On the Value of Birth Weight*

Damian Clarke
Universidad de Santiago de Chile

Sonia Oreffice
University of Surrey & IZA

Climent Quintana-Domeque
University of Oxford & IZA

March 5, 2017

Abstract

A large body of evidence documents the educational and labor market returns to birth weight, which are reflected in investments in large social safety net programs targeting birth weight and early life health. However, there is no direct evidence on the *private* valuation of birth weight. In this paper we estimate the willingness to pay (WTP) for birth weight in the US. Using a series of discrete choice experiments, we find that individuals are, on average, willing to pay $1.44 for each additional gram of birth weight. This marginal WTP is particularly high at low birth weights, and turns negative at higher weights. The WTP among parents is higher than among non-parents, and particularly than those who do not plan to have children. Nonetheless, a series of calculations suggest that even the parental WTP for birth weight falls short of the inferred public WTP from large social safety net programs, and is lower than the expected present value of birth weight in the labor market for a US-born child. We present a parsimonious model which is able to explain the different WTP by parental status and the discrepancy between our estimated private valuation and the returns in the labor market: Parents may underestimate the value of birth weight, opening the door for new policy interventions to increase health at birth via informational campaigns.

JEL Classification Codes: C9, I1, J1

Keywords: Discrete Choice Experiments; Early life health; Value of Health; Willingness to Pay

*The experiment documented in this paper has passed ethical approval at the Oxford Centre of Experimental Social Sciences (CESS), and been registered as project ETH-160128161. Any errors contained in the paper are our own.

1 Introduction

A baby's birth weight is a commonly measured and used endowment to capture the immediate stock of health early in life (Almond and Currie, 2011; Almond, Currie and Duque, 2017). The importance of the fetal period as a predictor of health throughout the life course has been recognized in a series of influential papers by Barker and coauthors on the fetal origins of disease (Barker et al., 1989; Barker, 1990, 1995), with considerable and ever-growing evidence in economics that insults to fetal health have enduring and significant costs throughout life (Case, Fertig and Paxson, 2005; Almond, 2006; Currie and Moretti, 2007; Black, Devereux and Salvanes, 2007; Almond, Edlund and Palme, 2009). These findings justify sizeable welfare programs targeted at babies with poor endowments early in life, such as those focusing on low birth weight infants (Almond, Chay and Lee, 2005; Bharadwaj, Løken and Neilson, 2013) and pre-natal nutrition programs, such as the Special Supplemental Nutrition Program for Women Infants and Children (WIC).[1]

Despite a large body of evidence on the importance of birth weight and considerable public investment, little is known regarding the private valuation of this birth outcome. Two main questions arise: First, how do taxpayers in general value birth weight? Second, how do parents perceive the value of birth weight? Knowing the value which people place on birth weight is of public concern and a fundamental policy issue, in particular as a key ingredient to policies focused on parental behavior *prior* to and *during* gestation. To the degree that a wide range of (costly) parental behaviors can positively impact birth weight (Rosenzweig and Schultz, 1983; Sexton and Hebel, 1984; Chevalier and O'Sullivan, 2007), the perceived importance of birth weight to parents may have significant effects on these behaviors, and on birth weight and other child outcomes throughout the life course.

In this paper we aim to provide the first estimate of the importance of birth weight to individuals (and parents), as measured by their Willingness to Pay (WTP) for birth weight. In order to do so, we conducted a series of discrete choice experiments on Amazon Mechanical Turk, an online labor market platform increasingly used both within and outside of the economics literature (Kuziemko et al., 2015; Jordan et al., 2016). We conducted these experiments with approximately 1,000 respondents, who were each asked to consider seven pairs of birth

[1] The total national spending on WIC in 2015 was 6.2 billion dollars. There are also non-targeted programs which have been reported to impact birth outcomes. Almond, Hoynes and Schanzenbach (2011) investigate the impact of the introduction of the modern Food Stamp Program (FSP) on birth outcomes. Using variation in the month FSP began operating in each US county, they find that pregnancies exposed to FSP three months prior to birth yielded deliveries with increased birth weight, with the largest gains at the lowest birth weights.

scenarios sequentially, amounting to around 14,000 different birth scenarios with a number of different characteristics. These characteristics were each orthogonally varied both within and between experimental subjects. Specifically, we performed conjoint analysis, a method first described by Lancaster (1966), which has recently been shown to perform more favorably than alternative experimental techniques when compared with real-world choice behavior (Hainmueller, Hangartner and Yamamoto, 2015).

These experiments allow respondents to reveal their preferences (or lack thereof) over a range of birth characteristics. In particular, we randomize a baby's birth weight, birth monetary costs, gender, and birth timing. Birth weight was randomized within the normal range of 2,500 to 4,000 grams. While many studies focus on low birth weight (LBW, or weights less than 2,500 grams) as their indicator of interest, we restrict our analysis to the "normal" range for two reasons. First, not only do continuous measures of birth weight have greater explanatory power for a large range of variables than a LBW indicator (Black, Devereux and Salvanes, 2007), but recent evidence also suggests that marginal increases in birth weight within the normal weight range are particularly important for well-being. Royer (2009) suggests that given this fact, babies born in the normal range of weights should receive *more* research attention.[2] This agrees with the findings of Currie and Moretti (2007) who observe that intergenerational links between a mother and her children's birth weights are particularly sensitive to changes above 2,500 grams. Second, from a purely practical standpoint, we focus on the normal range of birth weights to avoid priming effects (i.e., respondents linking low birth weight with other health conditions such as prematurity), thus confounding our estimates for the WTP of birth weight alone.

Our results suggest that, similar to researchers and policymakers, individuals view birth weight as an important and valuable characteristic. All else constant, a baby weighing 3,400 grams (7lbs, 8oz) is 18 percentage points (pp) more likely to be chosen than one weighing 2,500 grams (5lbs, 8oz). We estimate that over the normal range of birth weights, experimental participants would be willing to pay on average \$1.44 for each additional gram of weight, and that for those who are parents, this rises to \$1.72. Moreover, we observe a hump-shaped relationship in WTP, with the marginal WTP becoming negative at the top end of the range.

[2]In full, (Royer, 2009) reports (p. 52):

> "I find that the effects of birth weight on long-run outcomes are nonlinear and for educational attainment, in particular, are largest above 2,500 grams, the cutoff for defining low birth weight. These findings suggest that babies with birth weights outside the lower tail of the distribution (i.e., outside the range of low birth weight) should receive more attention."

This pattern mirrors the estimated effect of birth weight on health (Case, Fertig and Paxson, 2005) and on salary (Behrman and Rosenzweig, 2004). Depending on the point of the birth weight distribution considered, the WTP for birth weight is estimated to be as high as $12.50 per gram or as low as -$4.27 per gram (rising to $14.97 and -$7.58 for parents).

Despite the positive and economically significant estimate of the average WTP for birth weight, this is much smaller than the valuation inferred from public programs, which we estimate in a range from $12 to $18 per gram, and the present value of expected labor market returns to birth weight, which we compute at $14 per gram. The large gap between the average WTP ($1.44 among all individuals, $1.72 among parents) for birth and the range of inferred valuations [$12, $18] is robust to adjusting for the omission of the range of low birth weights. This large gap is puzzling for at least two reasons. First, parents have been shown to generally invest *more* in their child's health than their own health (Agee and Crocker, 2008). Second, the labor market returns to birth weight are *only* a subset of the full returns: Convincing evidence suggests that increases in birth weight reduce the prevalence of chronic morbidities (Barker, 1995; Almond and Mazumder, 2005; Johnson and Schoeni, 2011*a*), mortality (van den Berg, Lindeboom and Portrait, 2006), and a range of psychological outcomes (Fletcher, 2011).[3]

One potential explanation for our findings is that individuals, and even parents, underestimate the returns to birth weight. That parents have biased beliefs on the returns to human capital investments is consistent with recent work by Cunha, Elo and Culhane (2013) and Boneva and Rauh (2016). In an influential paper, Jensen (2010) demonstrates that correcting expectations regarding the returns to education increases investments in education. We present a parsimonious model which is consistent with both the different WTP by parental status and the discrepancy between our estimated private valuation and the returns in the labor market. If parents are not aware of the magnitude of the returns to birth weight, they may be investing sub-optimally in pre-natal behaviors. This opens the door to new policy options to augment health at birth via informational campaigns. Communicating the existing knowledge on the influence of health at birth on long-term outcomes to future parents is plausibly an intervention that may address sub-optimal parental investment behavior in early life health. Our finding and policy recommendation echo the research agenda on the early origins of inequality and its

[3]Here we focus largely on the long-run impacts of birth weight over an individuals' life. However, there is an additional even larger body of work documenting the importance of birth weight in explaining early life outcomes and human capital outcomes in childhood. These include Conley, Strully and Bennet (2003), Almond, Chay and Lee (2005), Oreopoulos et al. (2008), Lin and Liu (2009), Fletcher (2011), Torche and Echevarría (2011), Gupta, Deding and Lausten (2013), Figlio et al. (2014), and Bharadwaj, Eberhard and Neilson (forthcoming), among many others.

lifetime consequences.

In what follows, we describe the Mechanical Turk data and experimental set-up in section 2 and the methodology for estimating willingness to pay (WTP) in section 3. In section 4 we present our experimental estimates of the WTP for birth weight. In section 5 we compare our experimental estimates to inferred WTP for birth weight from US public programs, and to the value of birth weight in the US labor market. We provide a conceptual framework to explain the different WTP by parental status and the discrepancy between our estimated private valuation and the returns in the labor market in section 6. We briefly conclude in section 7.

2 Data Description

We collected data on preferences over birth characteristics by running discrete choice experiments on Amazon's Mechanical Turk (MTurk) online platform. This platform is a market place which provides access to a pool of US MTurk workers who are paid per completed HIT (Human Intelligence Task). We posted an HIT request to recruit respondents to complete a series of discrete choice experiments (described further below) as well as a series of demographic questions. These demographic questions were asked *after* the completion of the experiments to avoid any framing or experimenter demand effects (Zizzo, 2010), and the survey was advertised as a general demographic survey, not mentioning anything related to births, babies or birth weight. Mechanical Turk respondents have been documented to have desirable characteristics, and be more representative of the US population than other frequently-used subject pools such as college student samples (Berinsky, Huber and Lenz, 2012). Mechanical Turk samples are increasingly used both within and outside of the economics literature (Kuziemko et al., 2015; Jordan et al., 2016).

Our published HIT requested the completion of a short survey, and the experiment was conducted on a Monday in September 2016. Workers were paid $1.10 for a 6 minute experimental survey (average length), resulting in an effective hourly pay rate of approximately $11. The survey needed to be completed in order to be able to receive payment, and it was impossible to move forward if the question on the screen was not answered. We required that respondents must be from the United States,[4] and in order to maximize the likelihood that workers were

[4]Workers on Mechanical Turk are required to have a US Social Security Number.

based in the US at the time of completing the survey, this was launched at 9:00 AM East Coast Time. By 2:13 PM of the same day 1,002 valid responses were collected. We also required that workers had completed at least 100 tasks on MTurk in the past, and had achieved an approval rating of greater than 95% on these tasks. These restrictions are common in Mechanical Turk research (Berinsky, Huber and Lenz, 2012; Francis-Tan and Mialon, 2015). Of the 1,002 valid responses, we removed a small number based on a set of pre-defined consistency checks. These were: (a) workers whose geographical IP address placed them outside of the US at the time of survey (36 respondents, or 3.6%), any respondents who failed a consistency check where a question was repeated at the beginning and end of the demographic portion of the survey (8 respondents), and any respondents completing the entire exercise in under 2 minutes (6 respondents).[5] The final sample thus consists of 952 respondents.

The geographic location of these respondents within the US (based on their IP address) is provided in Figure 1. The geographic coverage is broadly representative of the US population. In Appendix Table A1 we compare our MTurk respondent coverage with the US population from 2015 (United States Census Bureau, 2015). In general, we see that our MTurk sample lines up well with the national population at the state level, however there are a number of exceptions, such as the lower number of respondents from California, most likely reflecting the earlier time zone on the West Coast. In our main specifications we always use the unweighted MTurk sample, however as an alternative we follow Francis-Tan and Mialon (2015) in re-weighting the sample so that results are representative of the US population.

Finally, summary statistics of the respondents are provided in Table 1. Slightly more than half of all respondents are female (56%), and the ages of respondents range from 18 to 75 years (with a mean age of 36 years); 84% are white and 6% are Hispanic. Approximately half of the respondents are parents (51%), and of those who are non-parents, 45% plan to have children or are already pregnant (implying that 27% of respondents are neither parents, nor plan to become so). In total, 45% of the respondents are married, 68% are employed and 89% have at least some college.

[5] In the appendix to this paper we demonstrate that nonetheless, our results remain largely unchanged if we do not remove these respondents.

3 Methodology

3.1 Discrete Choice Experiments: Attributes and Levels

In order to estimate the perceived importance of birth weight in terms of willingness to pay, we run discrete choice experiments on a large sample of US-based respondents. A discrete choice experiment (DCE) is a type of Conjoint Analysis (CA): An experiment in which respondents are asked to choose their preferred option from a set when a number of attributes are varied simultaneously. CA was borne from early work in consumer theory in which tastes for goods owe to the collection of their characteristics (Lancaster, 1966). In the past, CA has been used to measure preferences over medical care in a variety of contexts, including the valuation of waiting times (Propper, 1990, 1995), alternative miscarriage treatment options (Ryan and Hughes, 1997), asthma medications (King et al., 2007), or depression management (Wittink et al., 2010). Conjoint analysis has recently been shown to be the best-performing experimental design when compared with actual choices made (Hainmueller, Hangartner and Yamamoto, 2015).

