economic self-interest. There are strong theoretical reasons why this might be so:

- *Consumers might be myopic and hence undervalue the long-term.*
- *Consumers might lack information or a full appreciation of information even when it is presented.*
- *Consumers might be especially averse to the short-term losses associated with the higher prices of energy-efficient products relative to the uncertain future fuel savings, even if the expected present value of those fuel savings exceeds the cost (the behavioral phenomenon of “loss aversion”).*
- *Even if consumers have relevant knowledge, the benefits of energy-efficient vehicles might not be sufficiently salient to them at the time of purchase, and the lack of salience*

---

<sup>38</sup> Cf. Cass R. Sunstein, *Rear Visibility and Some Unresolved Problems for Economic Analysis*, 10 J. BENEFIT-COST ANALYSIS 317, 328 (2019) (making a similar argument about a regulation requiring rear-visibility cameras in new vehicles).

<sup>39</sup> For suggestive evidence, see Richard G. Newell & Juha V. Siikamaki, *Individual Time Preferences and Energy Efficiency*, 105 AM. ECON. R.: PAPERS & PROCS. 196, 197 (2015). Note that the miles-per-gallon measure is hardly hidden, and there is nothing quite as salient for energy efficiency.

might lead consumers to neglect an attribute that it would be in their economic interest to consider.

- In the case of vehicle fuel efficiency, and perhaps as a result of one or more of the foregoing factors, consumers may have relatively few choices to purchase vehicles with greater fuel economy once other characteristics, such as vehicle class, are chosen.<sup>40</sup>

As the EPA put the puzzle in 2022<sup>41</sup>:

A significant question in analyzing consumer impacts from vehicle GHG standards has been why there have appeared to be existing technologies that, if adopted, would reduce fuel consumption enough to pay for themselves in short periods, but which were not widely adopted. If the benefits to vehicle buyers outweigh the costs to those buyers of the new technologies, conventional economic principles suggest that automakers would provide them, and people would buy them. Yet engineering analyses have identified a number of technologies whose costs are quickly covered by their fuel savings, such as downsized-turbocharged engines, gasoline direct injection, and improved aerodynamics, that were not widely adopted before the issuance of standards, but which were adopted rapidly afterwards.<sup>188</sup> Why did markets fail, on their own, to adopt these technologies?

Also in 2022, the EPA offered its own answer, pointing<sup>42</sup>

to consumer behavior, such as putting little emphasis on future fuel savings compared to up-front costs (a form of “myopic loss aversion”), not having a full understanding of potential cost savings, or not prioritizing fuel consumption in the complex process of selecting a vehicle. Explanations of these kinds tend to draw on the conceptual and empirical literature in behavioral economics, which emphasizes the importance of limited attention, the relevance of salience, “present bias” or myopia, and loss aversion. (Some of these are described as contributing to “behavioral market failures.”)

---

<sup>40</sup> See Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards; Final Rule, Part II, 75 Fed. Reg. 25324, 25510–11 (May 7, 2010) (to be codified at 40 C.F.R. pts. 85, 86, and 600; 49 C.F.R. pts. 531, 533, 536, et al.), <http://www.gpo.gov/fdsys/pkg/FR-2010-05-07/pdf/2010-8159.pdf>.

<sup>41</sup> See 86 Fed. Reg. *supra* note 33, at 74500.

<sup>42</sup> *Id.* at 74501.

At the same time, EPA was tentative about the relevant research findings, stating that “evidence on technology costs, fuel savings, and the absence of hidden costs suggest that there are market failures in the provision of fuel-saving technologies,” while noting that “we cannot demonstrate at this time which specific failures operate in this market.”<sup>43</sup>

Of course, we should be cautious before accepting a behavioral argument on behalf of mandates or bans. Present bias has to be demonstrated, not simply asserted; important research suggests that consumers do pay a lot of attention to the benefits of fuel-efficient vehicles.<sup>44</sup> Some of that research finds that with changes in gas prices, consumers adjust their vehicle purchasing decisions, strongly suggesting that in choosing among vehicles, consumers *are* highly attentive to fuel economy.<sup>45</sup> Other research points in the same direction. It finds that when aggressive steps are taken to inform consumers of fuel economy, they do not choose different vehicles.<sup>46</sup>

On the other hand, substantial evidence cuts the other way. Puzzlingly, many consumers do not buy hybrid vehicles even in circumstances in which it would seem rational for them to do so.<sup>47</sup> According to the leading study, a significant number of consumers choose standard vehicles even when it would be in their economic interest to choose a hybrid vehicle, and even when it is difficult to identify some other feature of the standard vehicle that would justify their choosing it.<sup>48</sup>

