Is Utility Concave?

Nathan Mather*

October 28, 2022

Click here for latest version [1]

Online Appendix

Abstract

This paper empirically estimates the degree to which the marginal utility of income changes across income groups. The estimation is based on survey responses indicating willingness to pay to avoid unpleasant experiences and relies on the assumption that the associated disutility is equal on average across income groups. This assumption implies that any differences in average willingness to pay for relief are entirely driven by differences in marginal utilities of income. The results suggest that marginal utility of income is constant across income groups, implying that (cardinal) utility is roughly linear in dollars.

*Department of Economics University of Michigan: njmather@umich.edu
Thank you to my committee members, James Hines, Ash Craig, Joel Slemrod, and Tanya Rosenblat for their comments and support on this project. As well as comments from Dylan Moore and Stephanie Karol

[1] If links are not working, newest version and online appendix can be found on NathanMather.com

1 Introduction

What exactly is utility and is it comparable across people? Is it measurable? Despite playing a ubiquitous role in economic theory, the definition of utility and its use within that theory varies. Most economics theory models individual behavior as the result of utility maximization. When utility is defined as ordinal numbers corresponding to the preference ranking of a rational individual (with complete and transitive preferences) than utility maximization is true by definition. This is a great strength of ordinal utility as a flexible and adaptable modeling tool. It doesn't matter *why* people choose something, if they do in fact choose it consistently, we can assign it more ordinal utility. Despite ordinal utility's flexibility, normative economics frequently requires the additional assumption that utility correspond to an individual's welfare or well-being. By building on this stricter interpretation of utility and using a new survey on willingness to pay to avoid unpleasant experiences, I estimate how much the marginal utility of income changes across income groups. The key identifying assumption is that the associated discomfort, or disutility, from these carefully chosen experiences is equal, on average, across income groups. This assumption implies that any differences in average willingness to pay for relief are entirely driven by differences in marginal utilities of income since the average marginal utility for the good is the same across income. The results suggest that marginal utility of income is constant across income groups, implying that cardinal utility is roughly linear in dollars.

Measuring utility has been recognized as an important step in policy assessment as early as 1871 when Jevons postulated that "We cannot really tell the effect of any change in trade or manufacture until we can with some approach to truth express the laws of the variation of utility numerically" (Jevons, [1871] 1911). While Jevons was optimistic about the future, stating it will eventually be a "mathematical problem of no great difficulty how to disentangle the functions expressing the degrees of utility of various commodities" (Jevons, [1871] 1911), a somewhat common view among late 19$^{\text{th}}$ and early 20$^{\text{th}}$ century scholars, it has turned out to be much more difficult (Fisher, 1927; Moscati, 2018). Monotonic transformations of utility do not affect the marginal rate of substitution, people's preference rankings, or their observable choices. Despite Jevon's vision of modern computing power disentangling utility from observed behavior, it is simply not a problem that can be solved without some additional assumption.

I overcome this problem by finding a particular class of goods that enable estimating the concavity of utility over characteristics like income. For now, let's just call this special good a widget. The key characteristic of this widget is that the welfare benefit of receiving a widget is the same, *on average*, across income. That is, utility from widgets can vary between individuals but is independent from income. Now if we see that low income Americans are willing to pay \$X, and high income Americans are willing to pay \$Y for a widget, we know that they value X and Y dollars the same as well. This means the ratio of the average utility of money between high income and low income Americans is Y/X. That ratio identifies the average concavity of total utility.

While this may seem surprisingly simple, the difficulty lies in finding real goods like the widget described. The marginal utility of any given good almost always depends on the quantity already consumed and the consumption of compliments and substitutes[2]. To illustrate how difficult this really is, consider the following extreme example. Suppose everyone shares the same cardinal utility function and we know the taste of chocolate is the same for everyone. In this exceptional case, it may seem that chocolate would work well as the widget from the example above, but it is unlikely to satisfy the necessary assumption. The marginal utility of a chocolate bar still depends on how much chocolate I already have as well as how many other desserts, food, or any compliments or substitutes to chocolate that I consume. A change in income relaxes a consumer's budget constraint. With a relaxed budget constraint, consumers will change the quantity and quality of compliment and substitutes goods they consume, and, by extension, this changes the marginal utility of chocolate. Even in the exceptional case where everyone has identical cardinal utility functions, the marginal utility of most goods is unlikely to be independent from income. If even this extreme example fails, what goods fit this assumption in the realistic case where everyone has different utility functions?

I field a survey that elicits the willingness to pay for relief from common minor pains. These questions are carefully chosen to plausibly satisfy the assumption that the marginal utility of relief is actually independent from income. Consider the following example from the survey.

> Imagine you bump your shin badly on a hard edge, for example, on the edge of a glass coffee table. What is the most you would pay in U.S. dollars to completely and immediately eliminate any pain caused by the situation described, as if the event never happened?

Is the marginal utility of this hypothetical good, relief from the pain of hitting one's shin, independent from income? First, consider if there are any compliment or substitute goods that would make bumping one's shin more or less desirable. There is nothing available to

[2]This refers to compliments and substitutes in utility when utility is cardinal and is equivalent to a cross partial in cardinal utility. This is how the terms were originally conceptualized prior to the ordinal revolution (Auspitz & Lieben, 1889; Moscati, 2018)

buy to immediately relieve the pain. Any medications will not be accessible or take affect before the pain has subsided naturally. Given the lack of alternatives treatments, this means that no one, regardless of income, is already consuming some quantity of pain relief because it does not actually exist. Everyone is considering a change from zero relief after bumping their shin, to total relief. The marginal utility is not dependent on a person's consumption bundle, and so the only difference in marginal utility comes down to personal pain tolerance. While pain tolerance likely varies from person to person, it is arguably uncorrelated with income. I fielded a survey asking four open ended questions like the one above as well as three binary yes or no questions with a price randomly selected. The full survey can be seen in the online appendix.

Surprisingly, the results indicate that the willingness to pay for relief from these situations, and by extension the marginal utility of income, does not change significantly across income. This means that cardinal utility is roughly linear in dollars. One implication of linear utility is that a utilitarian policymaker, who values only the sum total of utility, would have linear welfare weights. By extension, any policymaker that favors redistribution is not a pure utilitarian and must care about the distribution of utility in addition to the sum total of utility. This conclusion relies on the assumption that the utility function people maximize with their individual decisions corresponds to the same definition of utility policymakers value for collective decisions. If a policymaker's philosophical interpretation of utility *requires* the marginal value of a dollar to diminish with income, then the results in this paper indicate that the utility individuals maximize with their behavior and the utility this policymaker values are not the same.

The rest of the paper precedes as follows. Before diving into the analysis, section 2, "Why Measure Utility", outlines why cardinal preferences and interpersonal comparisons are useful. While I expect some readers will need the extra convincing, others may not; in which case, this section can be skipped. Section 3, "Theoretical Foundation", outlines the main assumptions of the model and lays the theoretical groundwork for the empirical model. Section 4, "Empirical Model", outlines the empirical model used to recover estimates for the concavity of utility. Section 5, "Survey Data", outlines the data from the survey, the survey method, and summary statistics about the population. Section 6, "Survey Analysis", details the analysis and results from my survey while section 7, "Discussion of Results", discuses the implications of the findings. Finally, section 8, "Conclusion", provides some final thoughts.

2 Why Measure Utility

Why would we want to measure utility? What value does understanding the intensity of individual preferences bring? Much of economic theory is compatible with ordinal utility that simply ranks a consumer's preferences. Normative economics, however, often uses preferences to make interpersonal comparisons and engage in discussions about welfare or well-being. This requires more of utility than a simple ordering of choices.

Consider the following hypothetical scenario. A city has enough money set aside in their budget to either expand public transit or widen the roads to allow more cars through at a time. Of course, which policy they *should* implement is partially a normative question with a wide range of potential considerations. Despite being a question of ethics, economics can still objectively aid in decision making by laying out the outcomes from each policy in a way that allows policymakers to discern the option that best coincides with their ethical normative goals. While economic work can take a general instrumental rationality approach and consider ethical concerns like racial disparities, inequality, justice, rights, or fairness, it often focuses on consequentialist and welfarist approaches to policy assessment that rely on aggregating preferences.

Economists have tried to find a helpful way to aggregate individual ordinal preferences for social choices. Sticking to ordinal preferences has practical appeal since preference rankings are directly observable through actions and require assuming only that people have complete and transitive preferences[3]. Arrow (1950), however, showed that aggregating ordinal preferences is not possible without the social choice mechanism having properties generally considered undesirable. This motivated even him to look for a way to think about cardinal utility and interpersonal comparisons (Arrow, 1978). Beyond Arrow showing the necessity of cardinality for social choice, ordinal preferences simply do not capture all of the relevant information for many normative decisions.

Returning to the above example about expanding public transit or widening roads, suppose the only information we have is that more people prefer widening roads. With only two options Arrow's impossibility theorem does not bind, and I expect many policymakers would propose widening the roads to reflect the majority of preferences (Arrow, 1950). However, suppose the folks favoring widening the roads favor it because it shaves a few minutes off their morning commute and they do not use public transit at all. Now on the other hand, if the people in favor of expanding public transit would now be able to commute to more lucrative job opportunities or save considerable time on their commutes, reasonable policymakers may change their decision in favor of public transit expansion. The level of benefits going to

[3] Although these are not necessarily weak assumptions

each person in addition to their ranking often provides use-full, relevant information[4].

The need for more information than ordinal preferences provide is consistent with a welfarist approach to policy assessment. A welfarist policymaker has some weight they place on different people's utility and summing those weighted values up will lead to the outcome that coincides with their values. In the above example, a welfarist would, in their perfect world, take the utility gain for each person under each potential outcome, multiply that utility gain by that person's corresponding welfare weight, and sum that product for each person under both policies. Whichever sum is higher, is the policy the welfarist prefers. Doing this is, of course, easier said than done. There are many potential outcomes to consider and weigh against each other. Implicitly, the information about commuting times and job opportunities is conveying something about the intensity of the benefits, but what about other concerns like carbon emissions? How can policymakers compare or aggregate the impact of all the various relevant outcomes?

Economic analysis often tries to collapse these concerns into a single dimension using measures based on willingness to pay (WTP) [5]. Willingness to pay has a clear theoretical connection to an individual's decision utility since trading money for a good implies one prefers the money less than or equal to the good traded for it. This is also consistent with the economic practice of using people's own preferences to infer what increases their well-being. While the connection to an individual's decision utility is clear, how to handle interpersonal comparisons is not.

Rather than addressing the tricky issue of interpersonal comparisons, economic work often resorts to efficiency arguments that use willingness to pay but rely only on an ordinal conception of utility for aggregation. This practice stems from the idea put forth by Hicks (1940) and Kaldor (1939) who proposed measuring economic efficiency as the sum of aggregate real income. The core idea of their argument is that, of course policy creates winners and losers, but if the winners can compensate the losers, the size of the economic pie has increased, and everyone can be made better off. This idea has become widely accepted in economic analysis and modern practitioners often use such arguments to justify using the sum of consumer and producer surplus as a measure of economic efficiency. The practice is ubiquitous, but some examples include industrial organization's analysis of consumer surplus in merger litigation (Glick, 2018; Wilson, 2019), cost benefit analysis, or theoretical policy arguments for things like price gouging (Zwolinski, 2008).

While an ethics free[6] efficiency is appealing, the idea does not hold up in a setting with

[4]Pearce (2021) explained the insufficiency of ordinal preferences in a similar way that was very helpful

[5]compensated and equivalent variation or consumer surplus

[6]It's not really ethics free since Pareto efficiency is itself a fairly strong consequentialist ethical assumption, but it avoids normative interpersonal comparisons

heterogeneous preferences. Samuelson (1950) shows how even in the case where everyone is truly made better off, an outcome cannot be said to be efficient unless it completely expands the utility possibility frontier for every heterogeneous group. The logic here is that although everyone may be better off than the status quo after implementing a policy, that policy might also make redistributing utility to a particular group more difficult. If you are a policymaker that wants to redistribute to that group, the policy is a bad idea despite the Pareto gain relative to the status quo. This can be illustrated by a policy that shifts the slope of the utility possibility frontier between two groups and creates an intersection rather than efficiently expanding the utility possibility frontier. In his own words, Samuelson (1950, pg 10) says that without comparing an "Infinite number of points, no acceptable definition of an increase in potential real income can be devised at the non-ethical level of the new welfare economics."

Samuelson's critique applies even when allowing for theoretically cost-less lump sum transfers, but as he continues to explain, movement along a utility possibility frontier would require "an ideally perfect and unattainable system of absolutely lump-sum taxes or subsidies (Samuelson, 1950, pg. 18)". Lump sum transfers to increase equity simply do not exist. Rather than mapping out a utility possibility frontier, then, an efficiency approach would require mapping out utility feasibility frontiers. On this front, there has been some interesting work. Coate (2000) lays out a theoretical foundation that Hendren (2020) follows up on. The core idea being that, rather than use lump sum transfers as the benchmark, a given policy needs to be compared to redistribution through other government levers. Hendren (2020) looks at the cost of redistribution through the income tax code as a benchmark for other policies. While this is a promising improvement on the lump sum transfer arguments, it does not absolve the need for interpersonal comparisons. This approach is effective for eliminating re-distributive policies when redistribution could be more efficiently done though the income tax code. It eliminates some polices from consideration, but even a perfect efficiency measure can only partially order outcomes by eliminating Pareto dominated policies. A policymaker may still face a difficult and complex decision among efficient outcomes and objective measures can facilitate a decision that is in line with the policymaker's goals.

A slightly different type of efficiency argument is also frequently employed in an attempt to separate economic analysis from interpersonal comparisons. Introductory economics textbooks frequently present policy analysis as a balance between equity and efficiency. Betsey Stevenson and Justin Wolfers' book lays out the idea clearly

> One argument for focusing on efficiency is that whenever economic surplus rises it's possible for those who benefit to compensate those who were harmed, and to do so in a way that ensures everyone's better off... In reality, it's rare for

> new policies to compensate the people they harm. Thus, the argument that it's possible to make everyone better off is just that, a possibility...consequently, real-world policy debates are rarely just about efficiency. They also focus on equity (Stevenson & Wolfers, 2019, Section 7.1).

This logic is helpful, but it also requires making interpersonal comparison. If we are not actually achieving a Pareto gain, then we are not avoiding interpersonal comparisons or sticking to ordinal preferences. Weighing a gain in consumer surplus against a change in economic inequality, for example, requires assuming that surplus corresponds to welfare or well-being in a way that can be compared across people and can be compared against the distribution of that well-being, equity. This is a fine and common assumption to make, but it is not an ordinal efficiency argument. It is an interpersonal comparison.

Hendren and Sprung-Keyser (2020) provide a helpful framework for thinking about this type of trade-off. They propose measuring the marginal value of public funds for individual policies and using the following logic for policy comparisons: "Given two policies, A and B, suppose MVPFA = 2 and MVPFB = 1. Then one prefers more spending on policy A financed by less spending on policy B if and only if one prefers giving \$2 to policy A beneficiaries over giving \$1 to policy B beneficiaries." As we discuss in Eastmond, Mather, Ricks, and Betts (2022), this logic works well if policies impact homogeneous groups but places a high informational burden on policymakers for heterogeneous populations. Most importantly, it again assumes utility is cardinal and corresponds to individual well-being. Understanding the concavity of utility under this assumption will allow us to better conceptualize these trade-offs.

Interpersonal comparisons are an unavoidable aspect of collective decision making in most practical applications. I build on the typical economic framework that assumes individual behavior is the result of maximizing cardinal utility where utility corresponds to individual preferences and individual well-being. What does this assumption imply about how utility changes with income or across individuals? How does this assumption fit into the re-distributive preferences of policymakers? Are policymakers true utilitarians under this definition of utility, or do they consider the total distribution of utility as well? Our understanding of these questions and by extension the core assumption being made would be greatly improved if we knew the concavity of utility over income. In the next section, I state and explain the assumptions needed to measure that concavity.

3 Theoretical Foundation

Key considerations like, "what is utility", are not clearly communicable with just mathematical equations, but are a vital aspect of measuring the concavity of utility. This section covers the two assumptions needed to identify the concavity of utility and maps out the theoretical and philosophical groundwork for the empirical model to come. This conceptual map can be seen in the diagram in figure 1 which walks through a simple two income case. As I explained in the introduction, if the average marginal utility of pain relief is the same for the two income groups, and each is willing to pay \$X and \$Y dollars respectively, then the ratio of the marginal utility of money between the two income groups is Y/X. Each of these equalities linking the boxes, however, requires an assumption.