Our birth choice experiments consist of asking respondents to consider a series of paired birth scenarios, while focusing on four attributes of each birth scenario. We use a main-effects, orthogonal (all attribute levels vary independently) and balanced (each level of an attribute occurs the same number of times) experimental design. In the experiment the attributes are combined to form various (hypothetical) birth scenarios, all about a hospital birth of the first child with no complications. The attributes considered are the baby's weight at birth (5lbs, 8oz; 5lbs, 13oz; 6lbs, 3oz; 6lbs, 8oz; 6lbs, 13oz; 7lbs, 3oz; 7lbs, 8oz; 7lbs, 13oz; 8lbs, 3oz; 8lbs, 8oz; or 8lbs, 13oz), the out of pocket expenses associated with the birth ($250; $750; $1,000; $2,000; $3,000; $4,000; $5,000; $6,000; $7,500; or $10,000), the sex of the child (Boy or Girl), and the season in which the baby is born (Winter, Spring, Summer or Fall). The latter options are used in order to avoid priming respondents into thinking that we are interested in birth weight. As these attributes are all orthogonally varied, the effect of each characteristic on the likelihood that a particular birth is chosen is separately identified (Marshall et al., 2010).

Each respondent was asked to consider seven pairs of birth scenarios in an iterative fashion. In order to move forward in the experiment a choice must be made for each pair, and once the choice has been made the respondent may not go back and revise their choice. In each case the two pairs were displayed side-by-side on a single screen, and respondents were asked to indicate which was their preferred birth scenario. As well as randomizing the level of each attribute on each profile, the order of the attributes was randomized, however to reduce the cognitive load to respondents the ordering of attributes was only randomized once, and then fixed across the seven pairings that the respondent ranked. The DCE's framing and the explanation of the attributes shown to respondents is displayed in Appendix Figures A1 and A2. In Appendix Figure A3 we display an example of a pair of birth scenarios as presented to respondents. Appendix B describes the survey procedure followed by respondents in more detail.

The levels of attributes were chosen to represent plausible values from the US population (Ryan and Farrar, 2000), and extreme values were avoided to prevent the likelihood of "grounding effects" (or corner solutions) (Bridges et al., 2011). In order to minimize the likelihood that respondents would employ simple heuristics in answers, we limited the number of attributes (four) which need be considered. As discussed in Bridges et al. (2011), we observe in experimental responses that such heuristics are not employed given response sensitivity to all dimensions studied. Birth weights were always presented in pounds and ounces, as this experiment was run with a US-sample. As well as indicating that all births were complication-free, only birth weights over the normal range of 2,500–4,000 grams were included (11 evenly-spaced weights were defined in this range, and expressed in pounds and ounces in the experiment). This range includes the vast majority of all births in the US. According to full vital statistics of 2013 from the National Vital Statistics System (see Appendix Figure A4), 8.02% of births were low birth weight ($<$ 2,500 grams), and 7.89% were large for gestational age at birth ($>$ 4,000 grams). An opt-out option was not included in any of the discrete choices. This has been suggested to have desired properties such as avoiding non-random opt-out of all questions (Bekker-Grob, Ryan and Gerard, 2012; Veldwijk et al., 2014).

3.2 Estimating Equations

Consider a sample of $i \in \{1, \ldots, N\}$ individuals, each of whom considers K choice tasks in which they must decide between J options (profiles, or in our case, birth scenarios). Each profile contains L attributes, where each particular attribute l consists of discrete levels of the variable. In the case of the DCE described above, we have $N = 952$ respondents, $K = 7$ choice tasks per respondent, $J = 2$ profiles per task, and $L = 4$ attributes.[6] We follow Hainmueller, Hopkins and Yamamoto (2013) in defining a treatment vector T_{ijk}. This treatment vector has

[6] These four attributes have 2, 10, 11 and 4 levels respectively for sex, out of pocket costs, birth weight and season of birth.

L cells, and summarizes for individual i, at choice task k, for profile j, the full set of attributes observed. Each particular attribute T_{ijkl} is randomly assigned from among all the levels of l, the assignment of which is orthogonal to all other attributes the respondent sees. Using the potential outcomes framework, we define a binary variable $Y_{ijk}(\bar{\mathbf{t}})$ which takes the value 1 if respondent i would choose profile j on choice set k if faced with the set of attributes $\bar{\mathbf{t}}$, or 0 if the profile would not be chosen.

We are interested in estimating two quantities. Firstly, we would like to estimate, *ceteris paribus*, the likelihood that a birth scenario is chosen given that a particular birth weight is observed (compared with an omitted base category). Secondly, we would like to estimate the willingness to pay for birth weight, by combining the information from both variations in birth weight and variations in out-of-pocket costs.

Hainmueller, Hopkins and Yamamoto (2013) call this first quantity the Average Marginal Component Effect (AMCE) and demonstrate that under reasonably weak assumptions,[7] it can be recovered using a non-parametric sub-classification estimator, conditional regression, or a simple difference in means. The logic of the AMCE is to capture the change in the likelihood that a given profile would be chosen if the l^{th} component were changed from t_0 to t_1, or in our case, a change in birth weight.[8]

Under the controlled randomization in conjoint analysis, Holland (1986)'s fundamental problem of causal inference is resolved by construction, as on average there will be no correlation between observing the particular level of an attribute and individual correlates. Treatment units are thus those who observe a particular t_1, while those who do not act as controls. In practice, to estimate the change in the likelihood that a birth scenario is chosen given a change in birth weight (or any other attribute), we estimate the following two types of regression:

$$Pr(Y_{ijk} = 1) = \Lambda\left(\alpha + \beta Costs_{ijk} + \gamma BW_{ijk} + \sum_{r=2}^{4} \delta_r SOB_{ijk,r} + \kappa Girl_{ijk} + \mu_j + \phi_k\right) \tag{1}$$

[7]These assumptions relate to randomization of attributes, and stability of respondent behavior regardless of the number of profiles that they have seen or the order of the attribute in the profile. This first assumption holds by construction in our experiment. A benefit of the set-up of the DCE is that even if order and round effects are not completely neutral, these can be flexibly captured using fixed effects in a regression.

[8]Formally, the AMCE is defined as (Hainmueller, Hopkins and Yamamoto, 2013):

$$E[Y_i(t_1, T_{ijk[-l]}, \mathbf{T}_{i[-j]k}) - Y_i(t_0, T_{ijk[-l]}, \mathbf{T}_{i[-j]k}) | (T_{ijk[-l]}, \mathbf{T}_{i[-j]k}) \in \tilde{\mathcal{T}}]$$

which can be quite easily calculated by integrating over all of the other attributes and levels except for t_1 (the treatment of interest) and t_0 (the baseline level for the attribute). These other attributes and levels are denoted as the set $\tilde{\mathcal{T}}$ here.

and

$$Pr(Y_{ijk}=1)=\Lambda\left(\alpha+\beta Costs_{ijk}+\sum_{q=2}^{11}\gamma_q BW_{ijk,q}+\sum_{r=2}^{4}\delta_r SOB_{ijk,r}+\kappa Girl_{ijk}+\mu_j+\phi_k\right), \quad (2)$$

where $Y_{ijk}=1$ if the birth scenario j is chosen, Λ is the *cdf* of the logistic distribution, $Costs_{ijk}$ denotes the out of pocket expenses associated with the birth scenario j, BW_{ijk} is the birth weight associated with the birth scenario j, $BW_{ijk,q}$ is equal to 1 if the birth weight category of the birth scenario j is q, $SOB_{ijk,r}$ is equal to 1 if the season of birth category of the birth scenario j is r, $Girl_{ijk}$ is 1 if the gender of the baby of the birth scenario j is girl, and μ_j and ϕ_k are option-profile and choice-task order fixed effects, respectively. Standard errors are clustered at the level of the respondent to capture the (likely) positive correlations among choices based on attributes by a particular respondent. We estimate equations (1) and (2) and report average marginal effects. We omit from equation (2) the lowest birth weight category as the baseline level, implying that all marginal effects of each birth weight should be interpreted as the marginal likelihood of choosing a birth scenario given birth weight q in place of the lowest birth weight (2,500 grams).

Using equation (1), we are also able to estimate the willingness to pay for birth weight. The marginal effects on the likelihood of choosing a particular birth scenario given an increase in the particular attribute, conditional on all other attributes, are:

$$\frac{\partial Pr(Y_{ijk}=1)}{\partial Costs_{ijk}}=\beta\Lambda'(\cdot) \qquad \frac{\partial Pr(Y_{ijk}=1)}{\partial BW_{ijk}}=\gamma\Lambda'(\cdot),$$

where Λ' is the *pdf* of the logistic distribution. Given these marginal effects, the marginal rate of substitution (MRS) between birth weight BW and the price of a given birth (the out of pocket costs)—which measures the change in costs that a respondent would be willing to withstand for a marginal increase in birth weight—is given by:

$$MRS_{BW,Costs}=\frac{\frac{\partial Pr(Y_{ijk}=1)}{\partial BW_{ijk}}}{\frac{\partial Pr(Y_{ijk}=1)}{\partial Costs_{ijk}}}=\frac{\gamma}{\beta}.$$

Multiplying this quantity by minus 1 gives precisely the willingness to pay:

$$WTP_{BW}=-\frac{\gamma}{\beta}=-\frac{\partial Cost_{ijk}}{\partial BW_{ijk}}.$$

Note that in the above calculation we take the negative so that costs are interpreted as the positive amount that must be paid rather than the negative change in financial resources. This WTP can also be derived quite straightforwardly from a model of the indirect utility function as described in Zweifel, Breyer and Kifmann (2009).

The average WTP is obtained by taking the ratio of the average marginal effects. In order to calculate the confidence interval associated with the WTP we use the *delta method*, which is both simple and shown to perform well under simulation (Hole, 2007*a*). We also find that this confidence interval is quite comparable to that produced when using *block bootstrapping*, as outlined further in the following section.

4 Experimental Results

4.1 Main Results

4.1.1 Average Marginal Component Effects

In Figure 2 we present our experimental results. This figure displays point estimates of the likelihood of preferring a particular birth scenario given each characteristic, compared with an omitted base category for each characteristic. Along with each point estimate, the 95% confidence interval is plotted, clustering by respondent. While we present cost as a linear variable measured in 1,000s of dollars, in Appendix Figure A5 the same results are presented with costs displayed as the same categorical measure observed by respondents, and in Appendix Figure A6 we document that results are largely unchanged if we work with the full sample of 1,002 respondents rather than the preferred sample of 952 respondents meeting inclusion criteria.

The top panel displays the likelihood of choosing a birth scenario given a particular birth weight, compared to being shown the minimum sample birth weight of 5lbs, 8oz (2,500 grams). In each case, higher birth weights are associated with a greater likelihood of choosing the corresponding birth scenario. The most preferred birth weight (based on point estimates) is 7lbs, 8oz (3,400 grams), which results in a birth scenario being approximately 18 pp more likely to be chosen than the omitted base category. The magnitudes of the estimates are large. With the exception of 5lbs, 13 oz, all higher birth weights are at least 10 pp more likely to be chosen, and in each case the difference is statistically significant. In addition, there appears to be a hump-shaped pattern, with the most preferred births being those towards the middle of the normal birth weight range, and lower preferences for those at the extremes of the normal weight range.

4.1.2 Willingness to Pay for Birth Weight

As discussed in section 3, we can combine estimates of average marginal component effects to generate estimates of the WTP for each characteristic. In Table 2, column 1, we assume a linear functional form for birth weight. By comparing the change in the likelihood of choosing a birth scenario based on an increase in birth weight with the change in likelihood due to an increase in costs, we estimate that the average WTP for an additional 1,000 grams in the full sample is $1,438.3, or $1.44 per gram. However, as we observe in column 2, the relationship between birth weight and the likelihood of choosing a birth scenario is non-linear. In Figure 3 we document the WTP of all birth weight options, with respect to the minimum birth weight in the sample. We observe that the largest relative difference occurs at 3,400 grams (7lbs 8oz, compared with the omitted base of 2,500 grams), with a WTP of $2.14 per gram. Finally, and as expected, we observe that all else equal, higher costs result in a birth scenario being less likely to be preferred. On average, for each additional $1,000 in out of pocket expenses, the likelihood of choosing a birth scenario falls by nearly 10 pp. The non-linear estimates of these parameters are displayed in Appendix Figure A5. In Appendix Table A2 we find that *re-weighting* the population to be representative of the US at the state level leads to quantitatively and qualitatively similar results for each component as well as the estimated WTP.

Non-parametric WTP estimates. When calculating the average WTP of birth weight as a single figure, this is based on a specification in which birth weight (and costs) enter the estimating equation linearly. However, Figure 3 provides the WTP for each particular birth weight, as compared to a birth weight of 2,500 grams (or the relative willingness to pay). In Figure 4 we also present the marginal WTP, or the difference in WTP when moving between contiguous categories. It is observed that the largest marginal change occurs at the lower end of the distribution, when moving from approximately 2,600 to 2,800 grams. If we calculate the WTP over only this range this gives a maximal WTP of $12.49 per gram (95% CI: $8.21;$16.77) for the marginal increase in birth weight. On the other hand, we note that at the upper end of the birth weight range, the marginal WTP turns significantly negative, when moving from 3,800 to 4,000 grams, at -$4.27 per gram (95% CI: -$8.34;-$0.21). The hump-shaped relationship is statistically significant.[9]

[9] We have investigated this in three different ways: (a) If we run a base model with WTP and WTP squared, then the WTP squared term is highly significant in predicting choices; (b) The same results holds including the full controls; (c) If we run a linear model, and conduct a RESET test, we reject the H_0 that the model has no

Interestingly, this hump-shaped relationship lines up with a range of evidence related to the *benefits* of birth weight. For example, Case, Fertig and Paxson (2005) show a similar hump-shaped relationship for the effect of birth weight on health, with increases in self-reported adult health as birth weight increases, though a point of inflexion exists at very high birth weights (see for example their Figure 1).[10] Similarly, Behrman and Rosenzweig (2004) find a point of inflection of the labor market returns to birth weight which lines up quite closely with our estimates (specifically, they find that for births approximately 2 standard deviations above the mean of fetal growth,[11] labor market returns to a marginal increase in birth weight turn from positive to negative). We return to these estimates further below in section 5.3.