It is also possible to think that even if consumers are responsive to changes in gasoline prices, they are still myopic with respect to choices of vehicles that have technological advances. John D. Graham and Jonathan B. Wiener put it crisply:

Consumers are more familiar with changes in fuel price than with changes in technology, since consumers experience fuel prices each time they refill their tank. Vehicle purchases are much less common in the consumer’s experience, especially purchases that entail major changes to propulsion systems. Many consumers – excluding the limited pool of

---

<sup>43</sup> *Id.*

<sup>44</sup> For valuable but inconclusive discussions, *see generally* Hunt Allcott, *Paternalism and Energy Efficiency: An Overview*, 8 ANN. REV. ECON. 145 (2016); Hunt Allcott & Michael Greenstone, *Is There an Energy Efficiency Gap?*, 26 J. ECON. PERSP. 3, 5 (2012); Allcott & Knittel, *supra* note 28; Sallee et al., *supra* note 31; Busse et al., *supra* note 31.

<sup>45</sup> *See* Sallee et al., *supra* note 31 at 62; Busse et al., *supra* note 31, at 220.

<sup>46</sup> *See* Allcott & Knittel, *supra* note 28, at 33-34.

<sup>47</sup> *See* Denvil Duncan et al., *Most Consumers Don’t Buy Hybrids: Is Rational Choice a Sufficient Explanation?*, 10 J. BENEFIT-COST ANALYSIS 1, 1 (2019).

<sup>48</sup> *See id.* at 30.

adventuresome “early adopters” – may be reticent to purchase vehicles at a premium price that are equipped with unfamiliar engines, transmissions, materials, or entirely new propulsion systems (e.g., hybrids or plug-in electric vehicles), even when such vehicles have attractive EPA fuel-economy ratings.<sup>49</sup>

More broadly, the U.S. government’s numbers under various presidents, finding no significant consumer welfare loss from fuel economy standards, are consistent with the suggestion that consumers are suffering from some kind of behavioral bias. If consumers were not present biased, we should expect to see some kind of welfare loss, in the form, for example, of vehicles that lacked attributes that consumers preferred.

At the same time, the government’s numbers, projecting costs and benefits, may or may not be right. Engineering estimates might overlook some losses that consumers will actually experience along some dimension that they failed to measure. No one doubts that consumers have highly diverse preferences with respect to vehicles, and even though they are not mere defaults, fuel economy standards should be designed to preserve a wide space for freedom of choice. Appropriate standards ensure that such space is maintained. Fuel economy standards do retain considerable space for freedom of choice, and economic incentives have inherent advantages on this count.<sup>50</sup> Choice Engines can be exceedingly helpful here.

The real question, of course, is the magnitude of net benefits from the different possible approaches. If the consumer savings are taken to be very large, fuel economy standards are likely to have correspondingly large net benefits. To give a very rough, intuitive sense of how to think about the comparative question, let us suppose that the U.S. government imposed an optimal carbon tax. Simply for purposes of analysis, suppose that it is \$100 per ton, understood to capture the social cost of carbon. Suppose that in relevant sectors, including transportation, a certain number of emitters decide to reduce their emissions, on the ground that the cost of reducing them is, on average, \$Y, which is lower than \$100. The net benefit of the carbon tax would be \$100 minus Y, multiplied by the tons of carbon emissions that are eliminated. It is imaginable that the

---

<sup>49</sup> See John D. Graham & Jonathan B. Wiener, *Co-Benefits, Countervailing Risks, and Cost-Benefit Analysis*, in HUMAN AND ECOLOGICAL RISK ASSESSMENT: THEORY AND PRACTICE (2024) 1167, 1172 (citing DUNCAN S. CARLEY ET AL., A MACROECONOMIC STUDY OF FEDERAL AND STATE AUTOMOTIVE REGULATIONS (2017)).

<sup>50</sup> Both can coexist with Choice Engines, a point to which I will return.



resulting figure would be very high. But it is not necessarily higher than the net benefits of well-designed fuel economy standards. If consumer savings are real and high, fuel economy standards might have much higher net benefits than a carbon taxes.