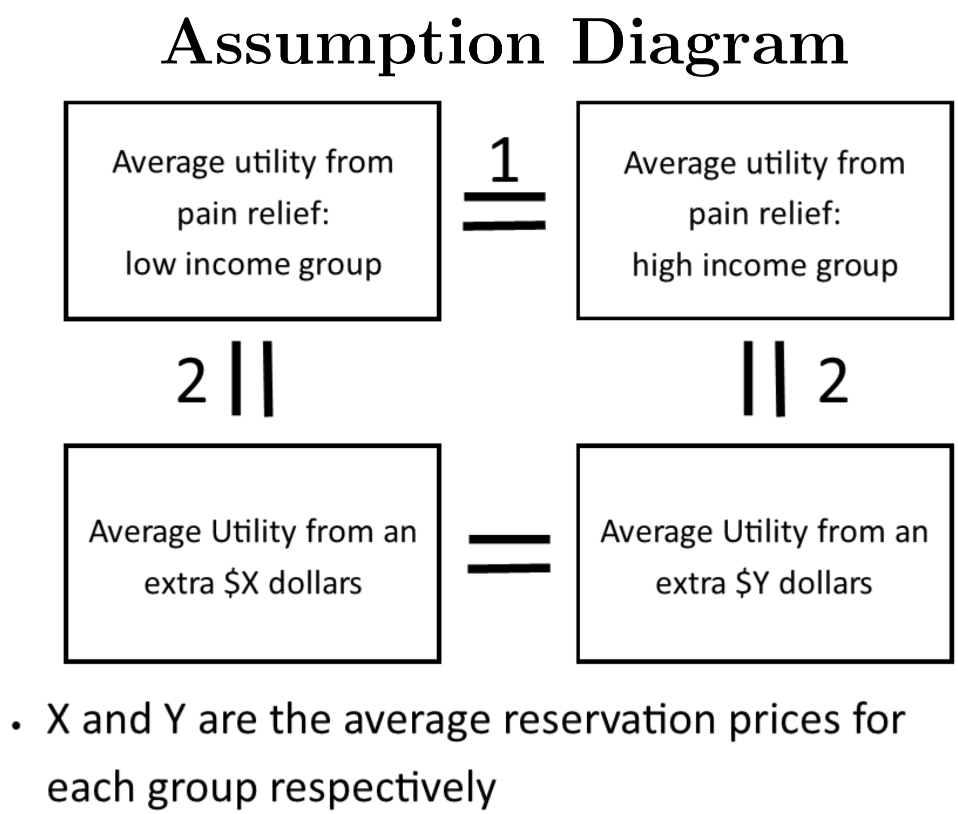

Figure 1

The first assumption needed for identification, and corresponding to equality 1 in the diagram, is the following

Assumption 1. *The Marginal utility of the pain relief described in the survey questions is independent of income.*

In other words, the average utility from pain relief in the survey scenarios is equal across income groups. Recall that the scenarios are things like bumping one's shin badly on a hard edge (full survey available here). There are no clear compliments or substitutes that make this experience better or worse and so different consumption bundles across income groups do not need to be considered to infer the marginal benefit of relief from this pain. We need only consider the experience itself.

Irving Fisher recognized that eliminating cross partials in cardinal utility, having no compliments or substitutes, facilitates identifying utility from revealed preferences[7] (Fisher, 1927). Instead of finding a particular good with this trait, however, he proposed using aggregates. The assumption being that "the utility derived from the consumption of each commodity group depends only upon the quantity of that commodity group that is consumed" (Morgan, 1945, pg 135). For example, the marginal utility of food must not depend on the type of housing or entertainment a person has. This simplifies the interpersonal comparison since we need only consider the marginal value of additional food, but richer people consume more and higher quality food; so, the marginal benefit of a bit more is conceivably different even without substitutes or compliments. To overcome this problem, Fisher proposes using different locations with a different price vector where people consume the same amount of food as the low-income group, but the same amount of housing has the high-income group. Morgan (1945) attempts to implement this method[8].

Using this method, however, requires solving a difficult if not impossible index number problem to determine equivalent food rations across regions with different populations and price vectors. What is the equivalent consumption of quality adjusted food for an American in dollars and an English person in pounds? If their preferences are not identical, the ability to purchase a given consumption basket in either country does not make the utility value of a given dollar expenditure on food equivalent to the average person in the two countries. This is a general problem with index numbers or measures of inflation under heterogeneous preferences (Samuelson, 1950; Samuelson & Swamy, 1974).

For the goods in my survey, like relief from bumping one's shin on a hard edge, there is no need to aggregate to avoid compliments and substitutes. Moreover, the marginal utility of relief is arguably equal across income levels without an intermediary comparison group. This avoids the index number problem requires in Fisher's approach. I have stated that marginal utility is arguably equal across income levels, but what does that actually mean? What is utility? An advantage of my survey questions is that they satisfy assumption 1 for

[7] For an interesting and more thorough overview of research on measuring utility see (Dimand, 2019) and (Moscati, 2018). I focus on Fisher's approach as it most closely mirrors my own.

[8] Or vise versa. The consumer with alternative prices needs to act as a link between any two broad class of goods

a wide range of philosophical definitions of utility.

Utility is a concept that is frequently used in economics and philosophy, but its definition is not consistent, and it is often used in differing and conflicting ways. Assumption 1 requires a cardinal definition where the experiences of two different people, or the average of two groups to be more exact, can be considered equal; but even among cardinal definitions of utility, there are a wide variety of views about what exactly utility is.

Hausman and McPherson (2006) provide a clarifying framework for grouping theories of well-being. Substantive theories say *what* things are inherently good. For example, happiness or pleasure could be considered the actual meaning of utility. The substantive approach fits with the work of utilitarian philosophers Jeremy Bentham, John Stuart Mill, and Henry Sidgwick (Driver, 2014). Assumption 1 is likely satisfied for these substantive theories of utility. It is reasonable to think, for example, that two people hitting their shins are experiencing the same pain or lack of pleasure.

Formal theories instead specify *how* to find out what is good, but not *what* is inherently good (Hausman & McPherson, 2006). In economics well-being is often considered to be the satisfaction of preferences and utility is the extent to which those preferences are satisfied. This is a formal theory that says what people want must give them utility. Irving Fisher actually preferred terms like "wantability" to avoid conflation of what economics is considering utility with the substantive theories of "Benthom and his school"(Fisher, 1927). However, some moral philosophers have accepted this as their preferred understanding of utility (Hare, 1981). If utility is the satisfaction of preferences how can we compare that across people?

Harsanyi (1955, 1986) proposed a way to think about how to compare the intensity of preferences across people using "extended preferences". This idea was even considered by Arrow (1978) as a way past the independence of irrelevant alternatives assumption in his impossibility theorem (Arrow, 1950). The idea proposes a thought experiment for interpersonal comparisons of utility. To compare the utility cost of two people hitting their shins on a coffee table we must imagine being each person and hitting our shin and compare those experiences. Harsanyi (1986) explains further that you must imagine yourself with that person's preferences. MacKay (1986) calls this the mental shoehorn trick (Hausman & McPherson, 2006).

While this seems to be the dominant theory for rationalizing interpersonal comparison of preferences, it is controversial. Arrow (2012, pg 115) says of it, "The principle of extended sympathy as a basis for interpersonal comparisons seems basic to many of the welfare judgments made in ordinary practice. But it is not easy to see how to construct a theory of social choice from this principle." Moreover, theoretical work has considered it unsatisfactory (Hausman, 1995; MacKay, 1986). The advantage of the particular assumption in this paper, however, is that we need not be capable of comparing every possible situation, only the ones in my survey which have been specifically chosen because the comparisons are relatively simple to make. Imagining oneself as a rich or poor person hitting their shin seems to fall into the category of basic welfare judgments arrow referred to. Since there are no compliments and substitutes, to compare the utility from relief across income we need only consider how pain sensitivity or preferences for pain relief differ across income.

The medical literature provides some helpful references for considering how painful experiences would change with income. First, what is pain? The International Association for the Study of Pain (IASP) defines pain as "An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" (Aydede, 2019, Section 1.2). Under this definition, people have "epistemic authority with respect to their pain: they seem to be incorrigible, or even infallible, about their pains and pain reports" (Aydede, 2019, Section 1.2). This means pain cannot be measured with a medical instrument, it must be elicited from the person experiencing it. The pain sensitivity questionnaire (PSQ) is a validated measure of pain sensitivity. First, respondents were asked to fill out the survey questions. Then, their responses were compared with questions asked during actual painful experiences (heat, cold, pressure, and pinprick) (Ruscheweyh, Marziniak, Stumpenhorst, Reinholz, & Knecht, 2009). These questions are used both for the willingness to pay questions, and to construct a control for general pain sensitivity in my survey.

While the very nature of pain makes it difficult to measure, the medical literature seems to suggest that, if anything, lower income people may be less tolerant to pain. Miljković et al. (2014) shows that participants with a lower household possession index, suggesting lower income, are more sensitive to pain. Research has also shown that chronic pain increases pain sensitivity (Ruscheweyh et al., 2012). These results align with the idea that richer folks may have access to goods that decrease pain sensitivity. While nothing can immediately relieve the pain from bumping my shin, perhaps the knowledge that I can reward myself for enduring that pain with an expensive treat makes the experience less difficult. Moreover, the research on chronic pain suggests that perhaps rich and poor folks are not coming from the same starting quantity of daily pain. If there are diminishing returns to pain relief, this would impact that benefit of relief from a marginal painful event.

If these relationships are true, estimates without controls will be biased since low-income folks will be getting more of a benefit from pain relief. Despite the findings described, I do not find the medical literature to be conclusive. The chronic pain research, for example, suffers from selection bias. If 10 people actually have a chronic pain condition, the people that are least tolerant to pain are the ones who will go and get a diagnosis. This is not really addressed in the medical studies I have read. This may be because it is not important for their purposes. Medical studies are interested in describing attributes of patients with a chronic pain diagnosis, and this is true even if it is because of selection. Additionally, studies assessing pain tolerance, like how long you can endure holding your hand in ice water, do not incentivize performance. Because of this, variance in performance may be attributable to variance in commitment levels to the study or a desire to look tough as much as differences in the epistemic experience (Stephens & Robertson, 2020). While the points above cast some doubt on how applicable the medical literature results are to this particular setting, if these studies are measuring pain sensitivity differences, then we can (and I do) control for pain sensitivity using the validated survey questions in the pain sensitivity questionnaire.

The second assumption needed for identification is about connecting our concept of welfare from assumption 1 to people's preferences and action. Specifically,

Assumption 2. *Individual behavior is the result of maximizing utility where utility holds the same meaning as in assumption 1. This implies if a person is willing to trade two goods, they provide the same cardinal utility.*

This assumption gives us the equalities labeled with a 2 in figure 1. Since people are willing to trade \$X or \$Y dollars for pain relief on average, they must provide the same utility.

For this to hold, people need to do what gives them the most utility under the same theory of utility used in assumption 1. The assumption that people maximize their own well-being goes hand in hand with the formal, preference utilitarian definition of utility commonly used in economics. If utility is defined to be higher for the things people prefer, then people's choices maximize their utility by definition. That being said, this assumption does not preclude a substantive theory, such as utility being pleasure, from also holding. If, for example "happiness is the ultimate object of preference, then it could be true both that well-being is the satisfaction of preference and that well-being is happiness" (Hausman & McPherson, 2006, Pg 119)

This assumption is a utilitarian theory of action, and it is important to clearly distinguish this from a utilitarian theory of ethics. Utilitarian ethics says the ethical action is the one that maximizes the most good for the most people. A utilitarian theory of ethics would apply to a policymaker who wants to maximize collective utility. A utilitarian theory of action requires people to act in their daily lives as if they ascribe to ethical egoism,[9] since an

[9] They must *act* like an egoist, but philosophical egoism requires a stronger statement on causality that is not important for our purposes. They may not choose an action *because* it maximizes self interest, but the assumption is that the actions happen to coincide with self interest regardless (Sen, 1977).

individual's utility can often come into conflict with what maximizes collective utility. The utilitarian theory of action is what is required by assumption 2.

This is a common assumption. The MVPF framework in Finkelstein and Hendren (2020); Hendren (2020); Hendren and Sprung-Keyser (2020) is a prominent example in public finance, but the consumer welfare standard in industrial organization or cost benefit analyses using a willingness to pay measures are all implicitly making this assumption.

Despite being commonly used in economics, this is also not an uncontroversial or unchallenged assumption. The potential problems come in two flavors. The first is broadly considered by the field of behavioral economics. People may not maximize their own well-being because they are bad at making choices. They would like to maximize their own well-being, but bounded rationality or behavioral mistakes get in their way. The second possibility is that, even if well informed, people may simply not act to maximize their own well-being. Sen (1977) makes this point clearer with a distinction between sympathy and commitment.

> The former corresponds to the case in which the concern for others directly affects one's own welfare. If the knowledge of torture of others makes you sick, it is a case of sympathy; if it does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment" (Sen, 1977, pg 326).

In either case, an ordinal utility defining choice would put a higher utility index on intervening to stop torture, but only in the former does the choice indicate a higher personal welfare from intervening. Assumption 2 assumes all actions are coming from sympathy, and are not being made in error.

The final equality in figure 1 follows by the transitive property and gives us the ratio of the average marginal utility of income between the two groups. While this diagram is a simple example with two groups, the same general logic applies to more groups or a continuous function across income. The next section builds on the theoretical foundation discussed above with an empirical estimation model.

4 Empirical Model

With these assumptions in mind, we can now consider a more explicit empirical model. The model starts with a specific functional form for cardinal utility, and thus, takes assumption 2 as given.

$$U(m_i, q_i, X_i, \epsilon_i) = \phi(m_i) + r(q_i, X_i, \epsilon_i) \tag{1}$$

Where ϕ is a function for the utility of income m and r is a function for the utility for relief from one of the painful experiences in my survey. q_i is the quantity of pain relief ranging from partial to total relief, X_i is a vector of observable characteristics that influence pain tolerance, and ϵ_i is an error term that represents differences in pain tolerance that are not captured by observable characteristics.

As a starting point, I have assumed that the marginal utility of income is constant across people, but I will relax that assumption below. It is crucial that $m_i \notin X_i$. That is, pain relief cannot also be a function of income. It is also crucial that the utility of pain relief and income are additively separable. This ensures there are not cross partials between income and pain relief. With these assumptions, we can see that the reservation price for person i for pain relief is the following

$$P_i^r = \frac{r'(q_i, X_i, \epsilon_i)}{\phi'(m_i)} \tag{2}$$

The price of money is normalized to one. $\phi'(m_i)$ is the marginal utility of income, and $r'(q_i, X_i, \epsilon_i)$ is the marginal utility of pain relief from one of the events in the survey questions. The equality is the indifference condition that comes from setting the marginal rate of substitution equal to the price ratio.

The math is simplified here by assuming that the change in pain relief is marginal. The questions in my survey are actually discrete, zero or total pain relief, but the responses are generally small relative to the concavity of utility and so should not pose a significant bias for estimation. Appendix B.2 shows this in more detail.

Building on equation 2, the following theorem shows when marginal utility of income can be identified with a conditional average.

Theorem 1. *If $r'(q_i, X_i, \epsilon_i) \perp\!\!\!\perp m_i$,*

then the following holds up to a normalization α

$$\phi'(m_i) = \frac{\alpha}{\mathbb{E}[P_i^r(m_i)|m]} \tag{3}$$

Note that the if statement in the theorem is equivalent to assumption 1. The proof follows from the assumption and is laid out below.

Proof. Start by taking the conditional expectation of both sides of equation 2. The expectations here are expectations across people for a given income level m. This gives

$$
\begin{aligned}
&\mathbb{E}[P^r(m_i)|m] = \mathbb{E}[\frac{r'(q_i, X_i, \epsilon_i)}{\phi'(m_i)}|m] \\
&= \frac{\mathbb{E}[r'(q_i, X_i, \epsilon_i)|m]}{\phi'(m_i)} \\
&= \frac{\mathbb{E}[r'(q_i, X_i, \epsilon_i)]}{\phi'(m_i)} \\
&= \frac{\alpha}{\phi'(m_i)} \\
&\implies \phi'(m_i) = \frac{\alpha}{\mathbb{E}[P_i^r(m_i)|m]}
\end{aligned}
$$

After conditioning on m, $\phi'(m_i)$ is a constant and so it can be removed from the expectation. Next, $r'_i \perp\!\!\!\perp m_i$ by assumption, and so we can remove the condition from the numerator. Finally, $\mathbb{E}[r'_i]$ is a constant that we can normalize to α

□

What is the interpretation of the above result given the normalization? The ratio of $\phi'(m_1)$ to $\phi'(m_2)$ is the same for any normalization regardless of m_1 and m_2. This means we can say things like the marginal utility at m_1 is twice the marginal utility at m_2. More generally, we have identified the concavity of utility.

An important point that is implicit in this model is that the price vector is fixed. In this simplistic model where money is a single good, this is not apparent, but in appendix B.1, I present the same model using indirect utility, preference heterogeneity, and infinitely many goods. With that approach, it is clearer that the model is conditional on a price vector, but the conclusions are otherwise the same. The implication of being conditional on a price vector is that a change in the relative price of goods, in particular the ratio of the average price of goods consumed by the rich and the poor, will shift the marginal utility of income function. I discuss the implications of this in more detail in section 7, "Discussion of Results".

The above equations say that the marginal utility of income is identified by the inverse of the conditional expectation of the reservation price for the special goods in my survey. This conditional average can be estimated in several ways. If income is categorical, we can simply take the average reservation price across income bins. Alternatively, the relationship between income and the reservation price can be estimating with any parametric or non-parametric estimation technique for a conditional average.