Alternative confidence intervals for WTP estimates. The confidence intervals estimated on the average WTP discussed above are always calculated using the delta method. However, we find relatively little difference if we estimate confidence intervals under block bootstrapping. In order to implement this block bootstrap procedure we perform 1,000 bootstrap replications, re-sampling with replacement over experimental respondents rather than over profiles. This block bootstrap calculation leads to a 95% confidence interval for willingness to pay of [$1,107.2;$1,769.3] for the main sample. This is marginally wider, though qualitatively very similar to the main calculation of [$1,119.4;$1,757.1] using the delta method and reported in Table 2. In all cases examined, block bootstrap confidence intervals lead to largely similar findings, and in some cases even lead to slightly *less* wide confidence intervals. Full comparisons of confidence intervals between methods are available in Appendix Table A3.

The value of other attributes. Table 2 also sheds light on preferences for other birth attributes. We find no evidence of any elicited preference for the baby's gender on average. Indeed, in both specifications displayed in Table 2 estimated coefficients on the baby being a girl are quite tightly estimated zeros (ranging from 0.000 to 0.001 with clustered standard errors of 0.010). However, when estimating separately by the gender of the respondent, we observe a preference for boy children among male respondents: In column 3 of Appendix Table A4 males are 3 pp more likely to choose a son than a daughter. This is in agreement with

omitted variables with $p < 0.001$.

[10] Indeed, this hump-shaped pattern is commonly observed in many morbidity and mortality measures across populations (Wilcox, 2001).

[11] Although these values can not be precisely converted to birth weight from the rate of fetal growth reported by Behrman and Rosenzweig (2004), our approximate calculation suggests that if fetal growth is a good proxy for birth weight and hence the turning point for birth weight is similarly two standard deviations above the mean, this would result in a turning point of $90.2\text{oz} + 2 \times 17.9\text{oz} = 126\text{oz}$, or 3.572 kg (all values from Behrman and Rosenzweig (2004) Table 1 and Figure 8).

the results of Dahl and Moretti (2008) who document a demand for sons, particularly among fathers.[12] When considering season of birth we observe a greater likelihood to choose birth scenarios in the spring (relative to winter), consistent with the existence of a demand for certain seasons of birth (Clarke, Oreffice and Quintana-Domeque, 2016).

4.2 Heterogeneous Effects

The headline estimated effect for average WTP suggests a value of \$1.44 per gram over the range examined (95% CI: \$1.12;\$1.76). This value is calculated using the entire sample of respondents. We briefly consider estimates for particular subgroups of interest, namely parents, non-parents, and non-parents who do and do not plan to have children. All these basic demographic characteristics were asked *after* the completion of the experiments.

Figure 5 displays outcomes of the discrete choice experiments for each group. Panels A and B split by parental status (parents versus non-parents), and then panels C and D further split non-parents by desired childbearing status (those who plan to have children or are already pregnant versus those who do not plan to have children). These figures reveal that, firstly, parents are the most sensitive to changes in birth weight, and also show the most clear non-linear pattern. Non-parents display a much flatter profile, and are consistently less likely to choose a birth scenario given a higher birth weight. When further splitting by those who plan to have children and those who do not, we observe that the profile for planners is comparable to that for parents (hump-shaped), while those who do not plan are significantly less likely to choose a birth scenario based on an increase in weight. We examine these results, along with precise values for WTP, in Table 3. We turn to these results now.

Parents vs. non-parents. In columns 2 and 3 of Table 3 we estimate the linear specification for birth weight and costs for parents and all non-parents. We observe, firstly, that although both groups are similarly impacted by increases in costs (a birth scenario is 6.2 pp or 6.4 pp less likely to be chosen for each \$1,000 increase in costs for parents and non-parents respectively), point estimates on birth weight are higher for parents than for non-parents. An increase in 1,000 grams of birth weight increases the likelihood that parents choose a profile by 10.6 pp,[13] while

[12] In column 5 of Appendix Table A4 we focus on fathers, rather than all males, and observe essentially the same coefficient but with a larger standard error on account of the reduction in sample size.

[13] The identical samples using a non-linear specification for birth weight are displayed in Appendix Table A6. The importance of birth weight among parents dominates that among non-parents at every point on the birth weight distribution with the exception of the extreme values, as observed with the greater curvature in Figure

only by 7.5 pp for non-parents. This is reflected in considerably different average WTP values. The average WTP for a gram of birth weight among *parents* is \$1.72, (95% CI: \$1.23;\$2.20), compared to \$1.17 among non-parents, (95% CI: \$0.75;\$1.59). If instead we consider the *largest* value of WTP for each group based on the non-parametric estimates described in Figure 5, these are \$14.97 per gram (95% CI: \$8.64;\$21.30) for parents, and \$10.38 per gram (95% CI: \$4.58;\$16.18) for non-parents, both based on the change between about 2,600 to 2,800 grams. Perhaps unsurprisingly, across the board parents are more likely than non-parents to be swayed by changes in non-pecuniary attributes: For parents birth weight and birth season are considerably more important than for non-parents. We estimate a pooled specification where we interact a dummy for being a parent with birth weight in Appendix Table A5. This recovers the 3.1 pp difference in birth weight between parents and non-parents discussed above, and also allows us to estimate the WTP differential and its 95% confidence interval. While the average WTP differential is considerable—at \$491 for an additional 1,000g—it is not statistically distinguishable from zero (95% CI: -\$119;\$1,101).

Parents vs. non-parents who plan to have children. Columns 4 and 5 of Table 3 display estimates for non-parents, separating by whether they plan to have children or do not plan to have children. If we compare figures for parents with those of non-parents who state that they *do* plan to have children, we see that the point estimate on birth weight is slightly higher among the former. As above, parents are 10.1 pp more likely to choose a birth scenario for each 1,000 grams increase in birth weight, while the same figure for non-parents who plan to have children is 9.1 pp. The average WTP of the non-parent planners is \$1.38 per gram (cf 1.72 for parents) with a 95% CI of \$0.76–\$1.99. Once again, if we refer to Appendix Table A5 we see that the difference in average WTP is not statistically significant (column 3). If instead of focusing on average WTP we focus on the maximum WTP from non-parametric estimates, this is \$8.35 per gram (cf 14.97 for parents) with a 95% CI of \$0.14–\$16.69, in this case observed when moving from 2,500 to 2,650 grams.

Non-Parents who plan to have children vs. non-parents who do not plan to have children. Finally, if we compare the two groups of non-parents, those who plan to have children and those who do not, we see a large average difference in the likelihood to choose a birth given an increase in birth weight. As above, non-parents who plan to have children are 9.1 pp more likely to choose a birth scenario for each 1,000 gram increase in birth weight,

5.

while non-parent non-planners are only 6.4 pp more likely. The average WTP for each group is \$1.38 per gram for those who plan versus only \$1.00 per gram for those who do not plan, or a 38% increase.

Bringing this all together, our headline result of an average WTP of \$1.44 per gram shows considerable heterogeneity by groups. This ranges from as high as \$1.72 per gram for parents, to as low as \$1.00 per gram for non-parents who have no stated desire to have children. Non-parents who do plan to have children fall somewhere in the middle at \$1.36 per gram. While only suggestive, it is noteworthy that individuals value more highly early-life health after becoming parents, even compared with prospective parents. We will return to this point in section 6.

5 Private WTP, Public WTP, and "The Returns to Birth Weight"

In this section we assess our experimental estimates for private WTP in light of a number of results from the economic literature on birth weight. In order to benchmark our estimates we ask two questions: Firstly, how does individual WTP compare to public WTP? We infer the public WTP for birth weight from two large social safety net programs. The first is WIC, a program which explicitly targets neonatal health, and the second is the Food Stamp Program, which, although not designed to target neonatal health outcomes, has been documented to have important impacts on early-life human capital measures. Secondly, how does the private WTP compare to the total expected (labor market) benefits accruing to birth weight over the life cycle? While the labor market returns to birth weight are a clear lower bound on the value of birth weight, these are all private returns, and so provide a benchmark value with which to compare the private WTP estimates discussed in the previous section.

5.1 Comparison with the Public WTP estimated from a Targeted Program

Our main findings suggest that on average individuals are willing to pay \$1.44 for each additional gram of birth weight over the normal birth weight range, and this increases to \$1.72 among parents.[14] It is of interest to ask how this private WTP compares with the inferred WTP from public investment. While much of the benefits of increases in birth weight accrue

[14] If instead of focusing on average WTP we focus on the highest WTP in any range studied, these are (respectively) \$12.49 and \$14.97.

to families,[15] increases in birth weight also have important public returns, including benefits flowing from reductions in public health care spending, and lower usage of means-tested public benefits programs (Almond, Chay and Lee, 2005).

We can provide a comparison between public and private WTP for birth weight using estimates from the WIC, which provides food and education to pregnant and postpartum breastfeeding women who earn less than 185% of the US federal poverty guideline.[16] By combining estimates of the cost per WIC user with estimates of the benefit in terms of additional birth weight, we can arrive at an estimate of the public WTP per gram of birth weight.

Ben-Shalom, Moffitt and Scholz (2011) document that WIC participation costs \$54 per enrollee per month, and according to WIC administrative data, 56.9%, 34.7% and 7.8% of participants enroll in the first, second or third trimester respectively (Johnson et al., 2013). Using trimester midpoints to calculate months of enrollment, this suggests approximate total costs of covering a single pregnant woman of \$321.[17] There exist a very large range of estimates of the impact of the WIC program. Recent studies suggest that the true impact may fall towards the lower end of the estimated spectrum. Among plausibly causal estimates, Rossin-Slater (2013) estimates that participation has a mean impact of 27 grams of birth weight, and Hoynes, Page and Stevens (2011) estimate impacts of 18-29 grams. In the case of the highest estimated impact, public WTP equates to \$321/29 grams = \$11.07 per gram, while for the lowest estimate, the WTP equates to \$321/18 grams = \$17.83 per gram. Both estimates of the public WTP exceed our experimental estimates of the private average WTP, for both the whole sample of respondents (\$1.44), and the sub-sample of parents (\$1.72).

5.2 Comparison with the Public WTP estimated from an Untargeted Program

The evidence from WIC discussed above estimates the inferred WTP using a targeted program which explicitly focuses on maternal and newborn health. Nevertheless, there are a range of other public programs which, while not explicitly targeting infant health, have been

[15] Such as private returns in the labor market, a reduction in out of pocket medical spending during childhood, and increases in education (Behrman and Rosenzweig, 2004; Oreopoulos et al., 2008).

[16] A comprehensive discussion of recent work on WIC is available in Bitler and Karoly (2015).

[17] This is calculated as

$$Cost = 54 \times (7.5 \times 0.569 + 4.5 \times 0.347 + 1.5 \times 0.078) = \$321.1.$$

documented to have unintended effects on these outcomes. Perhaps the largest of these is the Food Stamp Program (or FSP), now known as the Supplementary Nutrition Assistance Program (SNAP), which provided support for 44.2 million people in 2016 at a total cost of 70.9 billion dollars.

Almond, Hoynes and Schanzenbach (2011) provide a particularly well-identified estimate of the effect of the FSP on infant health, and in particular, on birth weight. Using county-level variation in program roll-out[18] and vital statistics data from 1968-1977 (a period of sharp program expansion), they estimate that the average county-level birth weight in program counties increased by between 2-2.6 grams for white pregnant women and 1.7-5.5 grams for black pregnant women. While this estimate is a county-level value, they also provide an *individual*-level effect based on county usage rates. These individual effects, which amount to 20.27 grams (white) or 31.69 grams (black),[19] allow us to estimate the inferred public WTP for a gram of birth weight when combined with the costs per pregnant women.

In order to determine the costs per pregnant women, we focus on data on current costs and users (in order to be comparable to our estimated WTP in current dollars). We assign the full cost per program beneficiary for three months of use, for two main reasons. Firstly, there is considerable evidence that nutrition can affect birth weight (rather than survival or other morbidities) only over a relatively short window, in particular during the third trimester (see for example Stephenson and Symonds (2002), as well as evidence from the timing of exposure to the Dutch Famine (Schwarz, 2004) and the Argentine crisis of 2001 (Bozzoli and Quintana-Domeque, 2014)). In particular, this is also shown to be the case with the FSP (Almond, Hoynes and Schanzenbach, 2011). Secondly as the FSP is not targeted to pregnant women or to promote neonatal health *per se*, expenses will be ongoing, begin before pregnancy, continue after pregnancy, and bring with them considerable additional impacts beyond only their effect on gestating infants.

Using the final three months of pregnancy to estimate the typical costs for a pregnant woman, and average per person monthly costs from 2016 of \$125.50, the inferred public WTP for an additional gram in birth weight is approximately \$16.[20] Once again, the inferred public

[18]Of note, this identification strategy is broadly similar to that used by Hoynes, Page and Stevens (2011) in estimating the impact of the WIC program, ensuring that estimates in each of the sections of this paper are comparable.

[19]Refer to column 2 of Table 1 from Almond, Hoynes and Schanzenbach (2011).

[20]This is calculated using $(125.5 \times 3)/31.69 = \$11.9$ based on Almond, Hoynes and Schanzenbach (2011)'s estimates for black mothers and $(125.5 \times 3)/20.27 = \$18.6$ for estimates for white mothers. In addition, we know that 40.2% of food stamp users are white and 25.7% are black (United States Department of Agriculture, 2014). Hence, we can get a weighted average estimate using $40.2/65.9 = 0.61$ as the weight for white mothers,

WTP exceeds our estimates for the average private WTP by an order of magnitude.

5.3 Comparison with the Returns to Birth Weight in the Labor Market

It is well accepted that higher birth weight is associated with reductions in morbidity and mortality, and greater educational attainment and achievement throughout childhood.[21] Moreover, these impacts have been well-documented to perdure into adulthood and impact labor market outcomes. In Table 4 we review the range of papers which have estimated the long-run returns to birth weight in the US.[22] The data requirements for such an exercise are quite demanding, requiring information on the weights of babies at birth, family linkages, and completed education or labor market outcomes many years later. Of particular interest for this paper are the returns to birth weight in the labor market.