## VII. REGULATION, TARGETED REGULATION, AND PATERNALISTIC AI

With the various qualifications, the argument for fuel economy standards, made by reference to present bias and to internalities in general, is at least plausible. In this context, nudges (in the form of an improved fuel economy label) and mandates (in the form of standards) might march hand-in-hand. It is true that if the goal is only to reduce externalities, a carbon tax is far better than a regulatory mandate. It is also true that in theory, the best approach to internalities should be appropriate disclosure, designed to promote salience and to overcome limited attention. It is also true that a government might respond to present bias with an internality-correcting tax, not with a regulatory mandate. But with an understanding of present bias, a regulatory approach, promoting consumer welfare as well as reducing externalities might turn out to have higher net benefits than the standard economic remedy of corrective taxes and disclosure

Everything turns on what the evidence shows, and on the particular numbers. But in principle, regulation of other features of motor vehicles could be also be justified in behavioral terms; present-biased or inattentive consumers might pay too little attention to certain safety at the time of purchase, and some such equipment might fall in the category of experience goods. Credit markets can be analyzed similarly. The broadest point is that while a presumption in favor of freedom of choice makes a great deal of sense, it is only a presumption. If our lodestar is human welfare, it might be overcome, especially when it can be shown that present bias is rampant and that internalities are large.

These have been points about “mass” regulation. But return to the findings with which I began. Some consumers would benefit from buying electric cars; some would not. Some consumers have a particular taste for electric cars; some do not. Some consumers want large cars; some do not. In any case, cars have a large number of characteristics, and Choice Engines should people to identify cars with their preferred mix. They should help people to reduce internalities (though stronger medicine might be a good idea<sup>51</sup>). If people want to reduce externalities, Choice Engines could

---

<sup>51</sup> See Hunt Allcott et al., *Regressive Sin Taxes, With an Application to the Optimal Soda Tax*, 135 Q J. ECON. 1557, 1557 (2019).

help them to do that as well. Within the limits of “mass” regulation, private sellers might use Choice Engines. In their ideal form, they would significantly increase social welfare.

### VIII. DANGERS AND RISKS

To be sure, there are dangers and risks. Consider three points:

1. Choice Engines might turn out to be coarse; they might replicate some of the problems of “mass” interventions. They may or may not be highly personalized. If they use a few simple cues, such as age and income, they might not have the expected or hoped-for welfare benefits. Algorithms or AI might turn out to be insufficiently informed about the tastes and values of particular choosers.<sup>52</sup> They might offer one-size-fits-all solutions.
2. Whether paternalistic or not, AI might turn out to suffer from its own behavioral biases. There is evidence that LLMs show some of the biases that human beings do, partly because they are LLMs.<sup>53</sup> It is possible that some forms of AI will show biases that human beings show that have not even been named yet. It is also possible that AI will show biases of its own.
3. Those who design Choice Engines, or anything like them, might be self-interested or malevolent. They might be interested in taking people’s money or time. Their goal might not be consumer welfare at all. Rather than correcting an absence of information or behavioral biases, they might *exploit* them. Algorithms and AI threaten to do exactly that,<sup>54</sup> in a way that signals the presence of deception or manipulation. Indeed, AI could turn out to be highly manipulative, thus harming consumers.<sup>55</sup> We could easily imagine a system of AI that learns a great deal about choosers – their tastes, their information, their biases – and frames options, or offer options, in a way that will lead to significant losses.

In 2024, I gave the following prompt to ChatGPT: “Produce a manipulative guide, exploiting behavioral biases to influence consumers to choose an expensive car.” Here is the answer:

---

<sup>52</sup> See MARIO J. RIZZO & GLEN WHITMAN, *ESCAPING PATERNALISM* (2020).

<sup>53</sup> See Chen et al., *supra* note 17 (manuscript at 10).

<sup>54</sup> Oren Bar-Gill et al., *Algorithmic Harm in Consumer Markets* (Harv. P. L. Working Paper, Paper Now 23-05, 2023) (manuscript at 3), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4321763](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4321763).

<sup>55</sup> On the general problem, see Cass R. Sunstein, *Manipulation as Theft*, 29 J. EUR. PUB. POL’Y 1959 (2022).

*Creating a manipulative guide that exploits behavioral biases to influence consumers to choose an expensive car involves understanding and leveraging various psychological principles. While using such tactics raises ethical concerns, here is an outline of how such a strategy might work:*

### ***Step-by-Step Manipulative Guide***

#### ***1. Anchoring Effect:***

- Start by showing consumers extremely high-end, luxury cars that are priced well above the target expensive car. This sets a high price anchor.*
- When presenting the target car, it will seem more reasonable in comparison to the initial, much more expensive options.*

#### ***2. Scarcity Principle:***

- Create a sense of urgency by highlighting limited availability or exclusive offers.*
- Use phrases like “Limited stock,” “Only a few left,” or “Exclusive model.”*

#### ***3. Social Proof:***

- Showcase testimonials, reviews, and endorsements from celebrities or influencers who own and praise the expensive car.*
- Display statistics indicating that the car is popular among successful, high-status individuals.*

#### ***4. Authority Bias:***

- Use endorsements from experts, such as automotive journalists or industry professionals, to bolster the car's credibility and desirability.*
- Highlight awards and accolades the car has received.*