One simplification made in the work above is that everyone in the same income group has the same marginal utility of income. This might be a reasonable normalization for some government policy like income tax that redistributes based mainly on income, but how does it hold up when we think about the mental model of utility we used for comparisons in assumption 1? Does extended sympathy, for example, tell us it is identical? Probably not. For example, suppose two people have the same income but one has a lot of wealth. Likely the experience of losing some amount of money will be worse the less wealth one has. We can enrich the model by making the marginal utility of income individual specific and estimating the conditional average. The following theorem shows that, with an assumption on the error distribution somewhat analogous to homoscedasticity, the estimation strategy is unchanged.

Theorem 2. *Let*

$$\phi_i(m_i) = \mu_i \phi(m_i)$$

be individual specific utility from income where $\phi(m_i) = \mathbb{E}[\phi_i(m_i)|m]$ is the average utility at income m and μ_i dictates the individual specific deviation from the average.

If $\frac{r'(q_i,X_i,\epsilon_i)}{\mu_i} \perp\!\!\!\perp m_i$ then the following holds up to a normalization α

$$\mathbb{E}[\phi_i(m_i)|m] = \frac{\alpha}{\mathbb{E}[P_i^r(m_i)|m]} \tag{4}$$

Proof.

$$\begin{aligned}
\mathbb{E}[P^r(m_i)|m] &= \mathbb{E}[\frac{r'(q_i, X_i, \epsilon_i)}{\phi_i(m_i)}|m] \\
&= \mathbb{E}[\frac{r'(q_i, X_i, \epsilon_i)}{\mu_i \phi(m_i)}|m] \\
&= \mathbb{E}[\frac{r'(q_i, X_i, \epsilon_i)\frac{1}{\mu_i}}{\phi(m_i)}|m] \\
&= \frac{\mathbb{E}[\frac{r'(q_i,X_i,\epsilon_i)}{\mu_i}|m]}{\phi(m_i)} \\
&= \frac{\mathbb{E}[\frac{r'(q_i,X_i,\epsilon_i)}{\mu_i}]}{\phi(m_i)} \\
&= \frac{\alpha}{\phi(m_i)} \\
&= \frac{\alpha}{\mathbb{E}[\phi_i(m_i)|m]} \\
\implies \mathbb{E}[\phi_i(m_i)|m] &= \frac{\alpha}{\mathbb{E}[P_i^r(m_i)|m]}
\end{aligned}$$

The second line comes from the definition of $\phi_i(m_i)$, the third is algebra, the fourth comes from $\phi(m_i)$ being a constant conditional on m, the fifth lines follows from the independence assumption, the sixth is just normalizing a constant utility level to α, and the seventh is just rewriting $\phi(m_i)$ using it's definition. $\square$

In words, the required independence assumption is saying that the utility from pain relief is independent from income, and that the deviation in individual marginal utility from the mean, as a multiple of the mean, is independent income. An implication of this is that if the marginal utility of income doubles, the variance of the marginal utility would double as well. While not a directly testable assumption, if this holds the variance of the reservation price will change in proportion to the marginal utility of income. Since we also expect the mean to change in proportion to the reservation price, the variance of the reservation price should change in proportion with the mean.

If we relax the assumption about how the error in the marginal utility of income is distributed, then we get the following:

Theorem 3. *Let $\phi_i(m_i)$ be defined as in Theorem 2*

If $r'(q_i, X_i, \epsilon_i) \perp\!\!\!\perp m_i$ and $r'(q_i, X_i, \epsilon_i) \perp\!\!\!\perp \phi_i(m_i)|m$ then the following holds up to a normalization α

$$\mathbb{E}[\phi(m_i)|m] = \alpha \mathbb{E}\Big[\frac{1}{P_i^r(m_i)}|m\Big] \tag{5}$$

The proof for this is in appendix B.3. For each empirical estimation in the paper, I include robustness check with an estimation strategy analogous to theorem 3. Before showing the results or building on this model, the next section reviews the data collection, population, and summary statistics.

5 Survey Data

The data for my main analysis comes from a new survey fielded to 1747 respondents through the survey panel company Centiment[10]. Respondents are recruited through Facebook, LinkedIn, and partner networks to fill out surveys for money. While not a truly random sample, my respondents are matched to the census on age, race, gender, and region and these demographics are provided to me by Centiment.

The final full set of survey questions can be seen in the online appendix . The first two sets of questions are variations of validated question on the pain sensitivity questionnaire (PSQ) Ruscheweyh et al. (2009). The first set are open ended questions asking respondents

[10]website: `https://www.centiment.co/`

to report their willingness to pay. The second set asks a simple "yes or no" question of the form, "would you pay $X to relieve that pain", where X is randomly selected. The next set of questions is a subset of the PSQ asking respondents to say how painful scenario's would be on a 0-10 scale. Many of these overlap with the WTP questions to get a sense of individual aversion to particular events.

I chose the PSQ questions that were reported as the most severe, but also that resolve in a relatively short amount of time. The pain resolving in a short amount of time helps satisfy the assumption that the marginal utility of relief from these experiences is uncorrelated with income because other goods do not impact the experience. To better see why these are a good fit, consider an excluded question that asks about pain from a sunburn. Sunburn is much less likely to satisfy the independence assumption because the respondent's job or access to soothing medication might make a difference on the actual pain sunburn causes.

One question in the first section and one in the third ask about scenarios that typically would not be painful at all. I refer to these questions as "catch questions". These "catch questions" are used to determine who is engaging in good faith in the questions and paying attention. How exactly these are used to exclude responses is outlined in section 5.1. Additionally, in the third section of the survey there is a question asking respondents to enter 9 to ensure their full attention. Respondents who did not enter 9 for this are, at that point, removed from the survey.

The next section asks a few basic demographics. Marital status, number of children, education, employment. The final section of the survey asks about income and financial health. The financial health questions come from the Fin-health survey[11].

5.1 Protest Answers and Outliers

It is common in open ended willingness to pay surveys to receive protest answers to questions, yet there is not a unique strategy in the literature to handle them (Boyle, 2017, pg.110-111). The general ad hoc solution is to include additional questions that allow analysis to identify respondents who are not engaging with the questions in the desired way. I remove respondents from the survey based on their response to the "catch" questions I mentioned above. Specifically, shaking hands with someone who has a normal grip should not be painful. It may make sense in theory to exclude anyone who did not answer 0 to both of these catch questions. The situations are not painful and so rational utility maximizing respondents should not pay anything to relieve the pain and should enter 0 when asked. I do not use this strict of a cutoff for my base specification. It is possible some people perceive some pain

[11]Website: `https://finhealthnetwork.org/`

from these situations or simply did not consider 0 to be a viable option.

Instead of removing respondents with non-zero answers to the catch questions, I use the following conditions. First, I look at the response to the open-ended catch question. I exclude anyone who answered higher on this than any of the other painful open-ended questions. The thought here is that, while respondents may not think to enter zero, they should not say a greater amount than for questions that are clearly painful. Second, to check for attention, I remove anyone who answered the same thing for every question. This includes 263 respondents who entered zero for every question. Finally, I drop anyone who said more than $5. This is, admittedly, a bit of an arbitrary cutoff but the goal is to allow people who defaulted to low, non-zero answers without included people who were not thinking about the question.

The last condition is based on the response to the pain sensitivity questionnaire catch question. Here, I ask again about shaking hands with someone who has a normal grip but ask them to rank the pain from 0-10. I drop anyone who's answer to this question is greater than 3. I tried a few other more complicated conditions, like being less than the mean or less than all other PSQ questions, but most of the cases overlapped with the three or less condition anyway. Table 1 shows how many responses were dropped for meeting each condition. 1021 responses remain after all of the drop conditions.

Drop Condition Counts

ALL IDENDICAL	NOT THE MINIMUM	WTP GREATER THAN 5	PSQ GREATER THAN 3	N
Pass	Pass	Pass	Pass	1021
Fail	Pass	Pass	Pass	258
Fail	Pass	Fail	Pass	6
Fail	Pass	Pass	Fail	22
Fail	Pass	Fail	Fail	20
Pass	Fail	Pass	Pass	19
Pass	Fail	Pass	Fail	11
Pass	Fail	Fail	Pass	47
Pass	Fail	Fail	Fail	112
Pass	Pass	Fail	Pass	84
Pass	Pass	Fail	Fail	63
Pass	Pass	Pass	Fail	84

Table 1: Each Row indicates what conditions are passed or failed. One failure leads to the response being dropped. The total for each combination of conditions are in the N column

For the open response questions, I also top coded responses. The maximum answer, for example, was one trillion dollars. I top-code them rather than dropping them since top-coding uses some of the information. Whoever entered such a high amount likely has a true amount that is high as well. A standard rule of thumb is to add 1.5 times the inter quartile range to the 75th percentile and treat anything above that as an outlier. As my data is skewed, I use 4.5 times the IQR for each income group. The number of top codes can be seen in table 2. The distributions for each open response question after removing protest answers and top coding outliers can be seen in figure 2.

Top Code Counts

QUESTION	OUTLIERS TOP CODED	PERCENT TOP CODED
Lemon Juice	62	6%
Hot Pot	68	7%
Burn Tongue	56	5%
Bump Shin	55	5%

Table 2

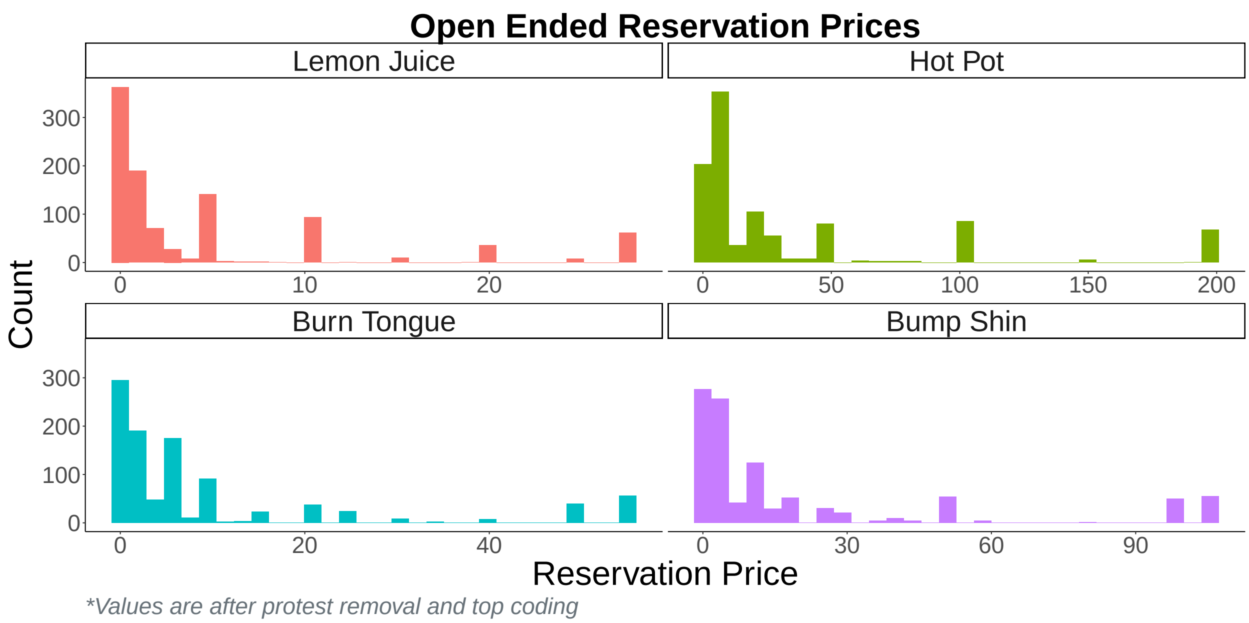

Figure 2

5.2 Summary Statistics

Before jumping into the analysis of willingness to pay and marginal utility, it is important to understand the population. While Centiment matched my survey sample to the census on age, race, gender, and census region, some respondents were dropped from the sample as described above. Table 3 shows rates for the matched demographic variables in my final sample compared to the census. Income was not matched explicitly, but figure 3 compares the results of my survey to the CPS family income. My survey is underrepresented in the $200,000 and above categories but appears roughly proportional in terms of income. Having a representative sample makes the results of the survey externally valid, meaning the mean WTP conditional on income should represent the actual mean in the United States. To gauge internal validity, we also want to consider how pain sensitivity differs across income.

Age Percent Comparison

AGE	CENSUS PERC	PERCENT
18 to 24 years	12.0	15.8
25 to 29 years	9.2	7.7
30 to 34 years	8.8	9.0
35 to 39 years	8.4	8.8
40 to 44 years	7.9	7.4
45 to 49 years	8.1	8.8
50 to 54 years	8.3	7.3
55 to 59 years	8.6	7.1
60 to 64 years	8.0	5.1
65 to 69 years	6.8	7.1
70 to 74 years	5.2	7.3
75 to 79 years	3.6	4.5
80 to 84 years	2.4	3.1
85 years and over	2.6	0.8

Gender Percent Comparison

GENDER	CENSUS PERC	PERCENT
female	51.3	49.8
male	48.7	50.2

Race Percent Comparison

RACE	CENSUS PERC	PERCENT
White	75.1	74.8
Black	14.2	11.5
Native Or Pacific Islander	2.2	1.5
Asian	6.8	3.0
Other	7.4	7.1
No Response	NA	2.2

Region Percent Comparison

REGION	CENSUS PERC	PERCENT
Midwest	24.0	21
Northeast	17.6	17
South	38.5	38
West	19.9	24

Table 3: *Values are after protest answer removal

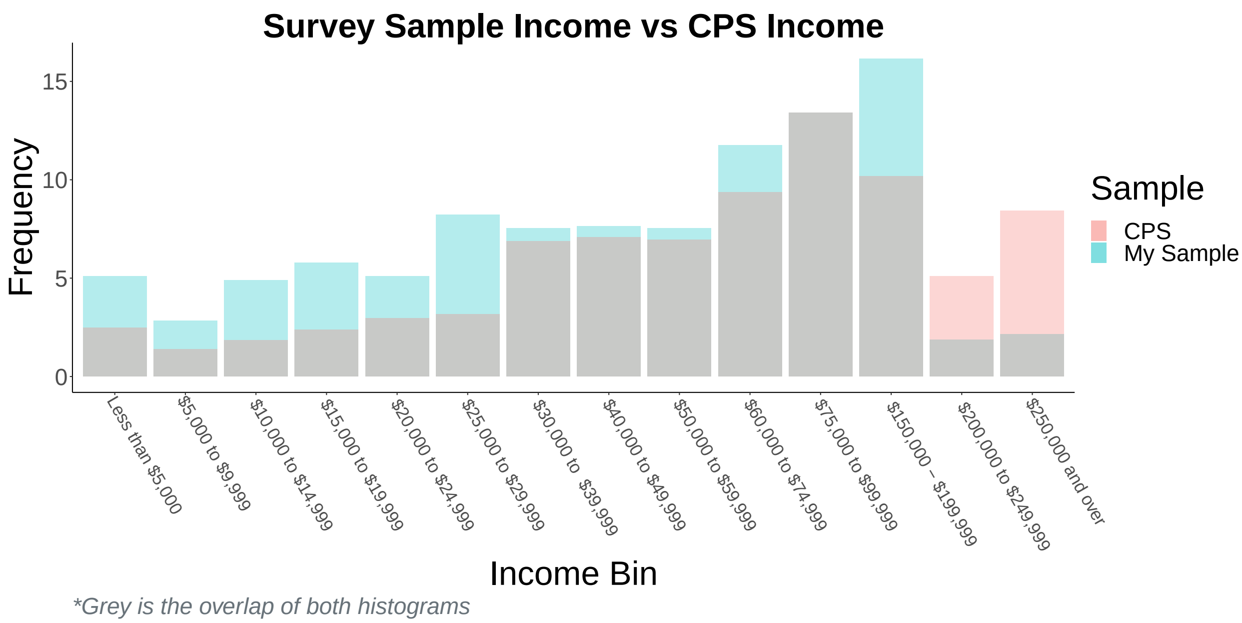

Figure 3

Figure 4 shows the relationship between pain sensitivity, measured by the mean pain sensitivity questionnaire score (PSQ) from 0-10, and income. While the point estimate for the slope is negative, it is not statistically significant. The table within the graph shows the results of the simple regression from the plot. A one hundred thousand dollar increase in income is associated with a .12 decrease in average PSQ response (on a scale from 0-10). Not only is this statistically insignificant, but it is also practically small relative to the within income variation. Nevertheless, I do control for the average responses to these questions in section 6.2.2.