One way to benchmark the parental average WTP for birth weight is to determine how it compares to the present value of the flow of expected benefits during the life of their child. Thus, considering these well-estimated cases of the labor market returns to birth weight, we can discount expected returns back to the start of an individual's life, and compare it with our experimentally estimated WTP.[23]

The most convincing empirical estimates produced from the literature come from within-sibling or within-twin methods, which can be viewed as the effect of shifting the smaller of two siblings (or twins) to the weight of the larger sibling (twin). Our experimental estimates of WTP are taken as the relative to the omitted baseline category of 2,500 grams. So, in the sense that we think of 2,500 grams as the lower weight of a comparison pair, all our experimental estimates line up to the effect of additional intrauterine growth reported by the within twin or within family literature reviewed in Table 4.

For this exercise, we are most interested in those papers which provide estimates of the

and $25.7/65.9 = 0.39$ as the weight for black mothers. This leads us to a weighted average of $\$16 = 0.61 \times \$18.6 + 0.39 \times \$11.9$.

[21] For example, on morbidity, Conley, Strully and Bennet (2003), Almond, Chay and Lee (2005), Oreopoulos et al. (2008), and Gupta, Deding and Lausten (2013), and on early-life education, Lin and Liu (2009), Fletcher (2011), Torche and Echevarría (2011), Figlio et al. (2014), and Bharadwaj, Eberhard and Neilson (forthcoming), demonstrate a strong and plausibly causal link.

[22] A number of similar estimates exist in a non-US setting (for example Rosenzweig and Zhang (2013) in China, Black, Devereux and Salvanes (2007) in Norway, and Currie and Hyson (1999) in Great Britain), however in order to benchmark our WTP results in the US population we do not focus on these here.

[23] This should of course be considered as a lower bound to the true value of birth weight. Labor market returns are a convenient financial metric, but do not include any of the additional pecuniary or non-pecuniary benefits which may flow to parents from a higher birth weight child such as lower expected costs associated with medical care (Almond, Chay and Lee, 2005).

long-run returns to birth weight in the labor market. This precludes studies which only observe completed education, but not labor market outcomes, described in panel B of Table 4. Among those papers which have estimated the effect of birth weight on earnings, there are three papers that use twin or sibling fixed effects to leverage within family variation in birth weight to estimate returns conditional on genetic material. These are Behrman and Rosenzweig (2004), Johnson and Schoeni (2011*b*), and Cook and Fletcher (2015). Behrman and Rosenzweig (2004) and Cook and Fletcher (2015) estimate the impact of increases in (continuous) birth weight rather than a binary indicator for low birth weight (LBW), or birth weights inferior to 2,500 grams. Both studies find that headline effects of birth weight on earnings are positive and economically significant, although the estimates of Cook and Fletcher (2015) are only statistically significant in certain specifications.

Behrman and Rosenzweig (2004)'s results provide a point estimate of the labor market returns to birth weight in the US which suggests that "augmenting a child's birth weight by a 1 lb. increases her adult earnings by over 7%". According to the US Census Bureau (2016), the median personal income in the US in 2015 was $30,240. If we assume a working life which begins at the age of 25 and ends at the age of 60, we can calculate the present value of a 7% increase in wages as a deferred annuity. This calculation suggests that the present value of an additional pound of birth weight is $10,235.[24] Dividing this value by the 454 grams in a pound gives the labor market value of a gram of weight of $23. If we assume that only approximately 60% of the working age population will actually be employed in the labor market (U.S. Bureau of Labor Statistics, 2017), scaling by this value still suggests a labor market return of approximately $14, an order of magnitude higher than our estimated values of average WTP.

This calculation using Behrman and Rosenzweig (2004)'s estimates relies on a number of assumptions that are unlikely to hold in practice. Chief among these is that the returns to birth weight are stable over the life course, and salary and labor market participation rates are also stable over the life course. Still, we believe this is an informative estimate, if only because the $14 per gram is close to the public WTP inferred from WIC and FSP ($11-$18 per gram), but 8-13 times larger than the private average WTP estimated among our respondents.

[24]We calculate the present value as

$$PVBW = (\$30,240 \times 0.07) \times \frac{1-(1+0.05)^{-35}}{0.05} \times \frac{1}{(1+0.05)^{25}} = \$10,235.46$$

5.4 Adjusting our experimental estimates of WTP for low birth weight omission

Inherent in the design of the DCE was the decision to focus only on the WTP for birth weight over the normal range of weights of 2,500-4,000 grams. It is pertinent to ask, then, whether the lower private valuation of birth weight owes to the fact that the WTP for this omitted weight range is unaccounted for, and particularly important. While the experimental design precludes the direct estimation of this WTP, we can estimate the minimum WTP over this range which would equate our private WTP to the range of estimates [$11, $18] from public programs and the labor market. Using the relative frequency of low birth weight and normal birth weight babies from the population of US births (see Appendix Figure A4), we estimate that in order for the average WTP to be between $11 and $18 per gram in the full population, while only being $1.44 in the normal birth range, the average WTP for the marginal gram of birth over the range of 500 to 2,500 grams would need to be between $111.3 and $191.8 per gram.[25] This is between a 77 and a 133 fold increase over our experimental estimates, and considerably more than the estimated hospital costs associated with LBW, which Almond, Chay and Lee (2005) place at $4.93 per gram. If we take the estimate by Almond, Chay and Lee (2005) as the average valuation of *low* birth weights, our adjusted average WTP for the whole range becomes $1.74 per gram.[26] This is still a very small magnitude compared to value of birth weight inferred from public programs and the labor market. In the next section, we provide a conceptual framework to shed light on such a discrepancy.

6 Understanding our Estimates

In this section we present a very stylized framework to understand (a) why the estimated average WTP for birth weight is much lower than the returns to birth weight in the labor

[25]The first value is calculated as follows:

$$\begin{aligned} WTP_{500-4000g} &= WTP_{500-2500g} \times \Pr(Birth_{500-2500g}) + WTP_{2500-4000g} \times \Pr(Birth_{2500-4000g}) \\ \$11 &= WTP_{500-2500g} \times 0.087 + \$1.44 \times 0.913 \\ \Rightarrow WTP_{500-2500g} &= \$111.3. \end{aligned}$$

And similarly for the second value. Even if we were not to take the relative weights of each types of births on the population (i.e., treating each group as of equal size), the average WTP would still be between $20.6 and $34.6 per gram in the 500-2,500 gram group as in the 2,500-4,000 gram group to drive average parental WTP to $11 and $18, respectively. In all of these calculations we ignore births which are over 4,000 grams, however note that this is a conservative decision, given that our non-linear estimates suggest that marginal WTP birth weight above 4,000 grams is most likely negative.

[26]This value is calculated as follows: $\$4.93 \times 0.087 + \$1.44 \times 0.913$.

market and (b) why the estimated WTP for birth weight varies by parental status.

6.1 A Parsimonious Model of the WTP for Birth Weight

Individuals in our sample are (i) parents, (ii) non-parents who are planning to have children, and (iii) non-parents who are not planning to have children. Suppose that individual preferences can be represented by:

$$U_i^{NP} = \gamma \log(X_i) \;\text{ if individual } i \text{ is a non-parent} \tag{3}$$

and

$$U_i^{P} = \gamma \log(X_i) + (1-\gamma)\log(Y_i^C) \;\text{ if individual } i \text{ is a parent} \tag{4}$$

where X_i is individual consumption, Y_i^C is the income of the individual i's child as an adult, and $0 < (1-\gamma) < 1$ is the altruism parameter. If individuals behave as von Neumann-Morgenstern expected-utility maximizers, then the expected utility of individual i can be written as:

$$EU_i = \pi_i U_i^P + (1-\pi_i) U_i^{NP} \tag{5}$$

where $\pi_i \in [0,1]$ is the probability of becoming a parent.

Finally, assume that the income of the individual i's child as an adult is generated by:

$$\log(Y_i^C) = \alpha + \beta \log(BW_i^C) + u_i \tag{6}$$

where β is the return to birth weight (BW) in the labor market and u_i is a random shock.

Given (3)-(6), the willingness to pay for BW for individual i is given by:

$$WTP_{BW_i^C} = \left(\frac{1-\gamma}{\gamma}\right)\pi_i \beta \left(\frac{X_i}{BW_i^C}\right) \tag{7}$$

That is, even if birth weight does not enter the utility function directly, it affects individual utility indirectly through its impact on the income of the individual i's child as an adult. *Ceteris paribus*, this simple model has three main predictions:

1. The higher is the probability of becoming a parent π_i, the higher will be the WTP for

birth weight.

2. The higher is the return to birth weight β, the higher will be the WTP for birth weight.

3. The higher is altruism $(1-\gamma)$, the higher will be the WTP for birth weight.

Let π_i vary with parental status, so that we have three π's: $\pi^P = 1$, $\pi^{NP,P}$, and $\pi^{NP,NP}$. It is natural to assume that $\pi^P = 1 > \pi^{NP,P} > \pi^{NP,NP} > 0$, so that among non-parents who state not to be planning to have children, we allow for their probability of having children ($\pi^{NP,NP}$) to be positive (e.g., failure in the contraceptive method). Thus, $WTP^{P}_{BW_i^C} > WTP^{NP,P}_{BW_i^C} > WTP^{NP,NP}_{BW_i^C}$.

Our parsimonious model shows that the differences in π (the probability of having children) by parental status may explain the differences in WTP for birth weight by parental status, even assuming that individuals have *both* the same preferences (in terms of altruism towards children) and the same information (in terms of knowledge about the returns to birth weight in the labor market). In addition, the fact that the WTP for birth weight among parents is much smaller than the returns to birth weight in the labor market can be explained by lack of information on β, the return to birth weight in the labor market. Indeed, that parents have biased beliefs on the returns to human capital investments is consistent with recent work by Cunha, Elo and Culhane (2013) and Boneva and Rauh (2016). Moreover, it would be counterfactual to explain the gap between the returns to birth weight in the labor market and parental WTP for birth weight by arguing that parents have a low $(1-\gamma)$. If anything, parents should be more altruistic towards children (i.e., have larger values of $(1-\gamma)$) than non-parents.[27]

If individuals, and even parents, under-estimate the returns to birth weight, there is scope for a public policy intervention. In particular, if parents are not aware of the magnitude of the returns to birth weight, they may be investing sub-optimally in pre-natal behaviors.[28] In an influential paper, Jensen (2010) demonstrates that correcting expectations regarding the returns to education increases investments in education. Our results potentially open the door to a similar phenomenon in early life health. Making available, through informational campaigns, the existing knowledge on the influence of health at birth on long-term outcomes,

[27] For instance, using experimental economics methods, Peters et al. (2004) show that parents contributed more to a public good when in groups with family members than when in groups with strangers.

[28] Combining the individual budget constraint with a production function for birth weight, it is straightforward to show that the higher is the degree of underestimation of β, the lower will be the optimal investment in birth weight (from the individual point of view).

is possibly a type of intervention that may address sub-optimal parental investment behavior in early life health. Our finding and policy recommendation echo the whole research agenda on the early origins of inequality and its lifetime consequences.

6.2 Allowing for Preference Heterogeneity

Our empirical analysis has used a traditional logit model, which assumes that the parameter associated with each birth attribute is *fixed* across individuals. In our case, this is tantamount to assuming *homogeneous* preferences over birth outcomes between individuals. However, as emphasized by our parsimonious model (equation (7)), the WTP for birth weight may vary by individual depending on the altruism parameter, the perceived return to birth weight, the probability of becoming a parent, and the ratio of individual consumption on child's birth weight. In this section we allow for preference *heterogeneity* in birth outcomes during our estimation process by specifying a *mixed* logit model.

The mixed logit model (Revelt and Train (1998), McFadden and Train (2000) and Train (2003)), also known as the random-parameters logit, offers a number of benefits for discrete choice experiments (DCEs) such as ours. Firstly, in a mixed logit the parameter associated with each component is allowed to be randomly (normally) distributed across respondents, implying that both the mean and the standard deviation of these parameters can be estimated. Secondly, the mixed logit permits efficient estimation when a given respondent makes repeated choices, precisely as in the DCEs examined here. These models explicitly account for the correlation in *unobserved* utility within an individual respondent between choices (Revelt and Train, 1998).[29]

We thus loosen specification (1), where rather than estimating this model using a standard logit, we estimate it using the mixed logit. This procedure requires the use of a maximum simulated likelihood in place of maximum likelihood, however is now available in many standard software packages.[30] The parameter vector now consists of each individual's specific parameters, which give rise to the mean parameter in the sample as well as measures of its variance.

In Table 5 we display the parameters estimated from the mixed logit, as well as the WTP for birth weight using the full sample and each sub-sample of interest. As is common in discrete

[29]The mixed logit does not exhibit the restrictive independence from irrelevant alternatives (IIA) property inherent in the logit model.

[30]See for example Hole (2007*b*) for a Stata implementation, or a series of packages made available by Kenneth Train in other languages (`https://eml.berkeley.edu/~train/software.html`).

choice applications with willingness to pay, we model the price (out of pocket expenses) as *fixed* across respondents, while allowing all other coefficients (and preferences) to *vary*. This ensures that the WTP for each attribute is identified, as outlined in Revelt and Train (1998). In panel A we display the mean estimates for each parameter, and in panel B the standard deviation of each parameter. As is typical with the mixed logit, the normalization of the parameters with respect to individual utility means that point estimates are significantly larger than those in the standard logit model. Nevertheless, we are more interested in the WTP of each parameter (as well as the distribution of parameters in the sample) rather than each parameter itself. *On average*, the WTP for birth weight is quite similar to that estimated in the standard logit model. For the full sample the WTP from the Mixed Logit model is \$1.68 per gram (95% CI: \$1.38-\$1.99). Similarly, we observe that this WTP is highest for parents at \$2.01 per gram (95% CI: \$1.55-\$2.45), followed by non-parents who plan to have a birth (\$1.57 per gram, 95% CI: \$0.98-\$2.16) and the lowest among non-parents who do not plan to have children (\$1.29 per gram, 95% CI: \$0.75-\$1.82). However, beyond the means of these estimated parameters, here we are also interested in their standard deviations, displayed in panel B. These standard deviations are at times quite large, consistent with substantial heterogeneity in tastes within the sample, particularly for the birth weight and child's gender. Such heterogeneity may reflect the non-linear valuation of birth weight and the asymmetric valuation of child's gender depending on respondent's gender.