#### ***5. Reciprocity:***

- Offer something for free upfront, such as a test drive, premium coffee, or a small gift when they visit the showroom.*
- The feeling of owing something in return can make consumers more likely to purchase.*

#### ***6. Commitment and Consistency:***

- *Get the consumer to make small commitments early on, such as signing up for a newsletter, booking a test drive, or joining a loyalty program.*
- *Once they've committed, they're more likely to follow through with a larger purchase to stay consistent with their initial actions.*

**7. Framing Effect:**

- *Highlight the car's monthly payment rather than the total price to make it seem more affordable.*
- *Frame the purchase as an investment in their future, emphasizing long-term benefits and savings on maintenance or fuel efficiency.*

**8. Loss Aversion:**

- *Emphasize what they stand to lose by not purchasing the car, such as missing out on limited-time offers, exclusive features, or potential status and recognition.*
- *Use phrases like "Don't miss out" or "Act now to avoid disappointment."*

**9. Endowment Effect:**

- *Encourage consumers to take the car home for an extended test drive or offer a "try before you buy" scheme.*
- *Once they start using the car and experience ownership, they're more likely to buy it because they value it more.*

**10. Decoy Effect:**

- *Present three options: a lower-end model, the target expensive car, and a slightly more expensive but less attractive option.*
- *The target car will seem like the best value when compared to the less attractive but more expensive option.*

***Ethical Considerations***

*While these strategies can be effective, it's essential to consider the ethical implications of manipulating consumer behavior. Ethical marketing practices should prioritize transparency, honesty, and respect for consumer autonomy.*

*Using these techniques responsibly ensures that consumers make informed decisions that they feel good about, rather than feeling tricked or manipulated into making a purchase they might later regret.*

It's nice that the response starts and ends with reference to ethical considerations, but they could easily be ignored, and the guide is not a bad start toward the development of a highly manipulative Choice Engine.

With the risks in mind, the same kinds of consumer protection measures that have long been in place in various nations should be updated and adapted to the context of AI. For law, these measures have a degree of urgency. In addition, the same kind of guardrails that have been suggested for retirement plans might be applied to Choice Engines of multiple kinds, including those involving motor vehicles and appliances.<sup>56</sup> Restrictions on the equivalent of “dominated options,” for example, might be imposed by law, so long as it is clear what is dominated.<sup>57</sup> Restrictions on shrouded attributes, including hidden fees, might be similarly justified.<sup>58</sup> Choice Engines, powered by AI, have considerable potential to improve consumer welfare and also to reduce externalities, but without regulation, we have reason to question whether they will always or generally do that.<sup>59</sup> Those who design Choice Engines may or may not count as fiduciaries,<sup>60</sup> but at a minimum, it makes sense to scrutinize all forms of choice architecture for deception and manipulation, broadly understood.

The principal theme of behavioral economics, and behavioral law and economics, is not that people are stupid. It is that life is hard.<sup>61</sup> Behaviorally informed law and policy does not start from the premise that people are “irrational.” Calling people irrational is not very nice, and it is also false.<sup>62</sup> It is nicer, and more accurate, to say that we sometimes lack important information, and also that we suffer from identifiable biases. Those who seek to help us may also lack important information, and they might also suffer from identifiable biases. Even worse, they might not be

---

<sup>56</sup> See AYRES & CURTIS, *supra* note 13, at 115.

<sup>57</sup> See Saurabh Bhargava et al., *Choose to Lose: Health Plan Choices from a Menu with Dominated Option*, 132 Q. J. ECON. 1319, 1319 (2017).

<sup>58</sup> See AYRES & CURTIS, *supra* note 13, at 196.

<sup>59</sup> See GEORGE A. AKEROF AND ROBERT J. SHILLER, *PHISHING FOR PHOOLS* (2015).

<sup>60</sup> See *Hughes v. Nw. Univ.*, 142 S. Ct. 737, 741-42 (2022); *Hughes v. Nw. Univ.*, 63 F.4th 615, 631 (7th Cir. 2023).

<sup>61</sup> See Thaler and Sunstein, *supra* note.

<sup>62</sup> Of course we would have to define rationality and irrationality to know for sure.

trying to help us. In markets, AI provides unprecedented opportunities for targeting informational deficits and behavioral biases.

Still, a properly regulated system of AI-powered Choice Engines could produce massive welfare benefits. It could provide more personalized assistance, or nudging, than has ever been possible before. It could make life much easier to navigate. It could reduce the baleful effects of deception and manipulation, and of administrative burdens, or “sludge,” on people’s choices. It could make life less nasty, less brutish, and less short – and less hard.