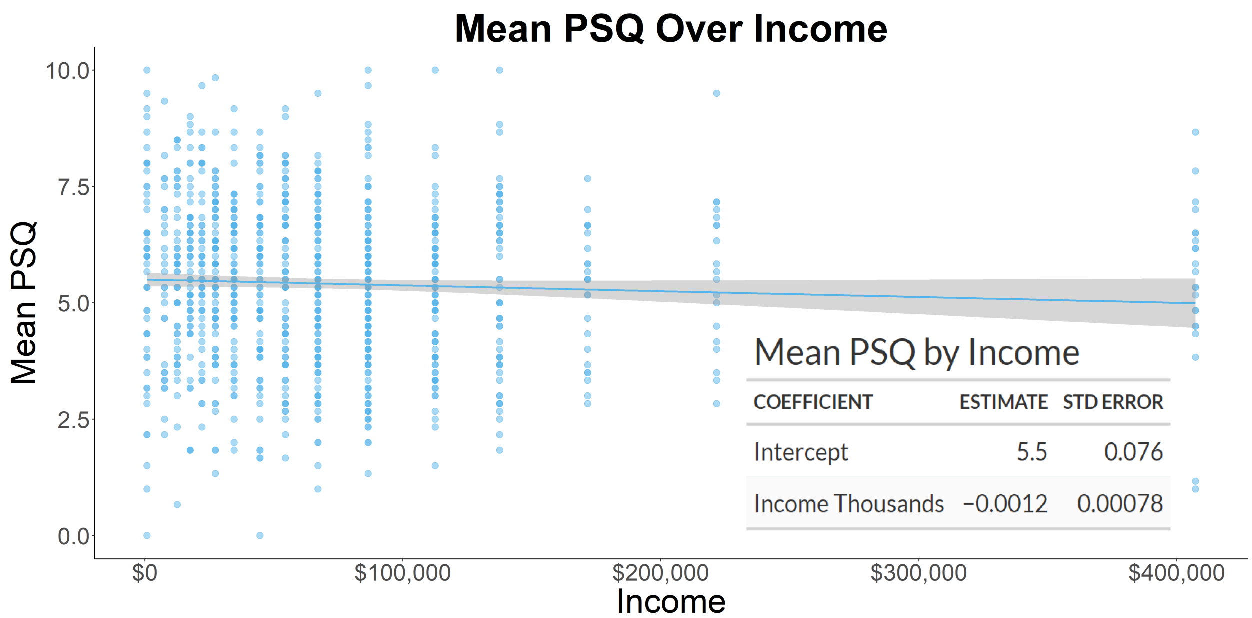

Figure 4: Mean PSQ is each respondents average over the 0-10 style pain sensativity questions

6 Survey Analysis

6.1 Open Ended Mean Results

As I showed in theorem 2, the average marginal utility of income as a function of income is equal to the reciprocal of the expected reservation price conditional on income. The conditional expectation of the reservation price can be calculated in a variety of ways. The first method is to estimate the mean for four income groups. The size of each group can be seen in table 4.

Aggregated Income Counts

Income	Count
0-25	242
25-50	239
50-100	334
more than 100	206

Table 4

Figure 5 plots the mean willingness to pay for relief for each income bin in each question. The vertical axis normalizes the lowest income group to 1. This is done because the willingness to pay only identifies the marginal utility of income, up to a normalization. So, what is actually of interest is the ratio of responses across incomes, not the absolute values. Normalizing to one makes it easier to compare the implied marginal utility of income across questions. The normalized mean response is marked with the large diamond, and bootstrap standard errors are included for each mean. The un-normalized values are displayed in text to the left of each point as well. The theory of diminishing marginal utility of income would predict a rise in WTP, but, surprisingly, the point estimates are relatively flat or even decreasing a bit.

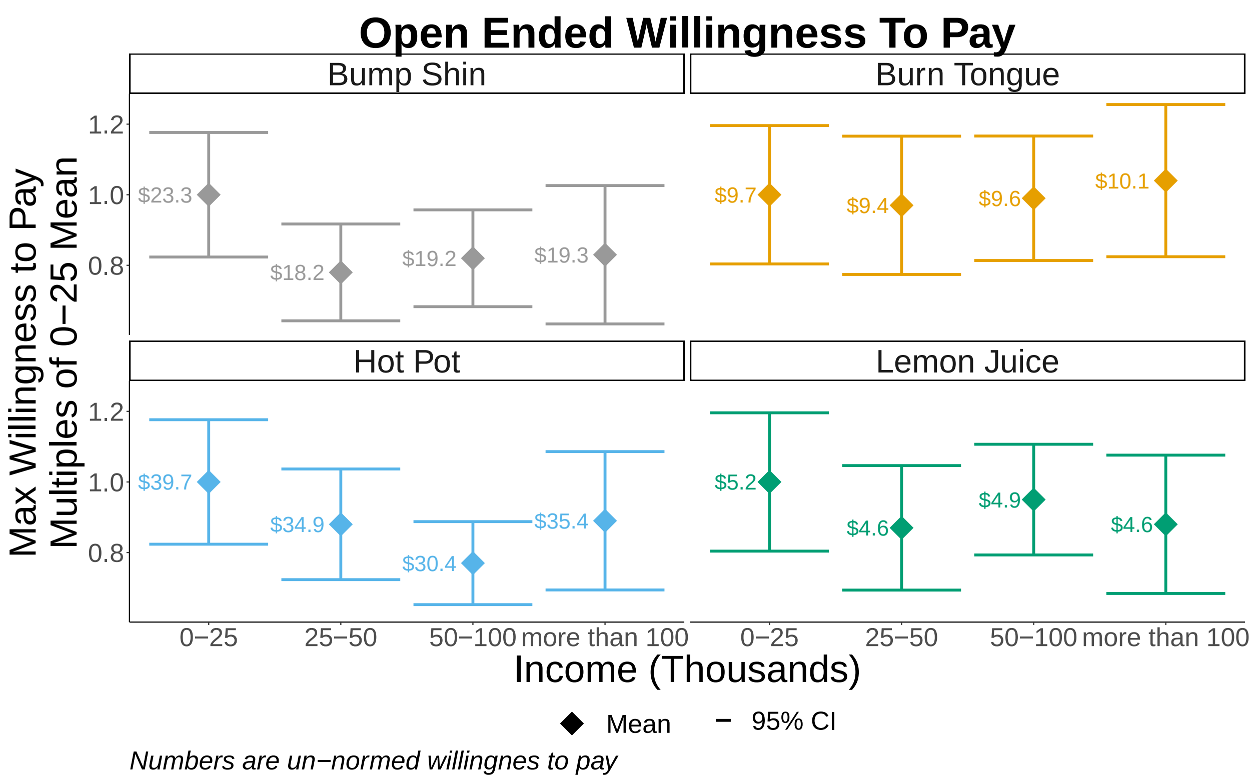

Figure 5

The second method for estimating the conditional expectation of the reservation price is to make a parametric assumption about the conditional average. Figure 6 shows every response to the open-ended questions across the full range of income categories. Each point is sized for the number of people in that income group with that response and a quadratic polynomial is fit for each question. Again, the results indicate little to no increase in WTP.

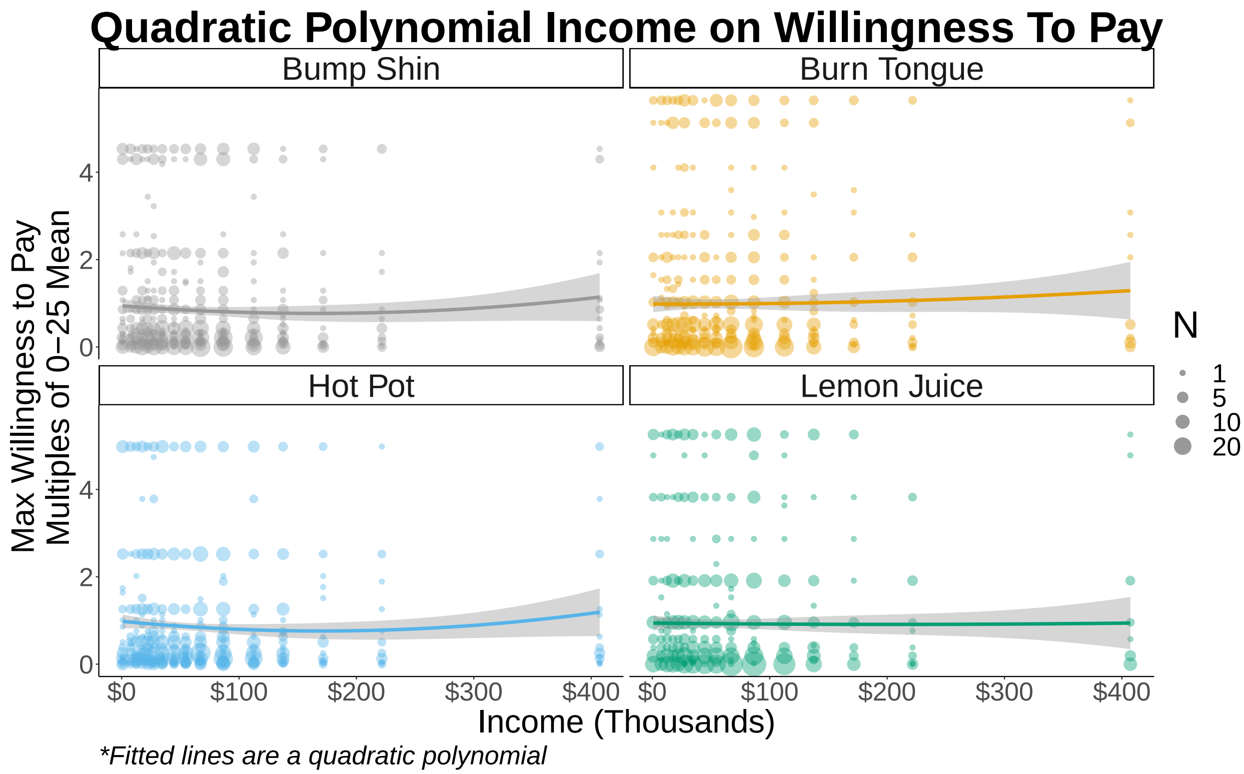

Figure 6

One concern with this approach discussed in section 4, is that the assumption from theorem 2, that the error in marginal utility is independent from income, does not hold. As I also explained above, if the error is independent from income, we would expect the standard deviation and mean to change in lock-step with one another over income. Table 5 in appendix A includes the mean and standard deviation for each income group and each question as a ratio of the $0-25k group. They both appear relatively constant across income, which supports the assumption. Given this support, I will continue to rely on this assumption for the main analysis, but I include robustness checks for each estimation strategy. The first of these alternative strategies can be seen in appendix A table 6 where the mean of the inverse of willingness to pay is shown. This is in line with the result from theorem 3.

Another concern with this simple approach is that perhaps pain tolerance is not independent from income. This would be an issue for every theorem in section 4. While income may not itself change the marginal utility of pain relief in these situations, there may be characteristics that both impact pain tolerance and correlate with income. Gender and age have been suggested to affect pain tolerance and by extension the utility from pain relief (Bartley & Fillingim, 2013; Lautenbacher, Peters, Heesen, Scheel, & Kunz, 2017). These are both related to income in my sample. Figure 7 shows that men are more concentrated in the high-income categories. Figure 8 shows that age increases some with income. The average goes from 42 in the lowest income group to 50 in the highest. Surprisingly, neither age nor gender has a significant relationship with the pain sensitivity questionnaire (PSQ). One way to further investigate pain tolerance is to look at the unconditional relationship with WTP for relief to see if there are differences across groups. There is not a significant relationship for gender and WTP on the open-ended questions. There is, however, a significant relationship with age. Older respondents are willing to pay less on average, as we can see in figures 9, despite having higher average incomes. These observations at the very least, motivate adding gender and age controls as a robustness check. While there is no apparent relationship between average PSQ responses and income, I include the PSQ as a control as well.

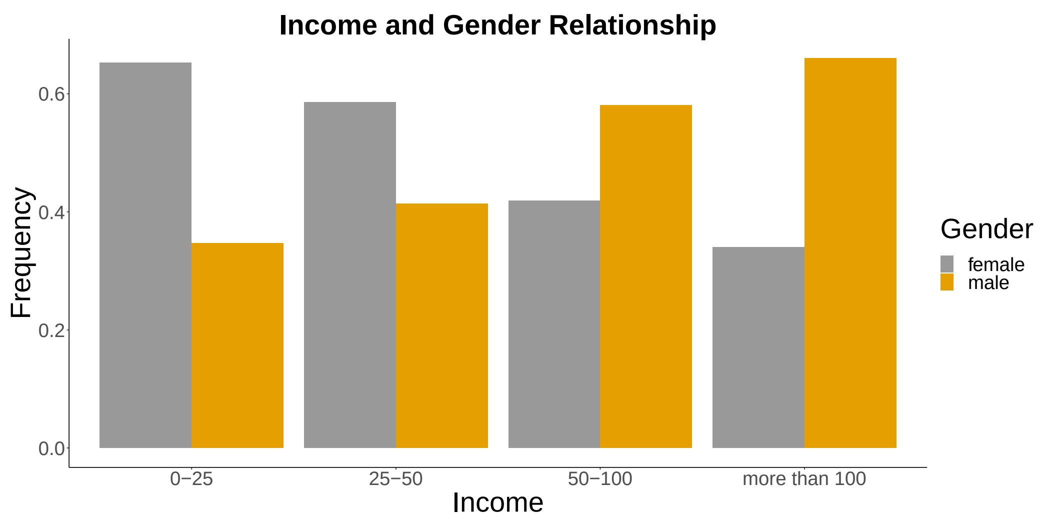

Figure 7

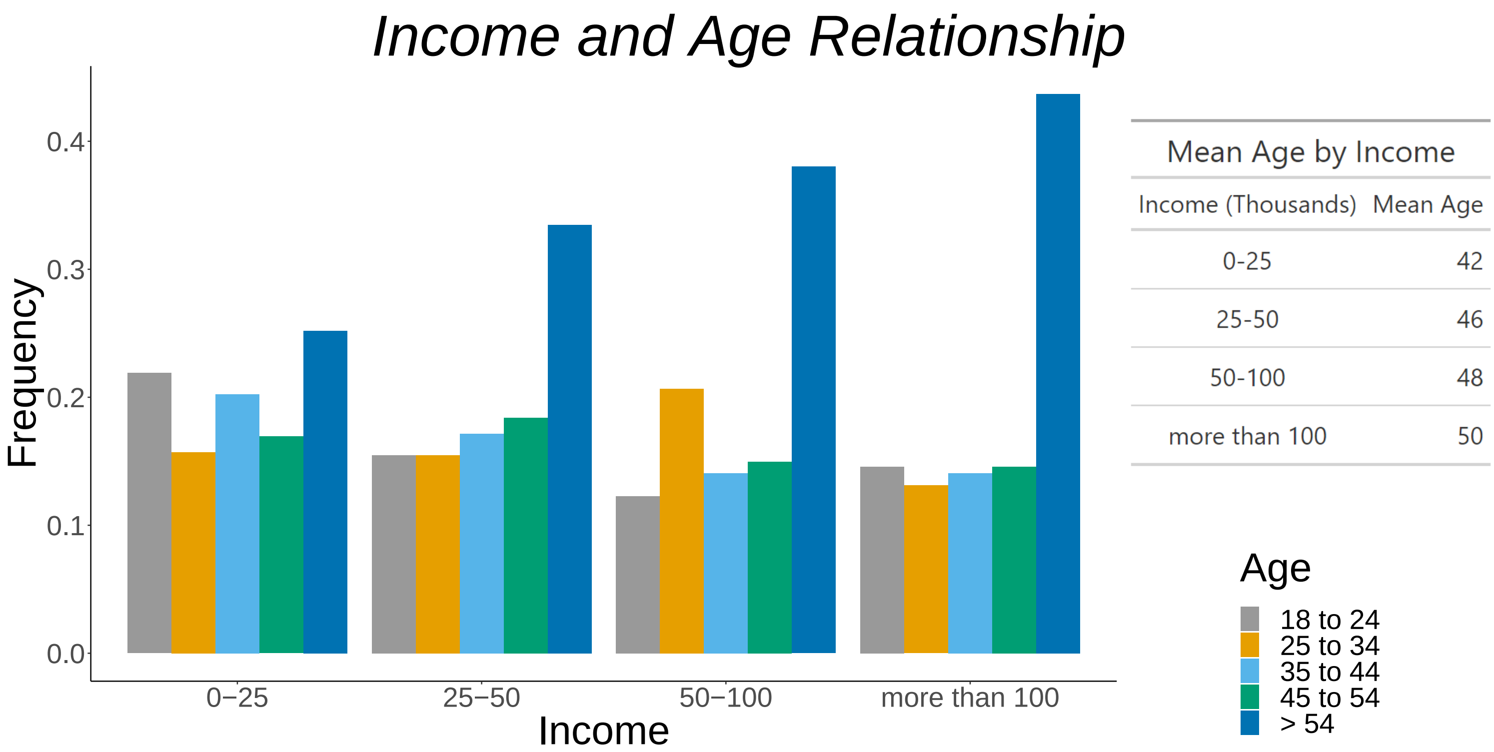

Figure 8

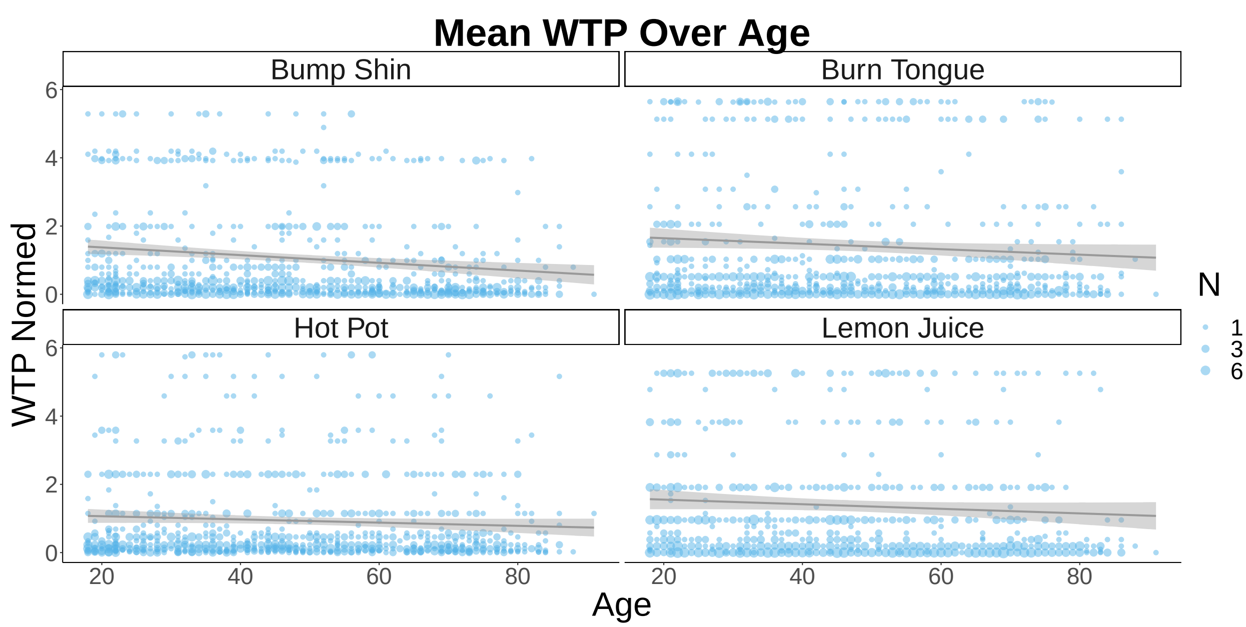

Figure 9

Unfortunately, we cannot simply add controls as a linear parameter into our model of conditional averages using ordinary least squares (OLS). This is because our outcome is in terms of WTP, but the difference between groups, like age, is in utility. Suppose men are more willing to endure pain on average. Pain relief would give men some fixed α fewer utils than women. Figure 10 shows a hypothetical relationship that assumes marginal utility is diminishing with income. As income increases, that α util gap translates into a larger and larger gap in WTP. In fact, that gap grows in proportion to the marginal utility of income. This means a simple binary control would not be sufficient. An OLS model with interactions would allow for the differing slopes in figure 10, but this sacrifices an opportunity to use the difference in the slopes to help identify the marginal utility of income. To fully take advantage of this relationship, I use a maximum likelihood estimation (MLE) technique.