Using both of these sets of parameters (mean and standard deviation), we are also able to determine the proportion of all respondents who positively value birth weight (and indeed any characteristic) in these linear specifications.[31] These values are displayed at the base of the table, indicating what proportion of respondents positively value birth weight. These values follow a similar pattern as those observed for WTP. Namely, parents are the most likely to place a positive value on birth weight (72.1% of respondents), followed by non-parents who plan to have children (69.6%), and by non-parents who do not plan to have children (68.1%). Using the conditioning of individual taste (COIT) method described in Revelt and Train (2000) we are able to estimate the entire distribution of WTP across respondents, which we display in Figure 6. Once again, this provides evidence of considerable heterogeneity in tastes for birth weight.

Finally, we extend the Mixed Logit to our non-parametric specification where birth weight enters in categories as observed by respondents. The results for the WTP, as well as the percent

[31]These can be calculated using the entire vector of parameters, or alternatively as $100 \times \Phi(-\mu_k/\sigma_k)$, where Φ is the cumulative normal distribution, μ_k is parameter k's mean, and σ_k is its standard deviation.

of respondents who value each birth weight positively, are displayed in Figure 7. These are all based on the mean and standard deviations of the parameters estimated from the mixed logit, as displayed in the footer of Table 5. As is the case with the standard logit, we observe a non-linear pattern in WTP, which is highest for moving from 2,500 grams to 3,400 grams. In turning to the proportion of respondents who positively value each birth weight category, we observe that this quickly rises as birth weights diverge from the baseline reference category. Once reaching approximately 2,950 grams, over 90% of all respondents prefer this to the baseline value of 2,500 grams, and this value rises to close to 100% once exceeding approximately 3,250 grams. We provide similar graphs for each of the groups displayed in Table 5 in Figure 8.

These mixed logit estimates provide additional evidence that our measures of the average private WTP are definitely much lower than the public values, or the associated labor market returns. All in all, our experimental private WTP estimation is robust to allowing for preference heterogeneity.

7 Conclusion

The use of birth weight as an individual's prominent measure of early-life endowment of human capital is now a well established practice in the economic literature. Birth weight has increasingly been shown to be a modifiable outcome, being particularly responsive to certain policy measures. Despite considerable public investment in policies to increase birth weight and health at birth, very little is known about the *private* willingness to pay for birth weight.

In this paper we fill this gap in the literature. We document that, firstly, individuals have a positive, economically and statistically significant WTP for birth weight. We demonstrate that this WTP is higher among parents than non-parents, and higher among non-parents that plan to have children than among non-parents who do not plan to have children. Secondly, we document that the WTP follows a hump shape. Among all respondents the average WTP for a gram of birth weight is estimated at $1.44, however this is as high as $12.49 at the lower range of weights examined and as low as -$4.27 at the upper end of the weight range. Similar values for parents are $1.72 (on average) with a range of nearly $15 for movements at the lower end to -$7.58 at the upper end of the birth weight distribution.

These values, while of interest in their own right, are more relevant when compared with a number of well-documented benchmark figures in the economic literature. We demonstrate that on average the *private* willingness to pay is considerably lower than the *public* WTP (estimated at between \$11 and \$18 per gram), and also considerably lower than the expected returns to birth weight for a US-born baby based on *labor market returns* (estimated at around \$14 per gram). While this is puzzling, we provide a parsimonious model which is able to explain the different WTP by parental status and the discrepancy between our estimated private valuation and the returns in the labor market. Our model highlights the possibility that even parents may under-appreciate the value of birth weight, echoing recent research on the importance of parental beliefs about the production of human capital (Cunha, Elo and Culhane, 2013; Boneva and Rauh, 2016). As Jensen (2010) documents with education, if parents and those planning to become parents do not appreciate the full value of birth weight, this opens the door to new, and perhaps comparatively cheap, policy options to improve birth weight (and early life health in general) using informational campaigns.

References

Agee, Mark D., and Thomas D. Crocker. 2008. "Does parents' valuation of children's health mimic their valuation of own health?" *Journal of Population Economics*, 21(1): 231–249.

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy*, 114(4): 672–712.

Almond, Douglas, and Bhashkar Mazumder. 2005. "The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data." *American Economic Review*, 95(2): 258–262.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives*, 25(3): 153–172.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *The Review of Economics and Statistics*, 93(2): 387–403.

Almond, Douglas, Janet Currie, and Valentina Duque. 2017. "Childhood Circumstances and Adult Outcomes: Act II." National Bureau of Economic Research Working Paper 23017.

Almond, Douglas, Kenneth Y. Chay, and David S. Lee. 2005. "The Costs of Low Birth Weight." *The Quarterly Journal of Economics*, 120(3): 1031–1083.

Almond, Douglas, Lena Edlund, and Marten Palme. 2009. "Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden." *Quarterly Journal of Economics*, 124(4): 1729–1772.

Barker, D J. 1990. "The fetal and infant origins of adult disease." *British Medical Journal*, 301(6761): 1111.

Barker, D J. 1995. "Fetal origins of coronary heart disease." *British Medical Journal*, 311(6998): 171–174.

Barker, D J, P Winter, C Osmond, B Margetts, and S Simmonds. 1989. "Weight in Infancy and Death from Ischaemic Heart Disease." *The Lancet*, 334(8663): 577–580.

Behrman, Jere R, and Mark R Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics*, 86(2): 586–601.

Bekker-Grob, E W, Mandy Ryan, and Karen Gerard. 2012. "Discrete Choice Experiments in Health Economics: A Review of the Literature." *Health Economics*, 21(1): 145–172.

Ben-Shalom, Yonatan, Robert A. Moffitt, and John Karl Scholz. 2011. "An Assessment of the Effectiveness of Anti-Poverty Programs in the United States." National Bureau of Economic Research, Inc NBER Working Papers 17042.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis*, 20(3): 351–368.

Bharadwaj, Prashant, J Eberhard, and C Neilson. forthcoming. "Do Initial Endowments Matter only Initially? Birth Weight, Parental Investments and Academic Achievement." *Journal of Labor Economics.*

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson. 2013. "Early Life Health Interventions and Academic Achievement." *American Economic Review*, 103(5): 1862–1891.

Bitler, Marianne P., and Lynn A. Karoly. 2015. "Intended And Unintended Effects Of The War On Poverty: What Research Tells Us And Implications For Policy." *Journal of Policy Analysis and Management*, 34(3): 639–696.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *The Quarterly Journal of Economics*, 122(1): 409–439.

Boneva, Teodora, and Christopher Rauh. 2016. "Human Capital Production and Parental Beliefs." University of Cambridge mimeo.

Bozzoli, Carlos, and Climent Quintana-Domeque. 2014. "The Weight of the Crisis: Evidence from Newborns in Argentina." *The Review of Economics and Statistics*, 96(3): 550–562.

Bridges, John F.P., A. Brett Hauber, Deborah Marshall, Andrew Lloyd, Lisa A. Prosser, Dean A. Regier, F. Reed Johnson, Josephine Mauskopf, Mandy Ryan, and Shelley Farrar. 2011. "Conjoint Analysis Applications in Health—a Checklist: A

Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force." *Value in Health*, 14(4): 403–413.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The lasting impact of childhood health and circumstance." *Journal of Health Economics*, 24(2): 365–389.

Chevalier, Arnaud, and Vincent O'Sullivan. 2007. "Mother's Education and Birth Weight." Institute for the Study of Labor (IZA) IZA Discussion Papers 2640.

Clarke, Damian, Sonia Oreffice, and Climent Quintana-Domeque. 2016. "The Demand for Season of Birth." Human Capital and Economic Opportunity Working Group Working Papers 2016-032.

Conley, Dalton, and Neil G. Bennett. 2000. "Is Biology Destiny? Birth Weight and Life Chances." *American Sociological Review*, 65(3): 458–467.

Conley, Dalton, Kate Strully, and Neil G Bennet. 2003. "A pound of flesh or just proxy? Using twin differences to estimate the effect of birth weight on life chances." National Bureau of Economic Research, Inc NBER Working Papers 9901.

Cook, C. Justin, and Jason M Fletcher. 2015. "Understanding heterogeneity in the effects of birth weight on adult cognition and wages." *Journal of Health Economics*, 41(3): 107–116.

Cunha, Flávio, Irma Elo, and Jennifer Culhane. 2013. "Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation." National Bureau of Economic Research, Inc NBER Working Papers 19144.

Currie, Janet, and Enrico Moretti. 2007. "Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight." *Journal of Labor Economics*, 25: 231–264.

Currie, Janet, and Rosemary Hyson. 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight." *The American Economic Review*, 89(2): 245–250.

Dahl, Gordon B, and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies*, 75(4): 1085–1120.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2014. "The Effects of Poor Neonatal Health on Children's Cognitive Development." *American Economic Review*, 104(12): 3921–55.

Fletcher, Jason M. 2011. "The medium term schooling and health effects of low birth weight: Evidence from siblings." *Economics of Education Review*, 30(3): 517–527.

Francis-Tan, Andrew, and Hugo M. Mialon. 2015. ""A Diamond Is Forever" And Other Fairy Tales: The Relationship Between Wedding Expenses And Marriage Duration." *Economic Inquiry*, 53(4): 1919–1930.

Gupta, Nabanita Datta, Mette Deding, and Mette Lausten. 2013. "The effect of low birth weight on height, weight and behavioral outcomes in the medium-run." *Economics & Human Biology*, 11(1): 42–55.

Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2013. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis*, 22(1): 1–30.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating vignette and conjoint survey experiments against real-world behavior." *Proceedings of the National Academy of Sciences*, 112(8): 2395–2400.

Hole, A R. 2007*a*. "A comparison of approaches to estimating confidence intervals for willingness to pay measures." *Health Economics*, 8(16): 827–840.

Hole, A R. 2007*b*. "Fitting mixed logit models by using maximum simulated likelihood." *The Stata Journal*, 3(7): 388–401.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association*, 81(396): 945–960.

Hoynes, Hilary, Marianne Page, and Ann Huff Stevens. 2011. "Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program." *Journal of Public Economics*, 95(3): 813–827.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand or Schooling." *Quarterly Journal of Economics*, 125(5): 515–548.

Johnson, Bryan, Betsy Thorn, Brittany McGill, Alexandra Suchman, Michele Mendelson, Kelly Lawrence Patlan, Brian Freeman, Rebecca Gotlieb, and Patty Connor. 2013. "WIC Participant and Program Characteristics 2012 Final Report." U.S. Department of Agriculture, Food and Nutrition Service Prepared by Insight Policy Research under Contract No. AG-3198-C-11-0010 Alexandria, VA.

Johnson, Rucker C, and Robert F Schoeni. 2011*a*. "Early-Life Origins of Adult Disease: National Longitudinal Population-Based Study of the United States." *American Journal of Public Health*, 101(12): 2317–2324.

Johnson, Rucker C, and Robert F Schoeni. 2011*b*. "The Influence of Early-Life Events on Human Capital, Health Status, and Labor Market Outcomes Over the Life Course." *The B.E. Journal of Economic Analysis & Policy*, 11(3): 1–57.

Jordan, Jillian J., Moshe Hoffman, Paul Bloom, and David G. Rand. 2016. "Third-party punishment as a costly signal of trustworthiness." *Nature*, 530(7591): 473–476.

King, Madeleine T., Jane Hall, Emily Lancsar, Denzil Fiebig, Ishrat Hossain, Jordan Louviere, Helen K. Reddel, and Christine R. Jenkins. 2007. "Patient preferences for managing asthma: results from a discrete choice experiment." *Health Economics*, 16(7): 703–717.

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review*, 105(4): 1478–1508.

Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy*, 74(2): 132–157.

Lin, Ming-Jen, and Jin-Tan Liu. 2009. "Do lower birth weight babies have lower grades? Twin fixed effect and instrumental variable method evidence from Taiwan." *Social Science & Medicine*, 68(10): 1780–1787.

Marshall, Deborah, John F. P. Bridges, Brett Hauber, Ruthanne Cameron, Lauren Donnalley, Ken Fyie, and F. Reed Johnson. 2010. "Conjoint Analysis Applications in Health — How are Studies being Designed and Reported?" *The Patient: Patient-Centered Outcomes Research*, 3(4): 249–256.

McFadden, Daniel, and Kenneth Train. 2000. "Mixed MNL models for discrete response." *Journal of Applied Econometrics*, 15(5): 447–470.

Oreopoulos, Philip, Mark Stabile, Randy Walld, and Leslie L. Roos. 2008. "Short-, Medium-, and Long-Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins." *Journal of Human Resources*, 43(1): 88–138.

Peters, H. Elizabeth, A. Sinan Ünür, Jeremy Clark, and William D. Schulze. 2004. "Free-Riding and the Provision of Public Goods in the Family: A Laboratory Experiment." *International Economic Review*, 45(1): 283–299.

Propper, Carol. 1990. "Contingent valuation of time spent on NHS waiting list." *Economic Journal*, 100(400): 193–199.

Propper, Carol. 1995. "The Disutility of Time Spent on the United Kingdom's National Health Service Waiting Lists." *Journal of Human Resources*, 30(4): 677–700.

Revelt, David, and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *The Review of Economics and Statistics*, 80(4): 647–657.

Revelt, David, and Kenneth Train. 2000. "Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier." University of California at Berkeley Economics Working Papers E00-274.

Rosenzweig, Mark R., and Junsen Zhang. 2013. "Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of Chinese twins." *Journal of Development Economics*, 104(C): 245–260.

Rosenzweig, Mark R, and T Paul Schultz. 1983. "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight." *Journal of Political Economy*, 91(5): 723–746.

Rossin-Slater, Maya. 2013. "WIC in your neighborhood: New evidence on the impacts of geographic access to clinics." *Journal of Public Economics*, 102(3): 51–69.