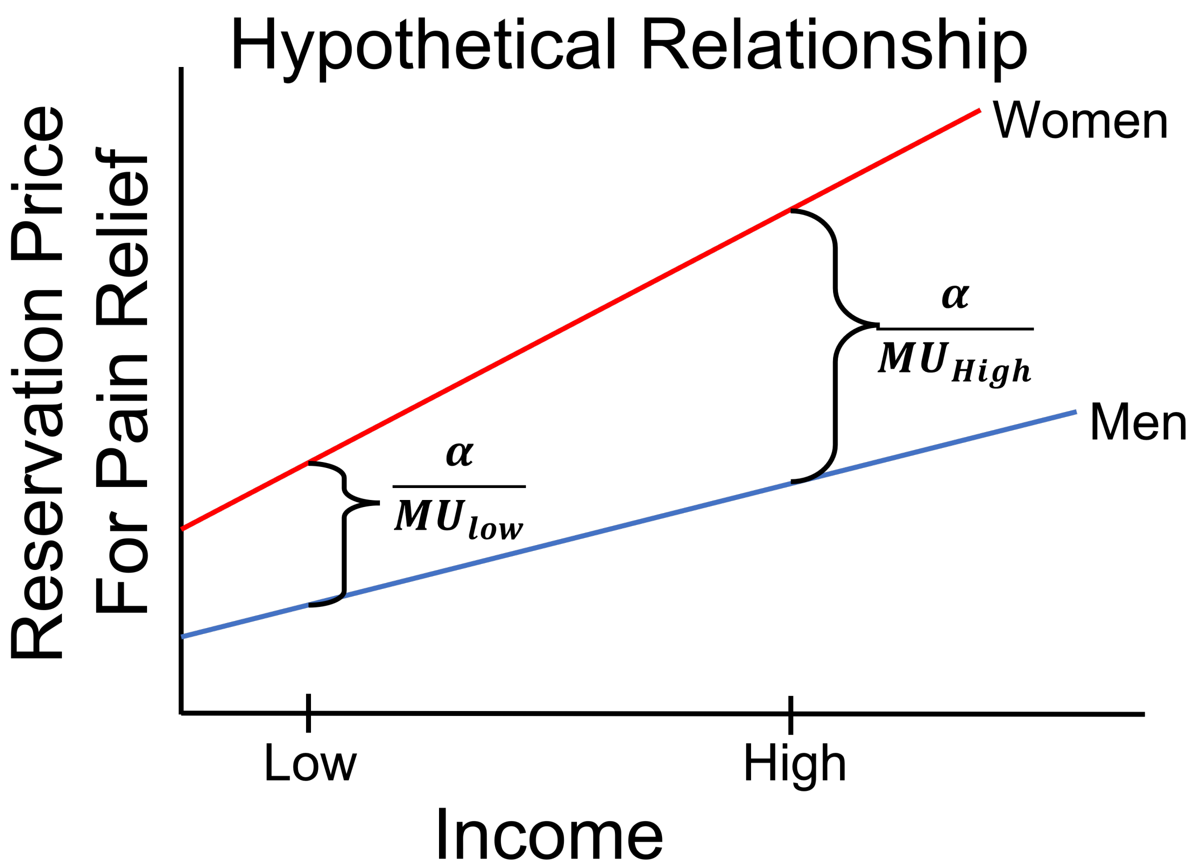

Figure 10

6.2 Open Ended Response With Controls

To move beyond simple averages, we need to further parameterize our model. As with the means, I aggregate income groups into b income bins. I assume the average marginal utility of income function is

Assumption 3.

$$\phi_i'(m_i) = \sum_{k=1}^{b} \mathbb{1}_{ik}(m_i \in k)\phi_k'$$

This gives b average marginal utility of income parameters, ϕ_k'. In this case b is four for the four income groups.

Now for the utility impact of r_i. I assume X_i and enters the model with linear parameters in the following form

Assumption 4.

$$r(X_i, \epsilon_i) = \beta_1 + \boldsymbol{\beta} X_i + \epsilon_i \tag{6}$$

Where

$$\epsilon_i \sim \mathcal{N}(0,\, \sigma^2) \tag{7}$$

and

$$\epsilon_i \perp\!\!\!\perp m_i \tag{8}$$

Together, these assumptions give the following equation for the reservation price of pain relief

Definition 1. *Given assumption 3 and 4 the reservation price for an individual is*

$$P_i^r = \frac{\beta_1 + \boldsymbol{\beta} X_i + \epsilon_i}{\sum_{k=1}^{B} \mathbb{1}_{ik}\phi_k'} \tag{9}$$

In matrix form, the expected price vector $\mathbb{P}^r$ for the full population can be written as

$$\mathbb{P}^r = (\mathbb{X}_i \boldsymbol{\beta} + \epsilon_i) \oslash \mathbb{M}\phi' \tag{10}$$

Where $\mathbb{X}$ is the matrix of X_i traits influencing pain tolerance and a constant for the intercept of the numerator. In matrix form, let $\boldsymbol{\beta}$ be $[\beta_1 \quad \boldsymbol{\beta}]$ from equation 9. Let $\mathbb{M}$ be an n by b matrix indicating what income bin each person i is in. Let $\boldsymbol{\phi'}$ be a b by 1 matrix of marginal utility parameters. $\oslash$ is element wise matrix division or Hadamard division.

Appendix B.4 proves that the model is identified, but the basic intuition is that it is identified so long as demographic characteristics do not show up in both the numerator and the denominator. For example, differences in pain relief across age cannot be separately identified from differences in marginal utility of income across age. This is why the independence of income and the utility from pain relief is crucial.

6.2.1 Likelihood Functions

We can estimate the model with MLE. Definition 1 gives the following probability density function

Definition 2.

$$P(P_i^r|\mathbb{X}_i, m_i \in k) = \pi \frac{1}{\sqrt{2\pi \frac{\sigma^2}{\phi'_{k_i}}}} e^{-\frac{(P_i^r - \hat{P_i^r})^2}{2(\frac{\sigma}{\phi'_{k_i}})^2}} \tag{11}$$

and the following log likelihood for a population of size n

Definition 3.

$$L(\mathbb{X}_i, m_i) = -\frac{n}{2} log(2\pi) + \sum_i^n -log(\frac{\sigma}{\phi'_{k_i}}) - \frac{(P_i^r - \frac{1+\boldsymbol{\beta}\mathbb{X}_i}{\phi'_{k_i}})}{2(\frac{\sigma}{\phi'_{k_i}})^2} \tag{12}$$

Where ϕ'_{k_i} indicates the marginal utility parameter for the corresponding bin for m_i. Recall that in assumption 4 the variance of the error term, and by extensions the variance in the utility from pain relief, is constant across income. This assumption allows the ϕ'_k terms to be identified off of mean differences in reservation prices, differences in the variance, and differences in reservation price gaps between control groups. We can see this since the marginal utility of income enters the same terms as σ and the term with $1 + \boldsymbol{\beta}\mathbb{X}_i$. The intuition here is that the underlying mean, variance, and coefficients in utility from pain relief are constant across income. So, observed differences in the mean, variance, or coefficients for the reservation prices across incomes must be because of differences in the marginal utility of income.

6.2.2 MLE Results

I run this model with controls for age, gender, and mean PSQ. The results of the model for each open ended question are shown in figure 11. The full coefficient table is also including in appendix A table 7. The vertical axis in figure 11 are direct estimates of marginal utility ϕ_k and not WTP like in the unconditional means above. Here, a higher value indicates a higher marginal utility. All the questions are generally flat. For example, none of even the largest income groups are statistically different than 1. A nice sanity check is that the coefficients on "Mean PSQ", which is the respondent's average response to the 0-10 pain sensitivity questions, are positive and significant for every question. For example, a one point higher average is associated with a .25 increase in utility from pain relief for bumping one's shin.

With utility normalize to one this means a one point increase in PSQ is correlated with a \$.25 increase in the reservation price for relief from bumping one's shin. This indicates that people who are more sensitive to pain will pay more to avoid it and indicates people are actually considering the questions and implications and answering in a logically consistent way throughout. Table 8 in appendix A shows the coefficients for a model where the variance of the error in utility is not assumed to be uniform across income groups, reflecting the conclusion from theorem 3. The results are similar.

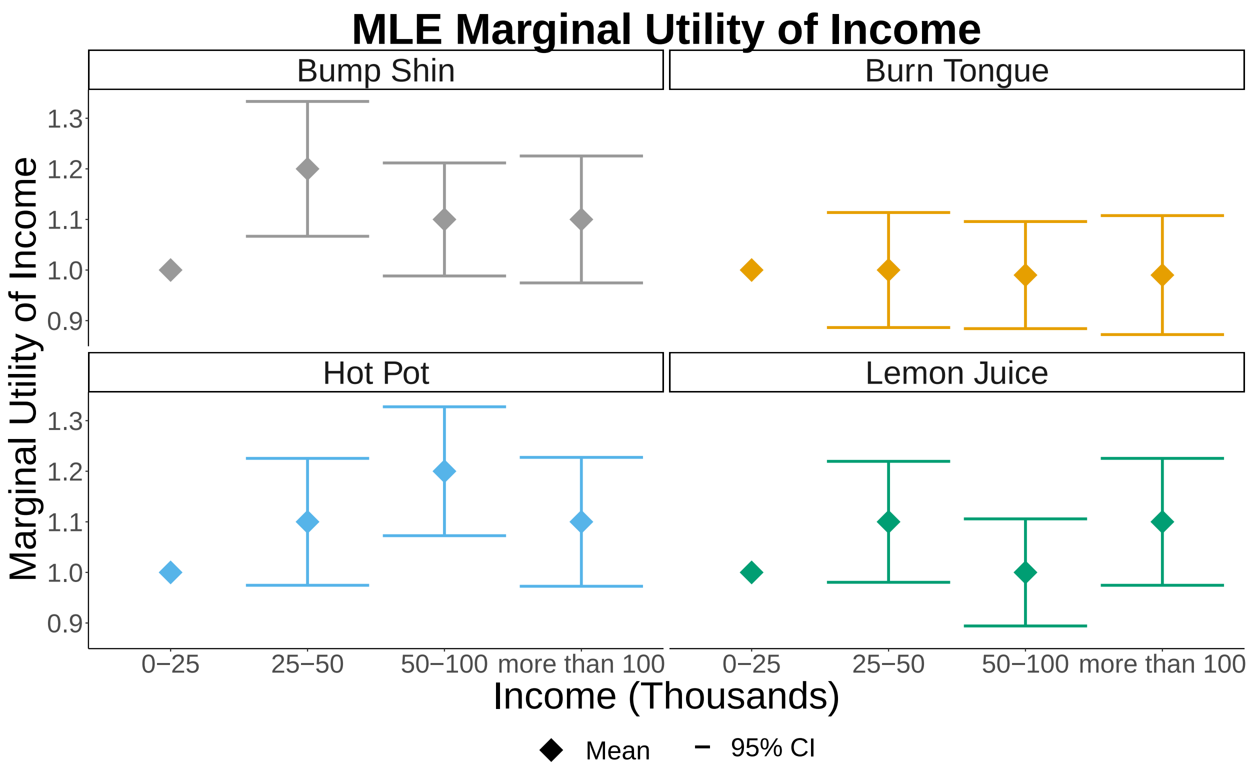

Figure 11

The contingent valuation literature suggests that open ended questions are more difficult to respond accurately to and have high rates of protest answers (Boyle, 2017, pg 110-111). With this in mind, I also ask a series of binary choice, yes or no, questions. The price proposed is randomized across respondents, but each respondent only answers each question one time, for one price. This is the opposite extreme to open ended questions in that it carries less information, but places much less burden on the respondents.

6.3 Binary Choice Model

While a binary choice question may be more familiar for respondents, leading to more accurate answers, assessing the results is more difficult and requires more assumptions. The underlying economic model is the same as in theorem 1 and 2, but now we must model and estimate the reservation price rather than just observing it in the data. I follow the general strategy from Hanemann (1984), but the technique outlined in that paper is to calculate an overall average or median WTP. I update the technique to analyze the change in WTP with respect to income rather than a single collective estimate. The first step is to run a random utility logit regression of the following form.

$$V_i = \sum_{j=1}^{4} \delta_j \mathbb{1}(M_i = j) + \gamma X_i + \sum_{j=1}^{4} \beta_j \mathbb{1}(M_i = j) * P_i + \epsilon_i \tag{13}$$

Where δ_j is an intercept coefficient for income level j, γ is a vector of coefficients for controls X_i and β_j is the price coefficient for income group j. It might seem incorrect, at first pass, to include unique intercepts for income in the utility model. This seems to imply different utility levels across incomes (the opposite of the identifying assumption). However, it is important to remember that V_i is ordinal utility that should not be compared across individuals and is, in my opinion, better thought of as just modeling choice probability. The differing utility intercepts allow for different choice probability levels across incomes while the differing price coefficient allows the choice probability to decrease deferentially with price across incomes. The intercepts allow the model to more flexibly estimate WTP with less strict functional form assumptions. A nice clarifying example is to imagine a good where people of different incomes buy it with the same probability at a price near zero (indicating a similar choice probability intercept). This observation would not in any way indicate that the good provides the same marginal benefit to both people. Similarly, the converse, having different purchase probability intercepts, does not imply different marginal benefits.

Putting our above equation in terms of Hanemann (1984), let α_1 and α_0 be the utility from pain relief and no pain relief respectively. We can write the difference in utility from paying for pain relief as

$$\Delta V = (\alpha_1 - \alpha_0) - \beta_j P \tag{14}$$

The CDF for the change in utility $F_n(\Delta V) = (1 + e^{-\Delta V})^{-1}$ gives the probability of purchasing relief. By extension the CDF of the reservation price is $F_n(\Delta V) = G_{P_r}(P)$. Now the mean reservation price, $\bar{P_r^j}$, for an income group j is represented by

$$\bar{P_r^j} = \int_0^\infty [1 - G_{P_r}(P)]dP \tag{15}$$

Connecting this back to the actual logit model, we can write this as

$$\bar{P_r^j} = \int_0^\infty [1 - \frac{1}{1 + e^{\delta_j + \bar{X}\gamma + \beta_j P}}]dP \tag{16}$$

Where $\bar{X}$ is the average value of the controls, age and gender and mean PSQ in this case, for the entire population. This assigns each group the ordinal utility for the average age and gender, but varies the ordinal utility associated with the income group and then scales the ordinal utility by the income specific ordinal marginal utility of income. The difference across groups shows the difference in WTP attributable to changes in income holding the controls fixed at the global means.

A slight variation of this I also use, which is used in Bishop and Heberlein (1979) and mentioned in Hanemann (1984), is to cap the integral at the maximum price.

6.3.1 Binary choice Results

A nice first check for the binary choice questions is to confirm that people are price sensitive. Do fewer people agree to pay for relief as the price increases? Figure 12 has log price on the horizontal axis and the percent of respondents who said they would pay that price on the vertical axis. This percent falls in all questions for all income categories. This indicates people did engage with the question and consider the price; however, a relationship between income groups is not clear.

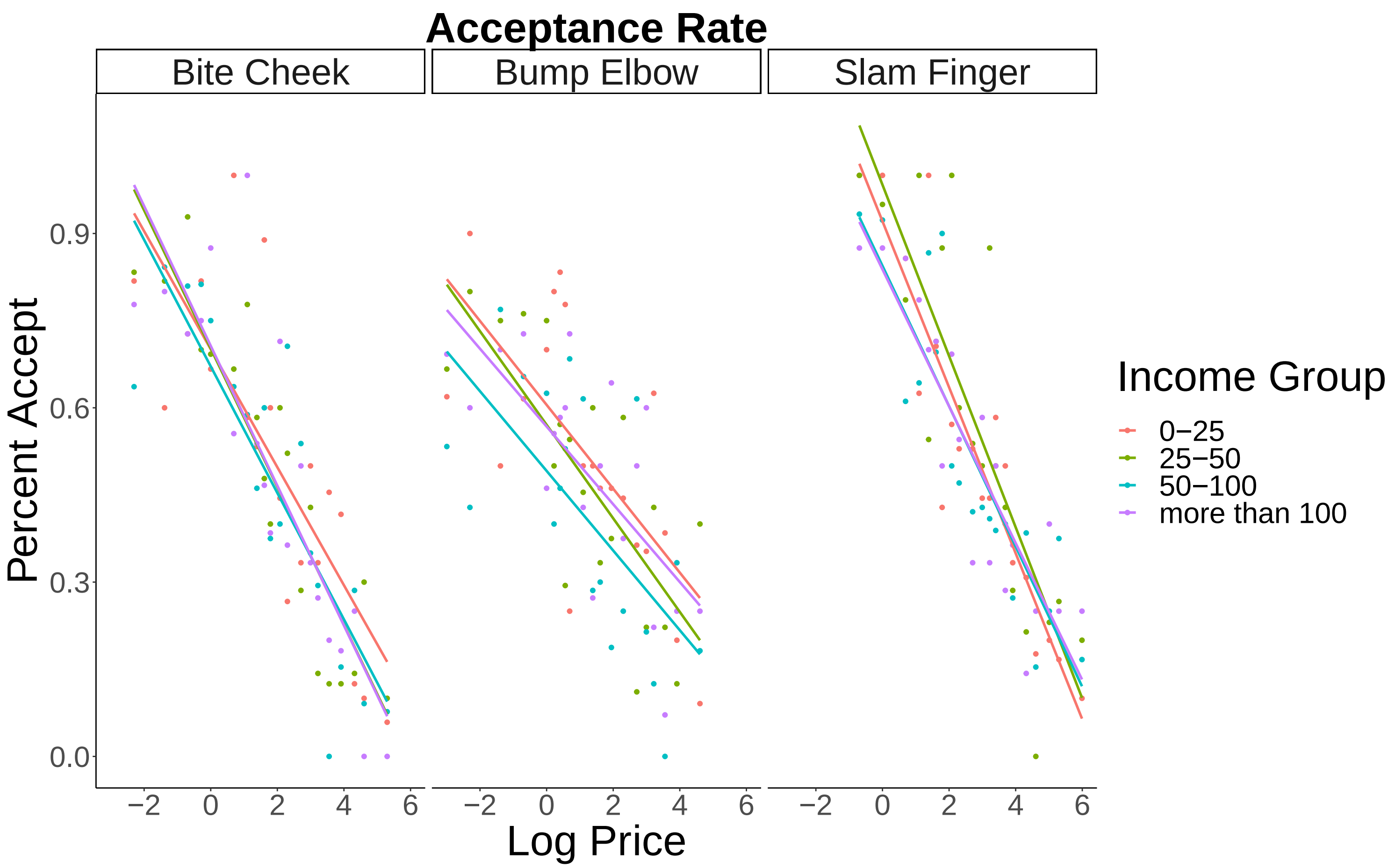

Figure 12

The willingness to pay estimates with controls for age, gender, and the mean response to the 0-10 PSQ questions and using the truncated mean (where the integral is capped at the highest bid) are displayed in figure 13. These estimates are normalized in the same way as the open ended means are. The first two questions seem to support the same story as the open-ended responses while the last suggest an increase in WTP, and so decrease in marginal utility. Between the lowest and highest income groups the WTP roughly doubles for the question regarding slamming your finger in a drawer, indicating marginal utility of income is halved. Table 9 in appendix A shows the results for the standard mean, truncated mean, and median willingness to pay estimates. The mean and truncated mean are what are described above, but the median estimate is described in Hanemann (1984).

For the binary choice questions, the error assumption in theorem 2 might not hold as well as in the free response questions. Appendix B.5 presents a model to estimate $\mathbb{E}\left[\frac{1}{P_i^r(m_i)}|m\right]$ in line with theorem 3, and the results can be seen in appendix A figure 14. Using this method, all three questions indicate little change over income.

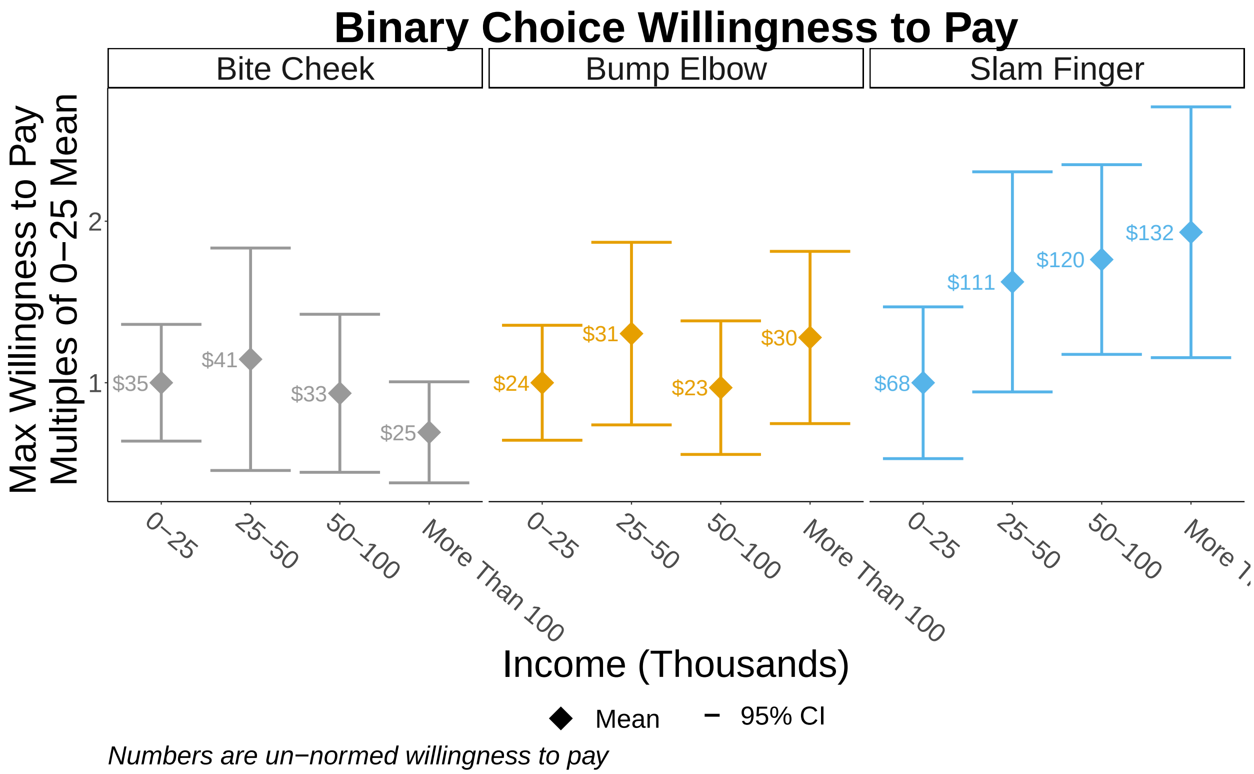

Figure 13

7 Discussion of Results

While there is some indication of diminishing marginal utility in one of the binary choice questions, the broad implication of the results is that marginal utility of income is constant across income groups, or at least diminishes slower than expected, implying that cardinal utility is roughly linear in dollars. This is a surprising outcome and certainly does not match my hypothesis going into this project.

To do be more precise, the results show that cardinal decision utility, which dictates individual choices, is linear. How does this relate to policymaker's choices and welfarist decision making? There are two possibilities. The first is to assume, as is standard in economics, that the utility policymakers value, typically assumed to correspond to well-being in some way, and the utility individuals maximize when making choices, decision utility, are one in the same. This is the conclusion most in line with standard economic assumptions. There is another possibility. Suppose a policymaker says that they are not a preference utilitarian[12]. They define utility to be some specific substantive thing, for example well-being. Furthermore, their definition of well-being *requires* that $100 creates more well-being

[12] They do not define utility as whatever people desire

when given to a poor person rather than a rich one. As an economist, if that is how they *define* well-being, I have no reason to persuade them otherwise. However, the results show that under this policymaker's definition of well-being, individuals do not maximize their own well-being[13]. This is a more radical departure from current practice in normative welfare economics, but it has precedence in even the economic literature. I discuss each option and their implications in turn below.

First, let's stick to the standard economic assumption that the utility policymaker's value is the same utility dictating individual preferences. The following welfare function will help to understand the possible implications of a constant marginal utility of income.

$$\sum_i W(m_i) = \sum_i \gamma(U(m_i)) \tag{17}$$

In this equation, $W(m_i)$ is the total weight the social planner places on a person with income m_i. $U(m_i)$ is the utility at income level m_i and γ is a function that expresses the policymaker's preferences over the distribution of utility. Using these equations, a pure utilitarian would have $\gamma(U(m_i)) = U(m_i)$ and value only the sum total of utility. A welfarist, but one who is not strictly utilitarian, might have a preference over not just the sum total utility, but how that utility is distributed. The policymaker's preference for the distribution of utility is represented with the γ function. Suppose I agree bumping one's shin on a table costs the same amount of utility regardless of income, but I would rather a richer person endure this pain than a poorer person since the richer person is at a higher total utility level. In this case $\gamma(U)$ will be concave in utility. I will refer to this type of policymaker as an egalitarian.

The results in this paper provide information about the shape of the $U(m_i)$ function, implying $U(m_i) = \alpha m_i + C$ where α and C are constant normalizations. What this implies about policymakers really depends on what it is we as economists know about those policymakers. Do we know more about W or γ?

Suppose we know more about γ. That is, we know a lot about the policymaker's preferences over the distribution of utility, but not much about their preferences over the end distribution of money. In this case, the knowledge that utility is linear rather than concave, i.e. $U(m_i) = \alpha m_i + C$, will update our beliefs about the policymaker's total welfare function $W(m_i)$. A pure utilitarian who only values the sum-total of utility, i.e. $\gamma(U(m_i)) = U(m_i)$, has a linear social welfare function. For an egalitarian who cares about the distribution of total utility the impact on the concavity of W is ambiguous. On the one hand, a linear

[13] Alternatively, this policymaker could reject assumption 1. This would mean the policymaker thinks that someone bumping their shin on a table lowers their well-being much less if they are rich than if they are poor. Even in this case we have learned more about how this policymaker thinks about well-being.

utility means that a marginal dollar to a rich person creates as much utility as a marginal dollar to a poor one. Compared to a concave utility, this would push our egalitarian towards a less concave W. However, the impact on the *distribution* of *total* utility is unclear. If the slope of the linear utility is large relative to the policymaker's priors, total utility may be much more unequal than under our concave prior assumption. This would push W towards being more concave. How exactly one should think about the level changes in utility across income is not addressed in this paper.

While knowing more about γ may make sense in some theoretical settings, it may not fit practical applications very well. I am not aware of work that elicits views about preferences over the distribution of utility or asks if policymakers really subscribe to a preference based welfarist or utilitarian ideology in accordance with assumption 2. What economic work has done is ask how much redistribution people would like and inferred redistributive goals from existing policy (Hendren, 2020; Kimball, Ohtake, Reck, Tsutsui, & Zhang, 2015). What this shows is that most people think giving $1 to a poor person is better than giving it to a rich person. People do use terms like "it would make a bigger difference to the poor person", but it's not actually obvious what they mean by that (Kimball et al., 2015). Do they mean the poor person desires it more in the utilitarian sense, or do they mean it in some other way? Perhaps they see the things it is being spent on as more worthy or more important beyond simple desire. In this case, it makes sense to think of policymaker's preference for redistribution, W, as fixed and update our beliefs about $\gamma(U)$. This informs us about how their motivations fit into the welfarist framework built on economic assumptions rather than informing us about their desired level of monetary redistribution. This approach indicates that the $\gamma(U)$ is more concave than typically assumed. In other words, the welfarist, but non-utilitarian, portion of people's re-distributive preferences, i.e. γ, is more concave than expected. This also implies that very few people subscribe to a pure utilitarian view (under the economic definition of utility) because very few people favor no redistribution. The typical policymaker will need to be modeled as an egalitarian.

Is a constant marginal utility of income plausible? The idea bumps up against some common intuition but does not cause as much conflict as I first expected. First, constant marginal utility of income says nothing about diminishing marginal utility for particular goods. For example, consuming a tenth apple is likely less enjoyable than the first. Part of the reason utility diminishes for apples is the assumption that people have convex preferences and so would prefer to spread their money out on other goods rather than concentrate it on only apples. This idea does not apply to money since doubling money would allow a consumer to double everything in an optimal way in accordance with their convex preferences. Convex preferences is one reason the tenth apple may be less enjoyable than the first, but it's also possible that goods actually do become less desirable on the whole the more a person has. In this case, doubling my income and doubling everything I own would lead to less than double utility. This conclusion has intuitive appeal, but in the real world, if someone's income were to double, they almost certainly would not just double their consumption of the same goods they already purchase. People do not have homothetic preferences and so the quality and type of goods will change significantly as their income changes. Both of these points show how little we really know about the marginal utility of income and why a linear marginal utility of income is plausible.

Non-homothetic preferences play another key role in why linear utility is plausible. The average marginal utility of income depends on the price vector in the economy. Thinking through how a price *change* could impact the marginal utility of income illustrates the role prices play. Suppose tomorrow all the goods lower income folks purchase on the margin, like necessities, become more expensive and lower quality. Suppose at the same time there are big efficiency gains to expensive luxury item purchased on the margin by richer folks (things low income folks could never afford). After these changes, a dollar might just buy a lot more value in the hands of a wealthy person than it did in the previous period[14]. This idea is also exacerbated by the related idea that it is expensive to be poor. Berkouwer and Dean, for example, show that households in Nairobi are only willing to pay \$12 for a stove that would save \$237 over two years, and that a low interest loan increases willingness to pay to the actual savings over the life of the loan (Berkouwer & Dean, 2021). Lower income folks may end up buying lower quality goods that do not last as long or provide as much value and so receive a lower quantity of quality adjusted goods per dollar on the margin.

I expect most people's prior intuition, including my own, is that utility must be concave, but our intuition about what utility is under assumption 2, that preferences reveal well-being, may simply not be very good. Under this assumption, the marginal utility of income is really about people's attitudes towards spending. How willing is a given consumer to part with money for a given gain? How flippant are they with cash? How much do they desire money relative to particular goods? It appears that richer people do still desire money. This desire just might not stack up well with policymaker's ethical beliefs about who is deserving of this money or where it will do the most "good". As explained above, we can adjust the $\gamma(U)$ in our welfare functions to match policymaker perceptions of fairness even with a linear utility function 2. This is possible so long as individuals are still maximizing their own well-being, but what if desire and individual actions are not even intrapersonally consistent with what a policymaker considers "good" or welfare maximizing? This cannot be reconciled by

[14]Similarly to the welfare function above, what this means for a policymaker's desire to redistribute between the two groups depends on if they value how utility is distributed in addition to its societal sum

manipulating $\gamma(U)$.

To make this idea more clear, consider two simple but conflicting definitions of utility, pleasure and happiness. Suppose individuals make choices to maximize their pleasure. So individual decision utility, and the utility corresponding to our utilitarian theory of action, is pleasure. At the same time, suppose a policymaker values the sum total of happiness. For this policymaker's utilitarian theory of ethics, which dictates what policy is the best and morally good policy, utility is happiness. This could cause a conflict for economists trying to use one utility function to represent both actions and policy assessment. How would we know if we were in this situation? How would we know if policymakers value something different than individuals? If this policymaker also knows that \$100 creates more happiness, under their definition, when given to a rich person than a poor person, then their definition conflicts with my estimates. This shows that their definition of happiness does not correspond to individual decision utility. While thinking in terms of somewhat concrete ideas like pleasure and happiness makes the distinction more apparent, any two different definitions of utility could lead to the same result. In these cases, the policymaker's definition of utility is "wrong" in the sense that it does not correspond to the way economists think about utility, through preferences and individual action, but from a moral or philosophical perspective they cannot really be "wrong". I have no desire to dictate their values, but the results show their values do not align to individual decision utility.