Royer, Heather. 2009. "Separated at Girth: US Twin Estimates of the Effects of Birth Weight." *American Economic Journal: Applied Economics*, 1(1): 49–85.

Ryan, Mandy, and Jenny Hughes. 1997. "Using Conjoint Analysis to Assess Women's Preferences for Miscarriage Management." *Health Economics*, 6(3): 261–273.

Ryan, Mandy, and Shelley Farrar. 2000. "Using conjoint analysis to elicit preferences for health care." *British Medical Journal*, 320(7248): 1530–1533.

Schwarz, Laura C. 2004. "The Dutch Hunger Winter and the developmental origins of health and disease." *Proceedings of the National Academy of Sciences*, 107(39): 16757–16758.

Sexton, M, and J R Hebel. 1984. “A clinical trial of change in maternal smoking and its effects on birth weights.” *Journal of the American Medical Association*, 251(7): 911–915.

Stephenson, T, and M E Symonds. 2002. “Maternal nutrition as a determinant of birth weight.” *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 86(1): F4–F6.

Torche, Florencia, and Ghislaine Echevarría. 2011. “The effect of birthweight on childhood cognitive development in a middle-income country.” *International Journal of Epidemiology*, 40(4): 1008–1018.

Train, Kenneth. 2003. *Discrete Choice Methods with Simulation.* Cambridge University Press.

United States Census Bureau. 2015. “State Population Totals Datasets: 2010-2015.” *https://www2.census.gov/programs-surveys/popest/datasets/2010-2015/state/asrh/scprc-est2015-18-pop-res.csv*, Accessed: 2017-01-26.

United States Department of Agriculture. 2014. “Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2013.” The Office of Policy Support Nutrition Assistance Program Report Series SNAP-14-CHAR.

U.S. Bureau of Labor Statistics. 2017. “Civilian Employment-Population Ratio [EMRATIO].” *retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/EMRATIO*, Accessed: 2017-01-26.

US Census Bureau. 2016. “PINC-01 Selected Characteristics of People 15 Years and Over, by Total Money Income, Work Experience, Race, Hispanic Origin, and Sex.” *http://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc/pinc-01.html*, Retrieved 25 January, 2017.

van den Berg, Gerard J., Maarten Lindeboom, and France Portrait. 2006. “Economic Conditions Early in Life and Individual Mortality.” *American Economic Review*, 96(1): 290–302.

Veldwijk, Jorien, Mattijs S. Lambooij, Esther W. de Bekker-Grob, H A. Smit, and G. Ardine de Wit. 2014. “The Effect of Including an Opt-Out Option in Discrete Choice Experiments.” *PLOS ONE*, 9(11): 1–9.

Wilcox, Allen J. 2001. “On the importance—and the unimportance—of birthweight.” *International Journal of Epidemiology*, 30(6): 1233–1241.

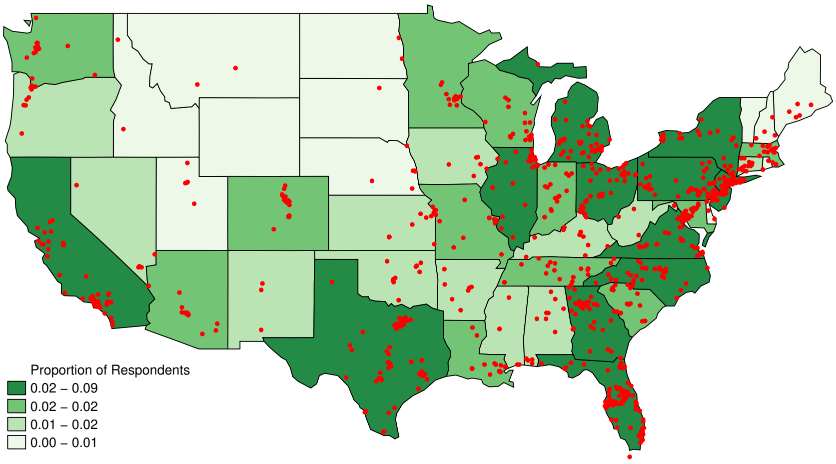
Wittink, Marsha N., Mark Cary, Thomas Tenhave, Jonathan Baron, and Joseph J. Gallo. 2010. "Towards Patient-Centered Care for Depression." *The Patient: Patient-Centered Outcomes Research*, 3(3): 145–157.

Zizzo, Daniel John. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics*, 13(1): 75–98.

Zweifel, Peter, Friedrich Breyer, and Mathias Kifmann. 2009. *Health Economics.* Springer.

Figures and Tables

Figure 1: Geographic Coverage of Respondents



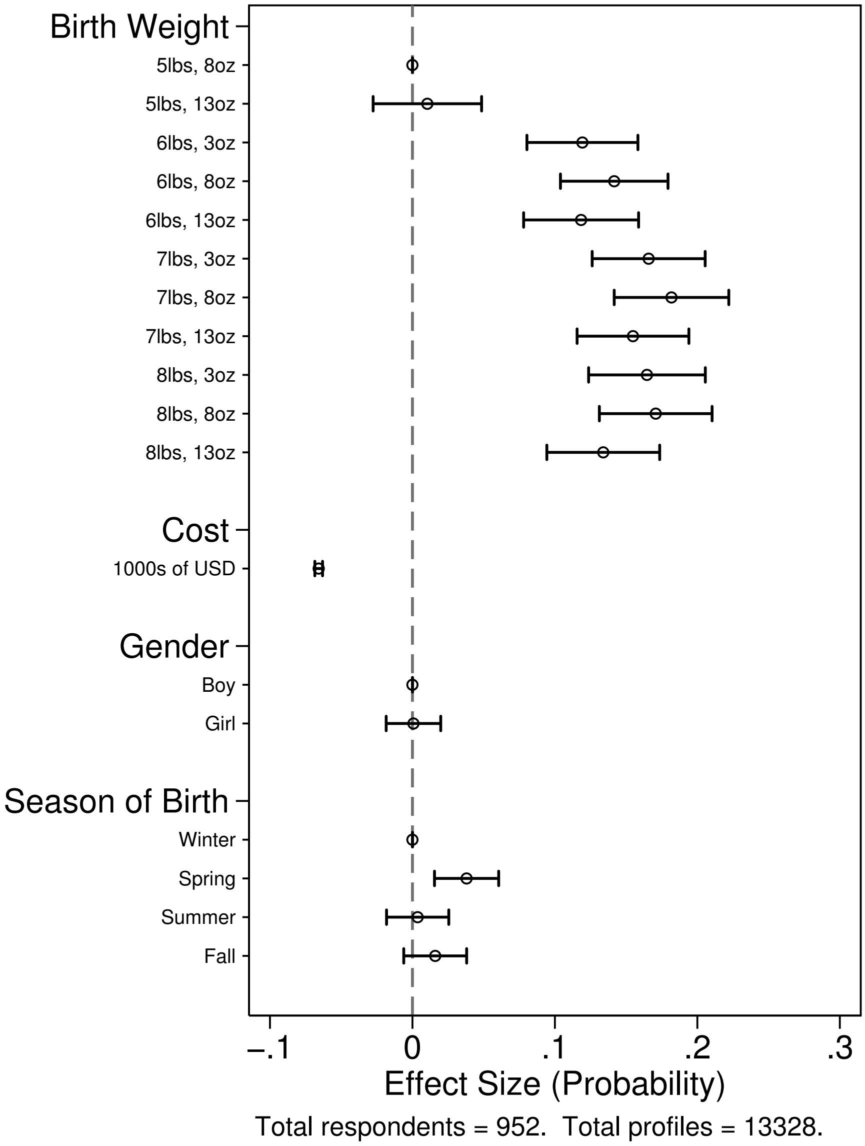
NOTES: The survey sample consists of 1,002 respondents. The final estimation sample consists of 952 respondents given that it removes respondents whose geographic IP suggested a non-US location (36 respondents, 3.6%), those who failed to respond that their educational attainment was identical at the beginning and end of the survey (8 respondents, 0.8%), and those who completed the discrete choice experiment in under two minutes (6 respondents, 0.6%).

Table 1: Summary Statistics of Respondents

	N	Mean	Std. Dev.	Min	Max
Female	952	0.56	0.50	0.00	1.00
Age	952	36.10	11.31	18.00	75.00
Black	952	0.08	0.26	0.00	1.00
White	952	0.84	0.37	0.00	1.00
Hispanic	952	0.06	0.23	0.00	1.00
Parent	952	0.51	0.50	0.00	1.00
Non-Parent Planning Children	469	0.45	0.50	0.00	1.00
Number of Children	952	1.07	1.32	0.00	6.00
Married	952	0.45	0.50	0.00	1.00
Employed	952	0.68	0.47	0.00	1.00
Some College +	952	0.89	0.31	0.00	1.00
Years of Education	952	14.60	1.76	8.00	17.00
Total Family Income (1000s)	952	57.17	38.25	5.00	175.00
Hourly earnings on MTurk	952	4.14	2.75	1.50	11.50

NOTES: Refer to Figure 1 for a discussion of the experimental sample. Years of education, total income and hourly MTurk earnings are calculated from categorical variables. Non-Parent Planning Children refers to any respondent who either answers that they are pregnant or plan to have children, and currently have no children.

Figure 2: Discrete Choice Experimental Results



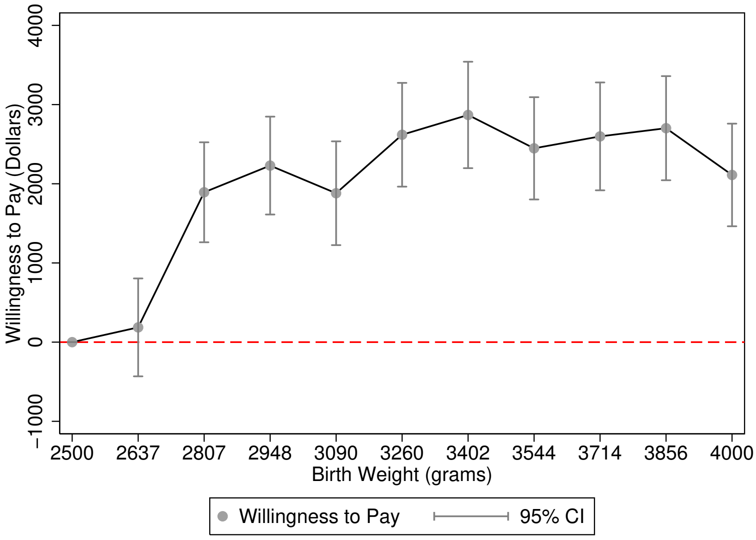
Notes: Point estimates and confidence intervals are displayed of the change in likelihood of choosing a birth profile given that a particular characteristic was seen. Each characteristic is compared to the omitted base case indicated on the zero line. Each respondent observes 7 paired birth scenarios, resulting in 14 profiles per respondent. 95% confidence intervals are clustered by respondent, and costs are displayed as a linear coefficient. Fully non-parametric costs are displayed in Appendix Figure A5.

Table 2: Birth Characteristics and Willingness to Pay for Birth Weight

	(1) Continuous	(2) Categorical
Birth Weight (in 1000s of grams)	0.091***	
	[0.010]	
Cost (in 1000s of dollars)	-0.063***	-0.063***
	[0.001]	[0.001]
5lbs, 13oz		0.012
		[0.020]
6lbs, 3oz		0.119***
		[0.020]
6lbs, 8oz		0.141***
		[0.019]
6lbs, 13oz		0.119***
		[0.021]
7lbs, 3oz		0.165***
		[0.020]
7lbs, 8oz		0.181***
		[0.021]
7lbs, 13oz		0.154***
		[0.020]
8lbs, 3oz		0.164***
		[0.021]
8lbs, 8oz		0.170***
		[0.020]
8lbs, 13oz		0.133***
		[0.020]
Girl	0.001	0.000
	[0.010]	[0.010]
Spring	0.039***	0.038***
	[0.011]	[0.011]
Summer	0.002	0.003
	[0.011]	[0.011]
Fall	0.015	0.015
	[0.011]	[0.011]
WTP for Birth Weight (1000 grams)	1438.3	
95% CI	[1119.4;1757.1]	
Observations	13328	13328

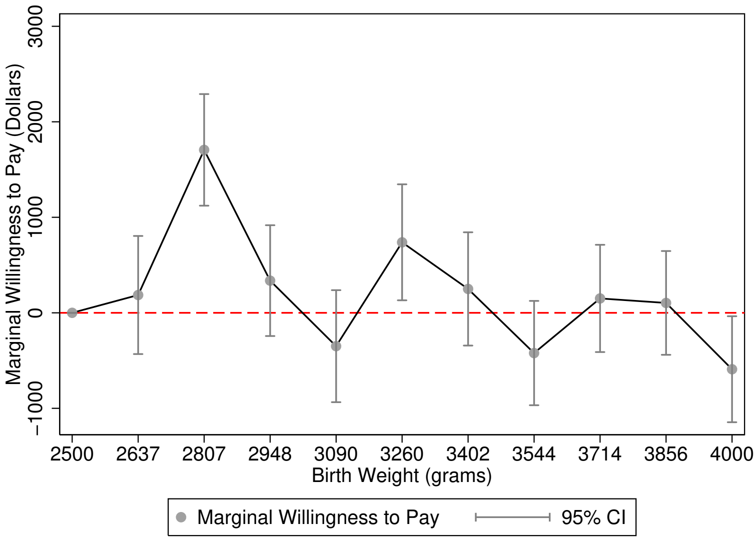
Average marginal effects from a logit regression are displayed. All columns include option order fixed effects and round fixed effects. Standard errors are clustered by respondent. Willingness to pay and its 95% confidence interval is estimated based on the ratio of costs to the probability of choosing a particular birth weight. The 95% confidence interval is calculated using the delta method for the ratio.

Figure 3: Relative Willingness to Pay for Birth Weight



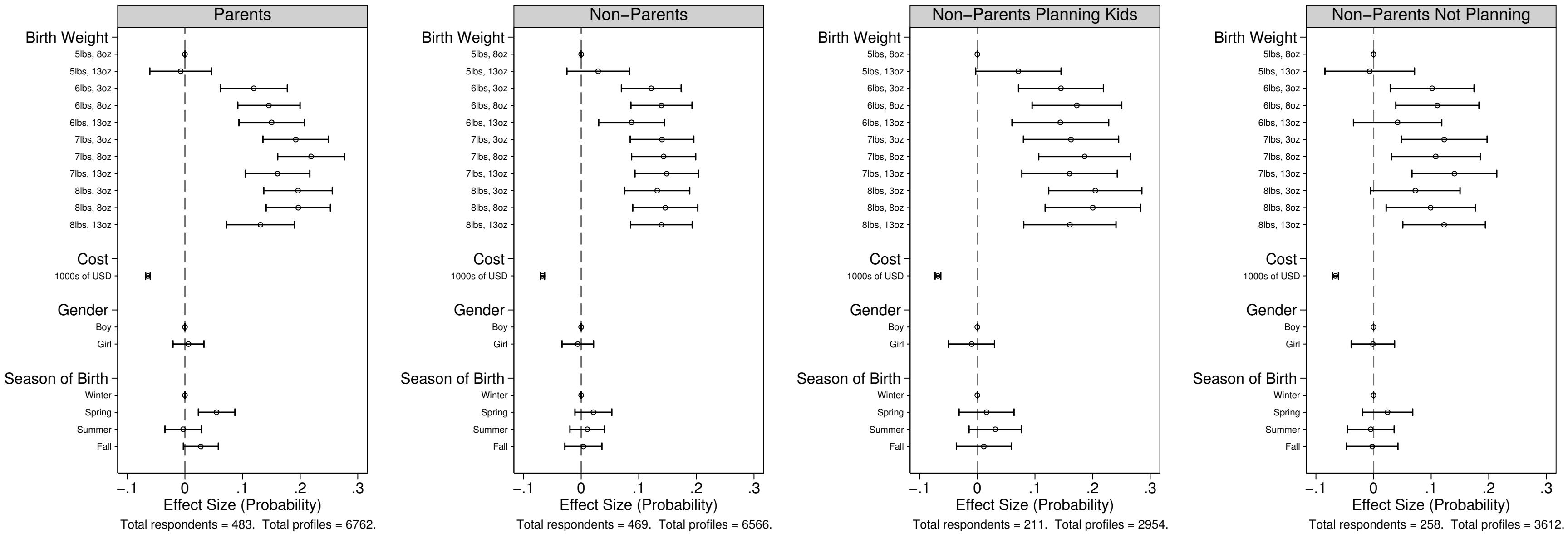
NOTES: Each point and confidence interval are with respect to the baseline (omitted) category of 2,500 grams, the minimum displayed birth weight. Willingness to pay is determined as the ratio between the particular birth weight and out of pocket costs estimated as average marginal effects in a logit regression. 95% confidence intervals displayed are calculated using the delta method.

Figure 4: Marginal Willingness to Pay for Birth Weight



NOTES: Each point and confidence interval compare the willingness to pay for a particular birth weight compared to the preceding birth weight. Willingness to pay is determined as the ratio between the particular birth weight and out of pocket costs estimated as average marginal effects in a logit regression. 95% confidence intervals displayed are calculated using the delta method.

Figure 5: Heterogeneity



NOTES: Methods are identical to those described in notes to Figure 2. The full sample is split by parents or non-parents (panels A and B), and then non-parents are split into those who report planning to have children (or already being pregnant) versus those who do not plan to have children (panels C and D).

Table 3: Birth Characteristics and Willingness to Pay for Birth Weight

	All	Parent		Non-Parents	
	(1)	(2) Yes	(3) No	(4) Planning	(5) Not Planning
Birth Weight (in 1000s of grams)	0.091***	0.106***	0.075***	0.091***	0.064***
	[0.010]	[0.014]	[0.013]	[0.020]	[0.018]
Cost (in 1000s of dollars)	-0.063***	-0.062***	-0.064***	-0.066***	-0.063***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]
Girl	0.001	0.007	-0.006	-0.012	-0.000
	[0.010]	[0.014]	[0.014]	[0.020]	[0.019]
Spring	0.039***	0.055***	0.022	0.018	0.026
	[0.011]	[0.016]	[0.016]	[0.024]	[0.022]
Summer	0.002	-0.004	0.008	0.030	-0.008
	[0.011]	[0.016]	[0.015]	[0.023]	[0.021]
Fall	0.015	0.026*	0.003	0.011	-0.004
	[0.011]	[0.016]	[0.016]	[0.024]	[0.023]
WTP for Birth Weight (1000 grams)	1438.3	1718.4	1172.4	1376.5	1002.5
95% CI	[1119.4;1757.1]	[1232.4;2204.4]	[753.0;1591.8]	[760.6;1992.4]	[433.6;1571.3]
Observations	13328	6762	6566	2954	3612

Average marginal effects from a logit regression are displayed. All columns include option order fixed effects and round fixed effects. Standard errors are clustered by respondent. Willingness to pay and its 95% confidence interval is estimated based on the ratio of costs to the probability of choosing a particular birth weight. The 95% confidence interval is calculated using the delta method for the ratio. Identical regressions with a continuous measure of birth weight are provided in Table A6. Planning and Not Planning in columns 4 and 5 refer to decisions regarding future children as outlined in Table 1.

Table 4: Estimates of the Long Run Returns to Birth Weight in the US

Authors	Weight	Geographic Area	Time Period	Dependent Variable	Estimated Return	Denominator	Estimation Strategy
Panel A: Labor Market							
Behrman and Rosenzweig (2004)	μ = 90.2oz ($\mu = 2,557$)	Minnesota	1936-1955	ln(Wage)	0.190(0.077)[a]	oz/week pregnancy	Between MZ twin
Cook and Fletcher (2015)	$\mu = 3,367$	Wisconsin	1957 HS graduates	ln(Wage)	0.0997(0.0788)	Birth Weight (1 sd)	Between siblings
Johnson and Schoeni (2011*b*)	NA	USA (PSID)	1951-1975	ln(Earnings)	-0.1667(0.097)	LBW	Between siblings (males only)
Panel B: Completed Education							
Royer (2009)	$\mu = 2,533$	California	1960-1982	Completed Education (Years)	0.16(0.07)	1,000g (3500-2500g)	Between twins (females only)
Currie and Moretti (2007)	$\mu = 3,268$	California	1970-1974	Completed Education (Years)	-0.079(0.014)	LBW	Between siblings (females only)
Conley and Bennett (2000)	Pr(LBW) =0.07	USA (PSID)	1968-1973	Timely Graduation	-2.024(0.764)[b]	LBW	Between siblings

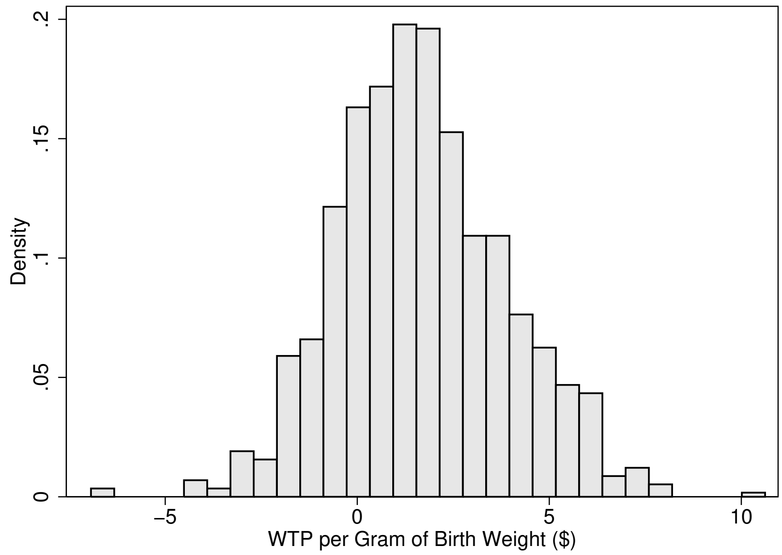
[a] Standard error is calculated based on t-statistic reported in original paper. [b] Results by birth weight groups are presented with respects to >3,500g.

Table 5: Allowing for Preference Heterogeneity with Mixed Logit

	All	Parent		Non-Parents	
		Yes	No	Planning	Not Planning
Panel A: Mean					
Cost (in 1000s of dollars)	-0.581***	-0.569***	-0.601***	-0.688***	-0.552***
	[0.020]	[0.028]	[0.029]	[0.054]	[0.036]
Birth Weight (in 1000s of grams)	0.978***	1.144***	0.821***	1.079***	0.711***
	[0.096]	[0.139]	[0.129]	[0.228]	[0.157]
Fall	0.161*	0.243**	0.106	0.302*	-0.044
	[0.083]	[0.120]	[0.116]	[0.180]	[0.155]
Spring	0.332***	0.380***	0.278**	0.322*	0.224
	[0.082]	[0.115]	[0.119]	[0.184]	[0.158]
Summer	0.046	-0.000	0.122	0.200	0.040
	[0.084]	[0.117]	[0.119]	[0.189]	[0.157]
Girl	-0.001	0.010	-0.053	-0.096	0.002
	[0.088]	[0.123]	[0.125]	[0.210]	[0.155]
Panel A: Standard Deviation					
Birth Weight (in 1000s of grams)	1.911***	1.955***	1.796***	2.102***	1.512***
	[0.141]	[0.192]	[0.185]	[0.296]	[0.247]
Fall	0.250	0.525**	0.070	0.256	0.134
	[0.264]	[0.230]	[0.303]	[0.342]	[0.418]
Spring	0.396**	0.332	0.577***	0.611*	0.662***
	[0.175]	[0.340]	[0.208]	[0.340]	[0.248]
Summer	0.421**	0.434	0.162	0.131	0.365
	[0.202]	[0.439]	[0.334]	[0.414]	[0.538]
Girl	1.943***	1.945***	2.005***	2.299***	1.791***
	[0.126]	[0.173]	[0.189]	[0.327]	[0.241]
WTP for Birth Weight (1000 grams)	1683.5	2008.3	1364.6	1567.4	1286.3
95% CI	[1377.1;1989.8]	[1554.9;2461.7]	[960.5;1768.7]	[975.5;2159.2]	[748.3;1824.3]
% Positively Impacted by Birth Weight	69.5	72.1	67.6	69.6	68.1
Observations	13328	6762	6566	2954	3612

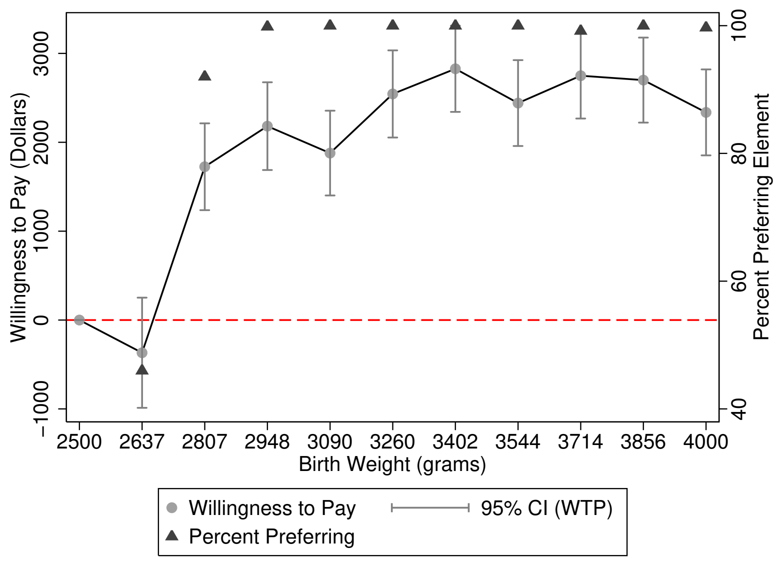
Panel A displays mean coefficients from the mixed logit, and panel B displays the estimated standard deviation of each coefficient. All coefficients with the exception of Cost are allowed to vary randomly throughout the sample. The WTP is calculated as the ratio of the coefficient on birth weight to that on costs, and confidence intervals are calculated by the delta method. The % of respondents who value birth weight positively based on individual coefficients is displayed at the foot of the table. Standard errors are clustered by respondent.

Figure 6: Distribution of Willingness to Pay in the Population



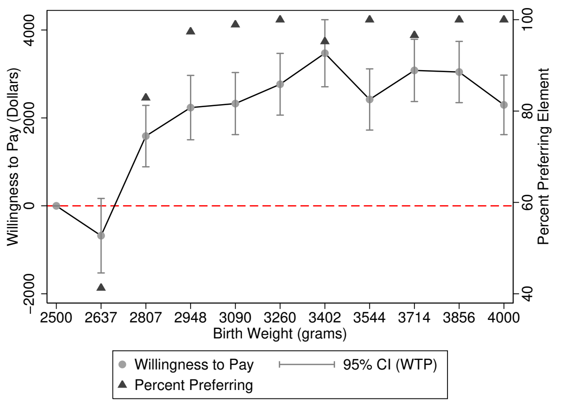
Notes: Distribution of WTP among all respondents is estimated from a mixed logit model using a linear specification for birth weight, and the conditioning of individual taste (COIT) procedure described in Revelt and Train (2000). All respondents are used.

Figure 7: Willingness to Pay and Proportion Positively Valuing Birth Weight (Mixed Logit)

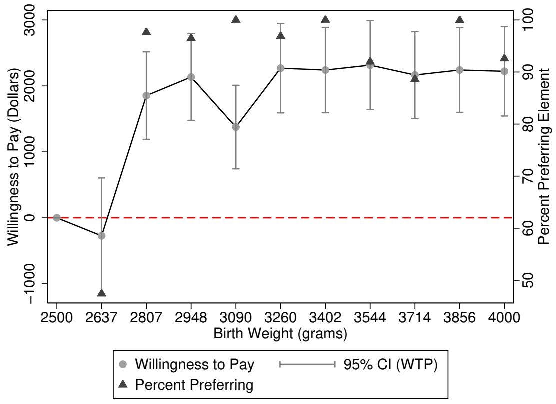


Notes: A mixed logit specification is estimated, however now birth weight is allowed to enter non-parametrically. Willingness to Pay for each component is with respects to the baseline birth weight of 2,500 grams. The Percent Preferring Birth weight refer to the percentage of all respondents who positively value a given weight versus the baseline category. Refer to notes to Table 5 for additional details.