What would it mean for there to be two different utilities. One for individual decisions and one for ethics and policy-making? Amartya Sen actually found this idea appealing, saying

> A person is given one preference ordering, and as and when the need arises this is supposed to reflect his interests, represent his welfare, summarize his idea of what should be done, and describe his actual choices and behavior (Sen, 1977).

A single utility function to describe what is, ultimately, multiple different things may just be too simple. Different definitions of utility are one way this could come about. Another is that people may be making mistakes. The field of behavioral economics draws a distinction between what would maximize utility and what people actually do given their psychological constraints. Rather than decisions simply being too hard, people may also deviate consciously from what maximizes their own well-being. Sen (1977) draws this distinction with the ideas of sympathy and commitment. Recall that sympathy is when an individual does a seemingly selfless act because it would make them feel badly not to or they feel good for doing it (like a warm glow). Commitment is when someone chooses an action despite it actually making them worse off than the alternatives.

Prior to the results in this paper, I would have been more confident that behavioral mistakes or acts of commitment, while certainly important in specific instances, do not frequently impact regular, everyday purchasing decisions. However, one interpretation of the linear utility result is that these occurrences are quite common. Consider the following possibility. Suppose people will only buy a good if it both provides more well-being than its price *and* if the price is seen as fair. Kahneman, Knetsch, and Thaler (1991) document what people consider fair pricing to be, and the results do not clearly relate to individual well-being. If the fair price limit binds prior to the welfare limit, the practice of deriving well-being from the maximum willingness to pay, or decision utility generally, will be hindered. This example could be viewed as either a commitment to fair prices, or a behavioral mistake that people are making. Either way, people are not doing what is in their best interest. The richer people in my survey might be better off if they paid more for relief, but a commitment not to pay what they view as an unfair price might hold them back. Or a behavioral mistake, only purchasing fair price, might keep them from a welfare enhancing purchase at a higher price.

A third possibility that I have not yet touched on is that the results are biased. I do not see this as the most likely outcome, but I do want to point out weakness in the analysis that I hope to improve on in future work.

The biggest area of concern is selection bias into the survey. What kind of rich person does a survey for a few dollars? One who is willing to do things for little money. This is a characteristic that will lead them to systematically give lower answers than their average peer. This motivated me to originally go with the company Pollfish. They collect responses from phone app users and pay the users with "in app" benefits related to the apps purpose like a free yoga lesson or news article. My expectation is that a wealthy person is more likely to do a survey to get free lives while playing a game like candy crush, because it is viewed more as an ad experience, than to sign up to do surveys for money, which may be viewed more as a job. This is speculation on my part, and the sample in the pilot I did with Pollfish was not representative of the United States on observables. The non-representative sample is what convinced me to switch to Centiment. However, the results from the Pollfish survey show more of an increase marginal utility of income with income. The online appendix shows the mean figures and MLE results for the questions in my pilot. For the pain questions, the marginal utility for the 0-25 group was high, which I expected was due to non-working spouses or temporary unemployment making up a disproportionate share of that group. However, after this initial income group, we see an increase in WTP, implying a decrease in marginal utility of income. The sample is smaller and does not reflect national demographics, and so I do not place much weight on these. However, the difference across these two panels suggests another sampling approach would be a worthy endeavor (Another important note is that the data cleaning and analysis is not completely harmonized between the two samples. Look for an update on this in the near future!). Other ways to address sampling bias in the future would be to increase the reward for participation in the survey so that most people would be willing to do it, or to decrease the monetary reward to zero so the selection into the survey is based on altruism only and not a willingness to do tasks for little monetary reward.

Another general concern is that it is a hypothetical survey and not properly incentivized. This is partially alleviated because the analysis does not require the absolute value of reservation prices to be at all accurate. I just need it to be the case that the ratio across incomes is consistent with real purchasing activity. This issue could be solved with an incentivized experiment, which I hope to follow up with in the future.

It is also possible pain tolerance changes significantly across income. Beyond showing that the PSQ does not change across income, there is not much that is possible to say on the matter. The best way to address this concern would be with follow up studies using other goods that are conceivably uncorrelated with income, but not based on pain. In the Pollfish pilot I also asked questions about disgusting scenarios. The variance in these questions was much higher. My intuition is that there is much more variance in what people find disgusting and how much of it they can handle. This convinced me to focus on the pain questions in my final survey in order to make the most of my limited budget, but perhaps another look at things like disgust is warranted in future work.

8 Conclusion

Measuring utility is a difficult problem with a long history. Without any structure, measuring the concavity of utility with respect to income is not even a well-defined problem, but with the right structure and data it is empirically possible. I present a model for a feasible identification strategy for the average marginal utility of income and implement that strategy with new survey data. The results suggest that the marginal utility of income is constant across income groups, implying that utility is roughly linear in dollars. this result has significant implications for our understanding of individual preferences, utility, well-being, and distributional ethics.

References

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of political*

economy, *58*(4), 328–346.
Arrow, K. J. (1978). Extended sympathy and the possibility of social choice. *Philosophia*, *7*(2), 223–237.
Arrow, K. J. (2012). *Social choice and individual values* (Vol. 12). Yale university press.
Auspitz, R., & Lieben, R. (1889). *Untersuchungen über die theorie des preises.* Duncker & Humblot.
Aydede, M. (2019). Pain. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2019 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2019/entries/pain/.
Bartley, E. J., & Fillingim, R. B. (2013). Sex differences in pain: a brief review of clinical and experimental findings. *British journal of anaesthesia*, *111*(1), 52–58.
Berkouwer, S. B., & Dean, J. T. (2021). Credit, attention, and externalities in the adoption of energy efficient technologies by low-income households. *Unpublished manuscript, University of Pennsylvania, Philadelphia, PA.*
Bishop, R. C., & Heberlein, T. A. (1979). Measuring values of extramarket goods: Are indirect measures biased? *American journal of agricultural economics*, *61*(5), 926–930.
Boyle, K. J. (2017). Contingent valuation in practice. In *A primer on nonmarket valuation* (pp. 83–131). Springer.
Coate, S. (2000). An efficiency approach to the evaluation of policy changes. *The Economic Journal*, *110*(463), 437–455.
Dimand, R. (2019). Irving fisher, ragnar frisch and the elusive quest for measurable utility. Retrieved from https://ssrn.com/abstract=3495946
Driver, J. (2014, Sep). *The history of utilitarianism.* Stanford University. Retrieved from https://plato.stanford.edu/entries/utilitarianism-history/
Eastmond, T., Mather, N., Ricks, M. D., & Betts, J. (2022). From value added to welfare added: A social planner approach to education policy and statistics.
Finkelstein, A., & Hendren, N. (2020, November). Welfare analysis meets causal inference. *Journal of Economic Perspectives*, *34*(4), 146-67. Retrieved from https://www.aeaweb.org/articles?id=10.1257/jep.34.4.146 doi: 10.1257/jep.34.4.146
Fisher, I. (1927). *A statistical method for measuring" marginal utility" and testing the justice of a progressive income tax.*
Flood, A. (n.d.). Terry pratchett estate backs jack monroe's idea for 'vimes boots' poverty index. *The Guardian.* Retrieved 6-27-2022, from https://www.theguardian.com/books/2022/jan/26/terry-pratchett-jack-monroe-vimes-boots-poverty-index

Glick, M. (2018). The unsound theory behind the consumer (and total) welfare goal in antitrust. *The Antitrust Bulletin*, *63*(4), 455–493.

Hanemann, W. M. (1984). Welfare evaluations in contingent valuation experiments with discrete responses. *American journal of agricultural economics*, *66*(3), 332–341.

Hare, R. M. (1981). *Moral thinking: Its levels, method, and point.* Oxford: Clarendon Press; New York: Oxford University Press.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of political economy*, *63*(4), 309–321.

Harsanyi, J. C. (1986). *Rational behaviour and bargaining equilibrium in games and social situations*. CUP Archive.

Hausman, D. M. (1995). The impossibility of interpersonal utility comparisons. *Mind*, *104*(415), 473–490.

Hausman, D. M., & McPherson, M. S. (2006). *Economic analysis, moral philosophy and public policy.* (Vol. 2nd ed). Cambridge University Press. Retrieved from `http://proxy.lib.umich.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=149844&site=ehost-live&scope=site`

Hendren, N. (2020). Measuring economic efficiency using inverse-optimum weights. *Journal of Public Economics*, *187*, 104198. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0047272720300621` doi: https://doi.org/10.1016/j.jpubeco.2020.104198

Hendren, N., & Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. , *135*(3), 1209-1318. Retrieved from `https://academic.oup.com/qje/artjournal={QuarterlyJournalofEconomics}` doi: 10.1257/jep.34.4.146

Hicks, J. R. (1940). The valuation of the social income. *Economica*, *7*(26), 105–124. Retrieved from `http://www.jstor.org/stable/2548691`

Jevons, W. S. ([1871] 1911). *The theory of political economy, 4th ed. london: Macmillan.*

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, *5*(1), 193–206.

Kaldor, N. (1939). Welfare propositions of economics and interpersonal comparisons of utility. *The Economic Journal*, *49*(195), 549–552. Retrieved from `http://www.jstor.org/stable/2224835`

Kimball, M. S., Ohtake, F., Reck, D., Tsutsui, Y., & Zhang, F. (2015). Diminishing marginal utility revisited. *Available at SSRN 2592935*.

Lautenbacher, S., Peters, J. H., Heesen, M., Scheel, J., & Kunz, M. (2017). Age changes in pain perception: a systematic-review and meta-analysis of age effects on pain and tolerance thresholds. *Neuroscience & Biobehavioral Reviews*, *75*, 104–113.

MacKay, A. F. (1986). Extended sympathy and interpersonal utility comparisons. *The Journal of philosophy*, *83*(6), 305–322.

Miljković, A., Stipčić, A., Braš, M., orević, V., Brajković, L., Hayward, C., ... Polašek, O. (2014). Is experimentally induced pain associated with socioeconomic status? do poor people hurt more? *Medical science monitor: international medical journal of experimental and clinical research*, *20*, 1232.

Morgan, J. N. (1945). Can we measure the marginal utility of money? *Econometrica: Journal of the Econometric Society*, 129–152.

Moscati, I. (2018). *Measuring utility: From the marginal revolution to behavioral economics.* Oxford Studies in History of E.

Pearce, D. (2021, Mar). Individual and social welfare: A bayesian perspectiveâ.

Ruscheweyh, R., Marziniak, M., Stumpenhorst, F., Reinholz, J., & Knecht, S. (2009). Pain sensitivity can be assessed by self-rating: Development and validation of the pain sensitivity questionnaire. *Pain*, *146*(1-2), 65–74.

Ruscheweyh, R., Verneuer, B., Dany, K., Marziniak, M., Wolowski, A., Çolak-Ekici, R., ... others (2012). Validation of the pain sensitivity questionnaire in chronic pain patients. *Pain*, *153*(6), 1210–1218.

Samuelson, P. A. (1950). Evaluation of real national income. *Oxford Economic Papers*, *2*(1), 1–29.

Samuelson, P. A., & Swamy, S. (1974). Invariant economic index numbers and canonical duality: survey and synthesis. *The American Economic Review*, *64*(4), 566–593.

Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy & Public Affairs*, 317–344.

Stephens, R., & Robertson, O. (2020). Swearing as a response to pain: Assessing hypoalgesic effects of novel "swear" words. *Frontiers in Psychology*, *11*. Retrieved from `https://www.frontiersin.org/article/10.3389/fpsyg.2020.00723` doi: 10.3389/fpsyg.2020.00723

Stevenson, B., & Wolfers, J. (2019). *Principles of macroeconomics.* Macmillan Higher Education.

Wilson, C. S. (2019). Welfare standards underlying antitrust enforcement: What you measure is what you get. In *In: Luncheon keynote address at george mason law review 22nd antitrust symposium. arlington: Antitrust at the crossroads.*

Zwolinski, M. (2008). The ethics of price gouging. *Business Ethics Quarterly*, *18*(3), 347–378.

Appendix A Auxiliary Tables and Figures

Average Willingess to Pay

	MEAN	SE	SD	MEAN NORMED	SD NORMED	N
LEMON JUICE						
0-25	5.2	0.48	7.8	1.0	1.0	242
25-50	4.6	0.49	7.3	0.87	0.94	239
50-100	4.9	0.42	7.8	0.95	1.0	334
MORE THAN 100	4.6	0.49	7.5	0.88	0.97	206
HOT POT						
0-25	40	3.6	59	1.0	1.0	242
25-50	35	3.3	53	0.88	0.89	239
50-100	30	2.7	47	0.77	0.80	334
MORE THAN 100	35	3.7	54	0.89	0.92	206
BURN TONGUE						
0-25	9.7	0.97	16	1.0	1.0	242
25-50	9.4	1.0	15	0.97	0.98	239
50-100	9.6	0.83	16	0.99	1.0	334
MORE THAN 100	10	1.1	16	1.0	1.0	206
BUMP SHIN						
0-25	23	2.2	34	1.0	1.0	242
25-50	18	1.8	28	0.78	0.82	239
50-100	19	1.8	31	0.82	0.93	334
MORE THAN 100	19	2.1	32	0.83	0.93	206

Table 5: Mean Open Response Table

Average Inverse Willingess to Pay

	MEAN NORMED	SE	SD
LEMON JUICE			
0-25	1.00	0.04	0.71
25-50	1.09	0.05	0.70
50-100	1.10	0.04	0.72
MORE THAN 100	1.12	0.05	0.71
HOT POT			
0-25	1.00	0.09	1.44
25-50	0.95	0.09	1.38
50-100	1.03	0.07	1.41
MORE THAN 100	0.94	0.09	1.32
BURN TONGUE			
0-25	1.00	0.05	0.84
25-50	0.99	0.05	0.83
50-100	1.01	0.05	0.83
MORE THAN 100	0.97	0.06	0.83
BUMP SHIN			
0-25	1.00	0.07	1.15
25-50	1.00	0.07	1.09
50-100	1.04	0.06	1.11
MORE THAN 100	1.10	0.08	1.14

*WTP < 1 was rounded to 1 to avoid dividing by zero Inverse WTP was normalized so lowest bin is 1

Table 6: Mean Open Response Table

Full MLE Coefficient Table

PARAMETER	LEMON JUICE		HOT POT		BURN TONGUE		BUMP SHIN	
	EST	SE	EST	SE	EST	SE	EST	SE
Income 0-25k	**1.0**	NA	**1.0**	NA	**1.0**	NA	**1.0**	NA
Income 25-50k	**1.1**	0.061	**1.1**	0.064	**1.0**	0.058	**1.2**	0.068
Income 50-100k	**1.0**	0.054	**1.2**	0.065	**0.99**	0.054	**1.1**	0.057
Income More Than 100k	**1.1**	0.064	**1.1**	0.065	**0.99**	0.060	**1.1**	0.064
intercept	0.61	0.21	0.64	0.20	0.051	0.22	-0.10	0.19
sigma	1.5	0.061	1.4	0.060	1.5	0.065	1.4	0.057
Mean psq	0.13	0.027	0.091	0.027	0.20	0.029	0.25	0.027
Age	-0.0090	0.0025	-0.00059	0.0024	-0.0041	0.0026	-0.0069	0.0023
Male	0.084	0.094	-0.27	0.092	0.11	0.099	0.040	0.088

Table 7: Full MLE Coefficient Table

Full MLE Coefficient Table With Unique Variance Terms

PARAMETER	LEMON JUICE EST	LEMON JUICE SE	HOT POT EST	HOT POT SE	BURN TONGUE EST	BURN TONGUE SE	BUMP SHIN EST	BUMP SHIN SE
Income 0-25k	**1.0**	NA	**1.0**	NA	**1.0**	NA	**1.0**	NA
Income 25-50k	**1.2**	0.15	**1.2**	0.16	**1.1**	0.15	**1.3**	0.17
Income 50-100k	**1.0**	0.14	**1.3**	0.15	**1.0**	0.12	**1.2**	0.13
Income More Than 100k	**1.1**	0.15	**1.1**	0.15	**0.99**	0.14	**1.1**	0.15
intercept	0.66	0.22	0.54	0.21	-0.29	0.22	-0.11	0.21
sigma 0-25k	1.5	0.068	1.5	0.067	1.6	0.074	1.4	0.062
sigma 25-50k	1.6	0.22	1.6	0.23	1.6	0.24	1.6	0.21
sigma 50-100k	1.5	0.21	1.5	0.19	1.6	0.20	1.5	0.18
sigma More Than 100k	1.4	0.22	1.4	0.22	1.5	0.23	1.5	0.21
Mean psq	0.14	0.029	0.11	0.028	0.24	0.034	0.26	0.031
Age	-0.0098	0.0027	0.00020	0.0026	-0.0020	0.0027	-0.0067	0.0025
Male	0.091	0.099	-0.26	0.097	0.14	0.10	0.019	0.094

Table 8: Full MLE Coefficient Table: Income specific Standard Deviation

Binary Choice Willingness to Pay

	BUMP ELBOW		BITE CHEEK		SLAM FINGER	
	EST	SE	EST	SE	EST	SE
MEAN WTP						
INCOME: 0-25K	**25**	(6)	**36**	(7)	**69**	(18)
INCOME: 25-50K	**37**	(21)	**41**	(15)	**115**	(31)
INCOME: 50-100K	**25**	(8)	**33**	(9)	**131**	(32)
INCOME: MORE THAN 100K	**34**	(13)	**25**	(6)	**147**	(52)
TRUNCATED MEAN WTP						
INCOME: 0-25K	**24**	(4)	**35**	(7)	**68**	(16)
INCOME: 25-50K	**31**	(7)	**41**	(12)	**111**	(24)
INCOME: 50-100K	**23**	(5)	**33**	(9)	**120**	(20)
INCOME: MORE THAN 100K	**30**	(6)	**25**	(6)	**132**	(27)
MEDIAN WTP						
INCOME: 0-25K	**10**	(4)	**16**	(6)	**39**	(11)
INCOME: 25-50K	**8**	(8)	**19**	(8)	**76**	(20)
INCOME: 50-100K	**3**	(4)	**20**	(6)	**70**	(19)
INCOME: MORE THAN 100K	**14**	(8)	**16**	(4)	**83**	(31)

*SE are bootstarp standard errors

Table 9

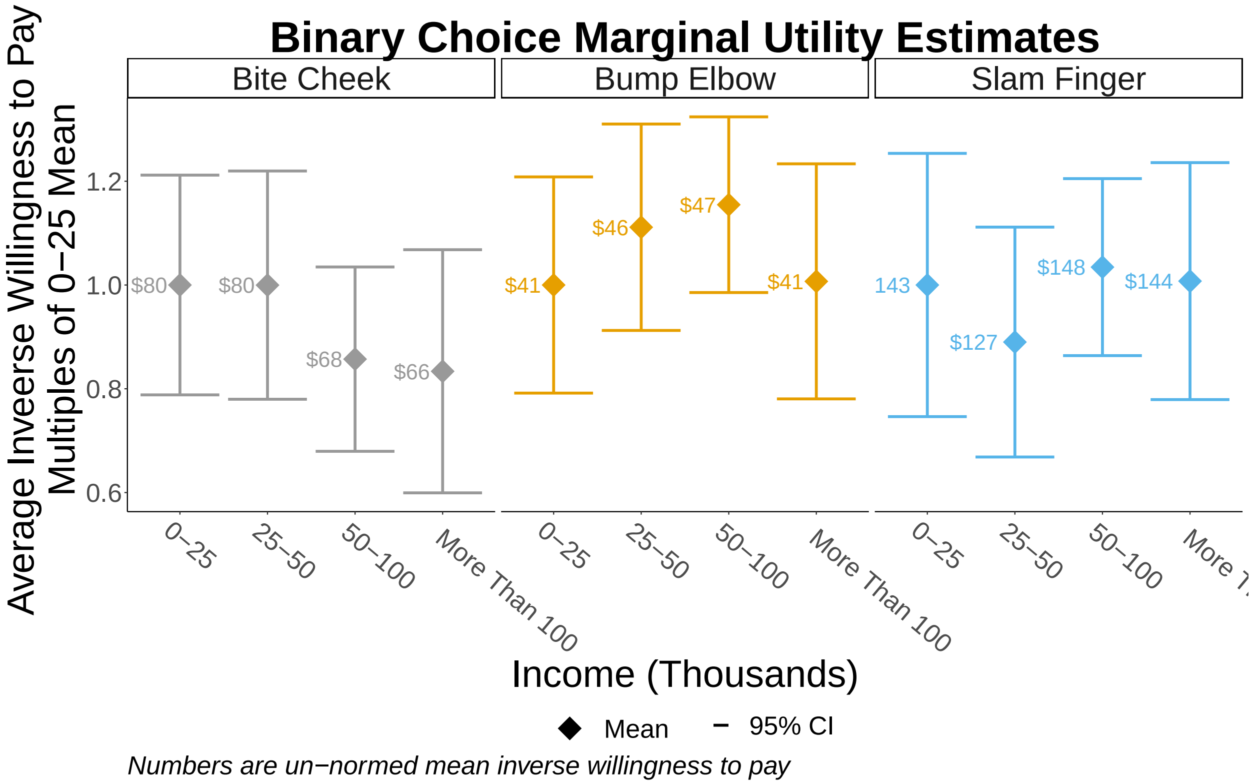

Figure 14

Appendix B Empirical Extensions and Clarifications

B.1 Empirical Model With Indirect Utility

The simple empirical model laid out in the body of the paper includes money as a numeraire good. However, the same estimation strategy can be supported using heterogeneous preferences and indirect utility. I think this model is more accurate to the real world where there are many goods and a whole set of prices, but the estimation is ultimately identical, and it is more confusing. Thus, it is here in the appendix.

First consider an indirect utility function V that is a function of prices P and income y, but also other characteristics θ like gender, age, pain sensitivity, or anything else that might influence preferences. This allows V to be heterogeneous across different people. This gives

$$V(P, Y, \theta) \tag{18}$$

Now for the pain relief described in our questions needs to be re-characterized as a price change. Let P be the price vector in our current world and let P' be a price vector where immediate pain relief is free. Now we can define the change in utility from pain relief as follows.

$$V(P', Y, \theta) = V(P, Y, \theta) = \Delta V^j \tag{19}$$

Since these are relatively small changes we can, as in the empirical model in the body, treat this as a marginal change. That means the following equality holds

$$EV(P', P, Y, \theta)U_y'(P, Y) = \Delta V^j \tag{20}$$

Where U_y' is the marginal utility of income and EV is equivalent variation. To shorten the notation, let $EV(P', P, Y, \theta) = EV^j$.

An important note regarding the previous step is that the marginal utility of income is not a function θ. So, while the preference for pain relief may vary by characteristics like age, the marginal utility of income cannot.

Now to identify the marginal utility of income we need the same assumption as in the simpler model. That is, we need the utility from pain relief to be independent of income. With that we get the following

Theorem 4. *if* $\Delta V^j \perp\!\!\!\perp Y$ *then*

$$\mathbb{E}[U_y'(P, Y)|Y, P] = \frac{\alpha}{\mathbb{E}[EV(P', P, Y, \theta)|Y, P]} \tag{21}$$

Proof. Given $\Delta V^j \perp\!\!\!\perp Y$ we get that

$$\mathbb{E}[\Delta V^j|P, Y] = \mathbb{E}[\Delta V^j|P] = \alpha \tag{22}$$

□

The first equation follows from the independence of the change in utility, ΔV^j and income Y. The second equality is a bit of an abuse of notation, but here the price vector is actually set to our current price vector P, and so this is a constant we can normalize to any level. Now, given this equality we get

$$
\begin{aligned}
&\alpha && (23)\\
&= \mathbb{E}[\Delta V^j | Y, P] && (24)\\
&= \mathbb{E}[EV(P', Y, P, \theta) U_y'(Y, P) | Y, P] && (25)\\
&= \mathbb{E}[EV(P', Y, P, \theta) | Y, P] E[U_y'(Y, P) | Y, P] && (26)\\
\mathbb{E}[U_y'(P, Y) | Y, P] &= \frac{\alpha}{\mathbb{E}[EV(P', P, Y, \theta) | Y, P]} && (27)
\end{aligned}
$$

The second line is the normalization explained above, the third line is the identify also explained above, the last line comes from the fact that $U_y'(Y, P)$ is only a function of Y and P and so conditioning on those makes it a constant. Rearranging the equation gives us the theorem.

One point this more complex model makes clearer is that the model estimates the marginal utility of nominal dollars, conditional on a price vector. A nominal dollar might not provide the same purchasing power to everyone in a world of non-linear pricing, quality variation, credit constraints, non-convex preferences, and other complications outside of basic economics models. In particular, a marginal dollar may have more, or less, purchasing power the more dollars someone has. A simple example would be to consider geographic sorting. Suppose richer people live in more expensive areas. In this case richer people value dollars less because the marginal value of consumption is lower, but also because a dollar literally buys less at stores in their area. This example might theoretically be controlled for with geographic price indices, but other examples are more complicated.

The following story popularized by novelist Terry Pratchett illustrating why the rich are able to spend less. Suppose a quality pair of boots that will last ten years is \$50, but a cheap pair that will last only a year is \$10. A poor person, with a marginal dollar, may only be able to purchase the cheaper option despite it being more costly in the long run (Flood, n.d.). The poor person may appreciate the boots more, in line with diminishing marginal utility of consumption, but the richer person is able to purchase boot years at half the price, making their marginal consumption per dollar higher. This example supports the colloquial saying[15], "it's expensive to be poor". Higher credit rates, an inability to buy in bulk or take advantage of off-peak sales might mean the poor just can't buy as much with an additional dollar. Berkouwer and Dean, for example, show that households in Nairobi are only willing to pay \$12 for a stove that would save \$237 over two years, and that a low interest loan increases willingness to pay to the actual savings over the life of the loan (Berkouwer &

[15] I've at least heard this a lot among family and friends. Not sure how common it actually is

Dean, 2021).

Perhaps a more fundamental consideration is that consumption is not homothetic and so the types of goods people consume with their first \$10,000 a year look very different than after making \$100,000. The first priorities for consumption are essentials like food and shelter. If food staples and housing become more expensive relative to luxury or entertainment goods, than we will see the marginal utility of a dollar for the poor fall relative to the rich. So, in general, the marginal utility of a nominal dollar also captures the relative cost of the differing consumption baskets of each income group. Interestingly, looking at these utility estimates over time could capture to what extent inflation has been concentrated on essentials or low-quality items compared to luxury goods.

B.2 Marginal Relief Assumption

While the empirical model in the body of the paper treats pain relief as a marginal change, in truth, the questions are a binary choice. Pay and receive total pain relief, or don't. How does this change the model? In words, I am making a local linearity assumption. I average the reservation prices within an income group and so a reservation price that is twice as high is treated as twice as much utility lost. If utility is concave, this is not correct since losing twice as much money should be more than twice as bad. The extent to which this biases the result is a function of the concavity of utility and the size of the difference. Over small changes, approximately marginal, the linearity assumption will be not so wrong. While our priors may be that someone with twice the income has very a different marginal utility of income, it is typically not assumed that say, a \$100 difference in income will drastically alter the marginal utility of income. To formalize this, we can re-characterize equation 1 as

$$U(m_i, q_i, X_i, \epsilon_i) = \phi(m_i) + r(X_i, \epsilon_i) \tag{28}$$

Where now r is a binary choice good for full relief or no relief. Now, the indifference condition is characterized by

$$\phi(m_i) = \phi(m_i - P_i^r(m_i)) + r(X_i, \epsilon_i) \tag{29}$$

Taking a first order Taylor approximation gives

$$\phi(m_i - P_i^r(m_i)) \approxeq \phi(m_i) - \phi'(m_i)P_i^r(m_i) \tag{30}$$

Inserting this into the indifference condition gives

$$\phi(m_i) \approxeq \phi(m_i) - \phi'(m_i)P_i^r(m_i) + r(X_i, \epsilon_i) \tag{31}$$

and finally rearranging gives us

$$P_i^r(m_i) \approxeq \frac{r(X_i, \epsilon_i)}{P_i^r(m_i)} \tag{32}$$

The Taylor approximation will be closer to correct the smaller the change in utility and the closer to linear utility is over the range from m_i to $m_i - P_i^r(m_i)$

An alternative way to see this is to consider the following exact equation

$$\frac{1}{P_i^r(m_i)} \int_{m_i - P_i^r(m_i)}^{m_i} \phi'(m_i) = \frac{r(X_i, \epsilon_i)}{P_i^r(m_i)} \tag{33}$$

Technically, what we are identifying with the inverse of the reservation price is the average marginal utility of income over the range from m_i to $m_i - P_i^r(m_i)$. Since the prices are small, this average is probably close to the marginal utility.

Suppose this really were a big concern. For example, suppose the reservation prices were larger and/or I was finding more concave utility. One potential solution would be to iteratively estimate the marginal utility of income function and, for each iteration, use the previous estimate to compute the integral in equation 33 until the estimates converge. Given my results indicate a linear utility, I have not done this or formalized the econometrics, but I expect it would provide more accurate estimates in this hypothetical case.

B.3 Proof of Theorem 3

Proof. Rearranging the indifference condition from equation 2 gives

$$\frac{\phi_i(m_i)}{r'(q_i, X_i, \epsilon_i)} = \frac{1}{P_i^r(m_i)}$$

Now taking the conditional expectation of both sides, we get

$$\begin{aligned}
&\mathbb{E}\Big[\frac{1}{P_i^r(m_i)}|m\Big] \\
&= \mathbb{E}\Big[\frac{\phi_i(m_i)}{r'(q_i, X_i, \epsilon_i)}|m\Big] \\
&= \mathbb{E}[\phi_i(m_i)|m]\mathbb{E}\Big[\frac{1}{r'(q_i, X_i, \epsilon_i)}|m\Big] \\
&= \mathbb{E}[\phi_i(m_i)|m]\mathbb{E}\Big[\frac{1}{r'(q_i, X_i, \epsilon_i)}\Big] \\
&= \mathbb{E}[\phi_i(m_i)|m]\frac{1}{\alpha} \\
&\implies \mathbb{E}[\phi(m_i)|m] = \alpha\mathbb{E}\Big[\frac{1}{P_i^r(m_i)}|m\Big]
\end{aligned}$$

where the third line comes from $r'(q_i, X_i, \epsilon_i) \perp\!\!\!\perp \phi_i(m_i)|m$, the fourth line comes from $r'(q_i, X_i, \epsilon_i) \perp\!\!\!\perp m_i$, and the fifth line comes from normalizing utility.

$\square$

B.4 Maximum Likelihood Identification

Equations 9 and 10 lead to the following theorem specifying the identification of the parameters. Let $\theta = (\boldsymbol{\beta}, \boldsymbol{\phi'})$ be the full set of parameters in the model.

Theorem 5. *If the conditions in assumption 3 and 4 and definition 1 hold and we also have that the matrix* $[\mathbb{X} \quad \mathbb{M}]$ *is full rank, then the ratio of any two parameters in* θ *is identified. If we normalize the marginal utility of income for the lowest income group to one, that is* $\phi'_1 = 1$*, than the remaining parameters in* θ *are identified.*

To prove this Suppose there exists a $\theta^* \neq \theta$ s.t. $\mathbb{E}[\mathbb{P}^r|\theta^*] = \mathbb{E}[\mathbb{P}^r|\theta]$. This implies

$$\mathbb{X}_i\boldsymbol{\beta}^* \oslash \mathbb{M}\phi'^* = \mathbb{X}_i\boldsymbol{\beta} \oslash \mathbb{M}\phi' \tag{34}$$

or that

$$\frac{\beta_1^* + X_i\boldsymbol{\beta}^*}{\sum_{k=1}^b \mathbb{1}_{ik}\phi_k'^*} = \frac{\beta_1 + X_i\boldsymbol{\beta}}{\sum_{k=1}^b \mathbb{1}_{ik}\phi_k'} \quad \forall \quad i \tag{35}$$

Now it is true that $\theta = \alpha\theta^*$, where α is any constant, satisfies the condition since α cancels out in the numerator and denominator. Once we have normalized $\phi_1 = 1$, however, α no longer appears in the denominator for i's in income group 1 and so does not cancel.

If $\phi_2' = .5$, for example, it implies the marginal utility of a dollar for income group 2 is half that of group 1.

With the $\phi_1 = 1$ normalization, any change to a parameter in the numerator would alter the expected reservation price for those in group one and violate the equality in equation 35, assuming, as in a regression, that $\mathbb{X}$ is full rank. Any change to the other marginal utility parameters ϕ_k' could be cancelled out for that group by appropriately scaling the numerator, but, since the marginal utility of group one is fixed and they share the numerator parameters, this would again change the expected reservation price for group 1 and violate 35. Thus, there does not exist a $\theta^* \neq \theta$ satisfying the condition.

Can this same identification strategy be used for any good? No, the marginal utility of income cannot be identified from just any reservation price. Recall the assumption that $\epsilon_i \perp\!\!\!\perp m_i$ and notice that income does not enter the utility function for pain relief and the factors impacting pain relief to not impact the marginal utility of income. If either of these appeared in both the numerator and denominator, we would not not be able to uniquely identify the parameters.

B.5 Binary Choice Expectation of Inverse Price

The average reservation price is found by using the integral of one minus the CDF for the reservation price $G_{P_r}(P)$ like so:

$$\bar{P_r^j} = \int_0^\infty [1 - G_{P_r}(P)]dP \tag{36}$$

To get an estimate of the average inverse of the reservation price in line with theorem 3 we need the CDF of $Y = \frac{1}{x}$. Using the following, we get

$$\begin{aligned}
&F_n(Y) = P(Y < p) \\
&= P(\frac{1}{x} < p) \\
&= P(\frac{1}{p} < x) \\
&= 1 - G_{P_r}(\frac{1}{p}) \\
&\implies \mathbb{E}[\frac{1}{p}] = \int_0^\infty [G_{P_r}(\frac{1}{p})]dp \\
&= \int_0^\infty [\frac{1}{1 + e^{\delta_j + \bar{X}\gamma + \beta_j \frac{1}{p}}}]
\end{aligned}$$

The results of this estimation are in figure **??**. In my Monte Carlo Simulations for this estimate, the truncated estimate, which limits the integral to the highest bid, was the only stable and accurate estimator for the average inverse reservation price. This is what is presented in the figure.