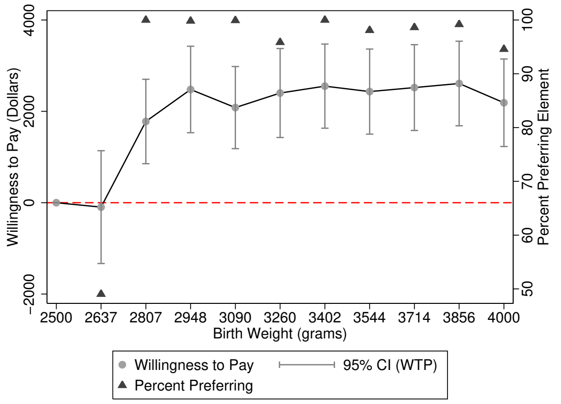
Figure 8: WTP and Proportion Positively Valuing Birth Weight by Group (Mixed Logit)



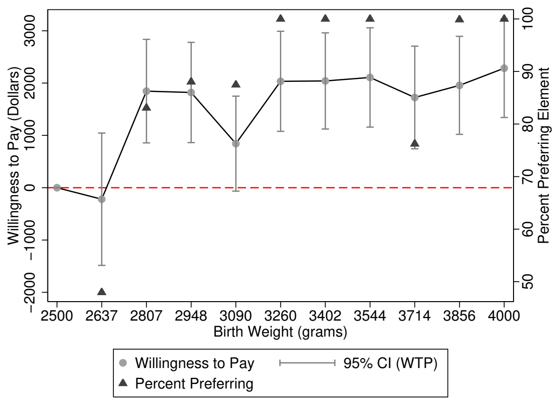
(a) All Parents



(b) All Non-Parents



(c) Non-Parents Planning Children



(d) Non-Parents Not Planning Children

NOTES: Refer to Figure 7 for full notes. Each panel is the output for WTP and percentage positively valuing birth weight as estimated by a Mixed Logit model. Each panel is for a different group displayed in Table 5.

A Appendix Figures and Tables

Table A1: Geographical Coverage

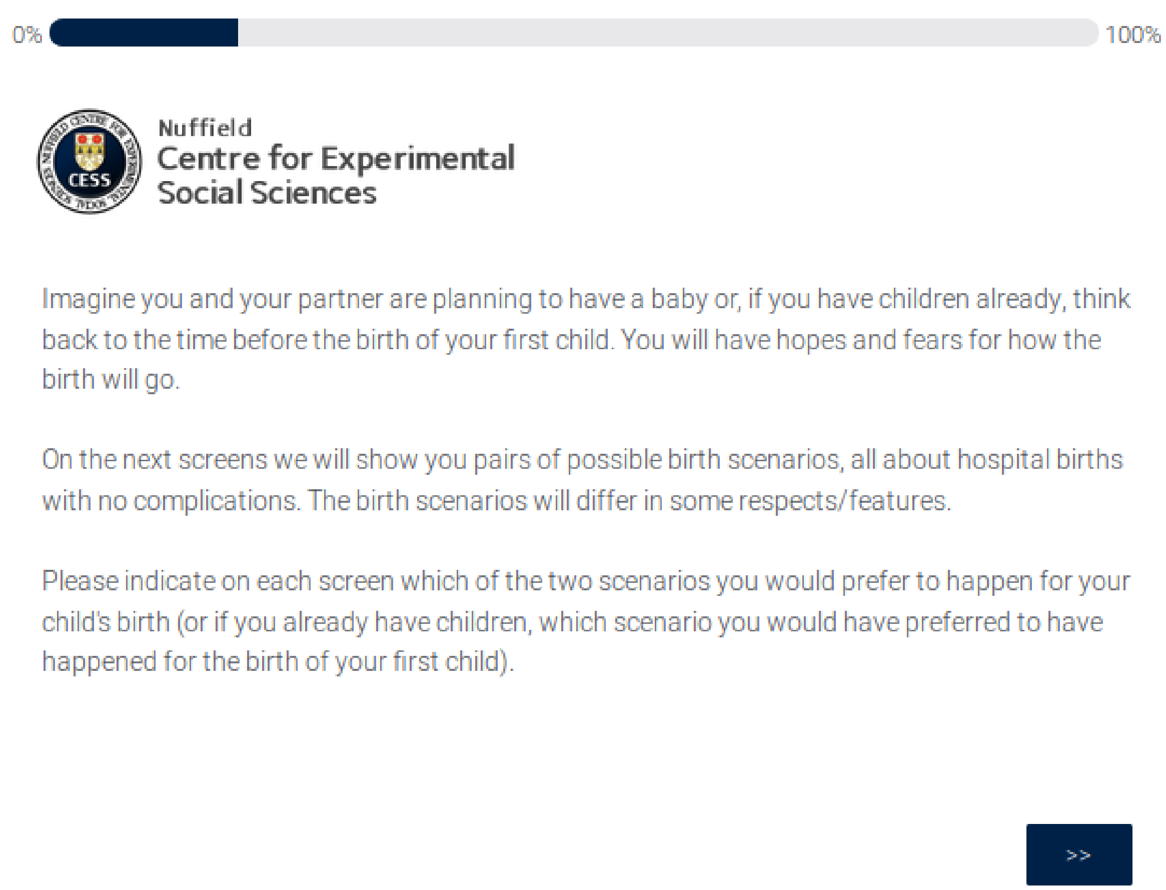
State Name	Percent MTurk	Percent Census Bureau	Difference (%)
Alabama	1.58	1.51	0.06
Alaska	0.11	0.23	-0.12
Arizona	1.79	2.12	-0.34
Arkansas	1.16	0.93	0.23
California	5.04	12.18	-7.14
Colorado	2.31	1.70	0.61
Connecticut	1.47	1.12	0.35
Delaware	0.11	0.29	-0.19
District of Columbia	0.11	0.21	-0.10
Florida	9.45	6.31	3.15
Georgia	4.62	3.18	1.44
Idaho	0.11	0.51	-0.41
Illinois	2.94	4.00	-1.06
Indiana	2.00	2.06	-0.06
Iowa	0.84	0.97	-0.13
Kansas	0.74	0.91	-0.17
Kentucky	1.16	1.38	-0.22
Louisiana	1.68	1.45	0.23
Maine	0.53	0.41	0.11
Maryland	1.79	1.87	-0.08
Massachusetts	2.21	2.11	0.09
Michigan	4.31	3.09	1.22
Minnesota	1.68	1.71	-0.03
Mississippi	0.63	0.93	-0.30
Missouri	1.68	1.89	-0.21
Montana	0.21	0.32	-0.11
Nebraska	0.53	0.59	-0.06
Nevada	0.74	0.90	-0.16
New Hampshire	0.42	0.41	0.01
New Jersey	3.05	2.79	0.26
New Mexico	0.63	0.65	-0.02
New York	6.41	6.16	0.25
North Carolina	3.78	3.12	0.66
North Dakota	0.21	0.24	-0.03
Ohio	4.83	3.61	1.22
Oklahoma	1.05	1.22	-0.17
Oregon	1.37	1.25	0.11
Pennsylvania	5.99	3.98	2.00
Rhode Island	0.32	0.33	-0.01
South Carolina	1.89	1.52	0.37
South Dakota	0.11	0.27	-0.16

Table A1: Geographical Coverage

State Name	Percent MTurk	Percent Census Bureau	Difference (%)
Tennessee	2.31	2.05	0.26
Texas	7.04	8.55	-1.51
Utah	0.53	0.93	-0.41
Virginia	4.41	2.61	1.80
Washington	1.68	2.23	-0.55
West Virginia	0.63	0.57	0.06
Wisconsin	1.89	1.80	0.10
Hawaii	0.00	0.45	-0.45
Vermont	0.00	0.19	-0.19
Wyoming	0.00	0.18	-0.18

NOTES: Columns present the percentage of respondents from the MTurk sample, the percentage of residents according to US Census Bureau records (2015), and the difference between the percentage of MTurk respondents and residents.

Figure A1: Discrete Choice Experiment Framing


0% 100%

Nuffield Centre for Experimental Social Sciences

Imagine you and your partner are planning to have a baby or, if you have children already, think back to the time before the birth of your first child. You will have hopes and fears for how the birth will go.

On the next screens we will show you pairs of possible birth scenarios, all about hospital births with no complications. The birth scenarios will differ in some respects/features.

Please indicate on each screen which of the two scenarios you would prefer to happen for your child's birth (or if you already have children, which scenario you would have preferred to have happened for the birth of your first child).

>>

Figure A2: Discrete Choice Experiment Options

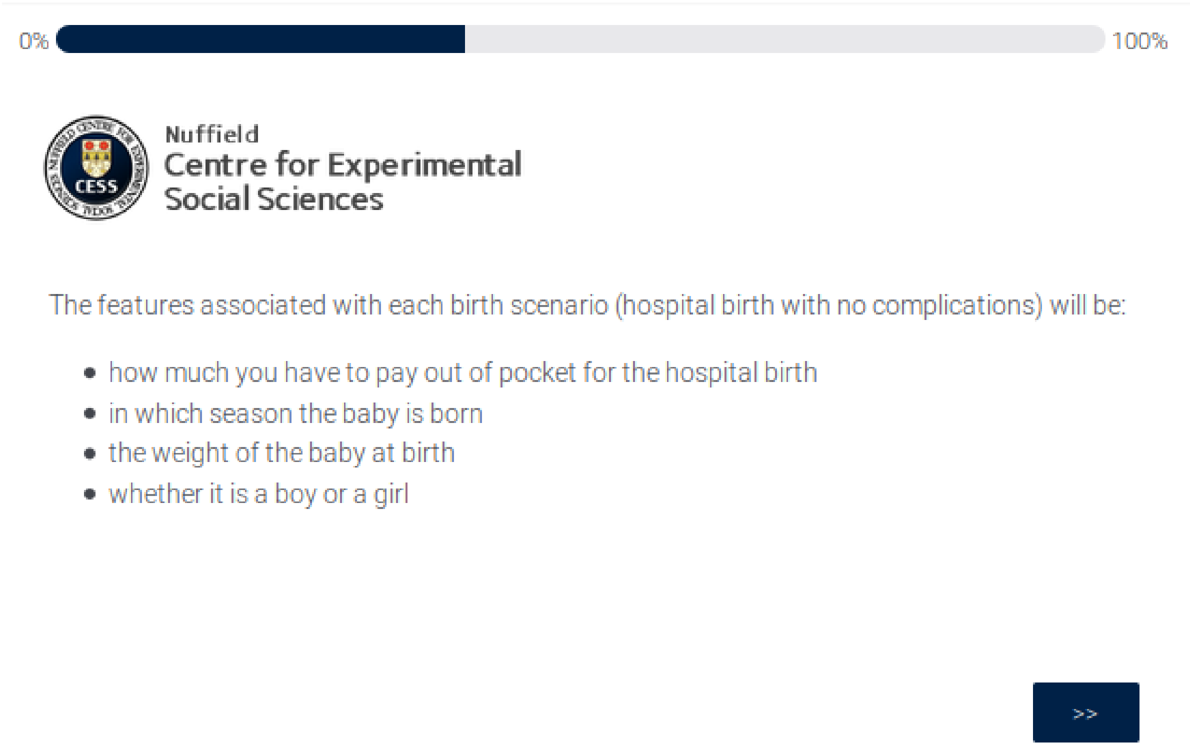


Figure A3: Discrete Choice Experiment Example

0% 100%

Nuffield
Centre for Experimental
Social Sciences

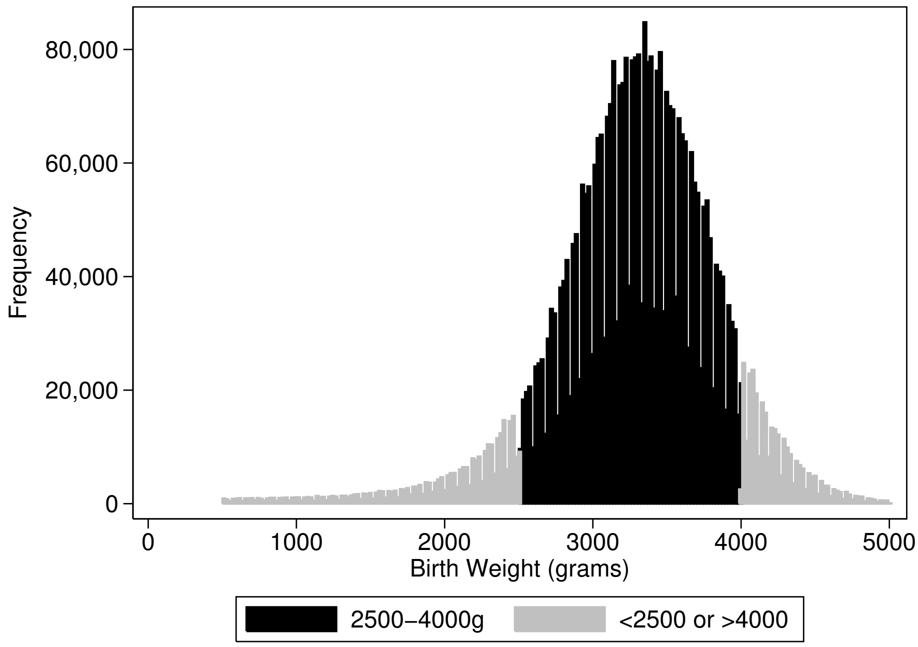
Which of these two birth scenarios would you choose?

	Scenario 1	Scenario 2
Gender	Girl	Boy
Out of Pocket Expenses	$250	$5,000
Birth Weight	7 pounds 8 ounces	6 pounds 13 ounces
Season of Birth	Spring	Winter

Scenario 1 Scenario 2

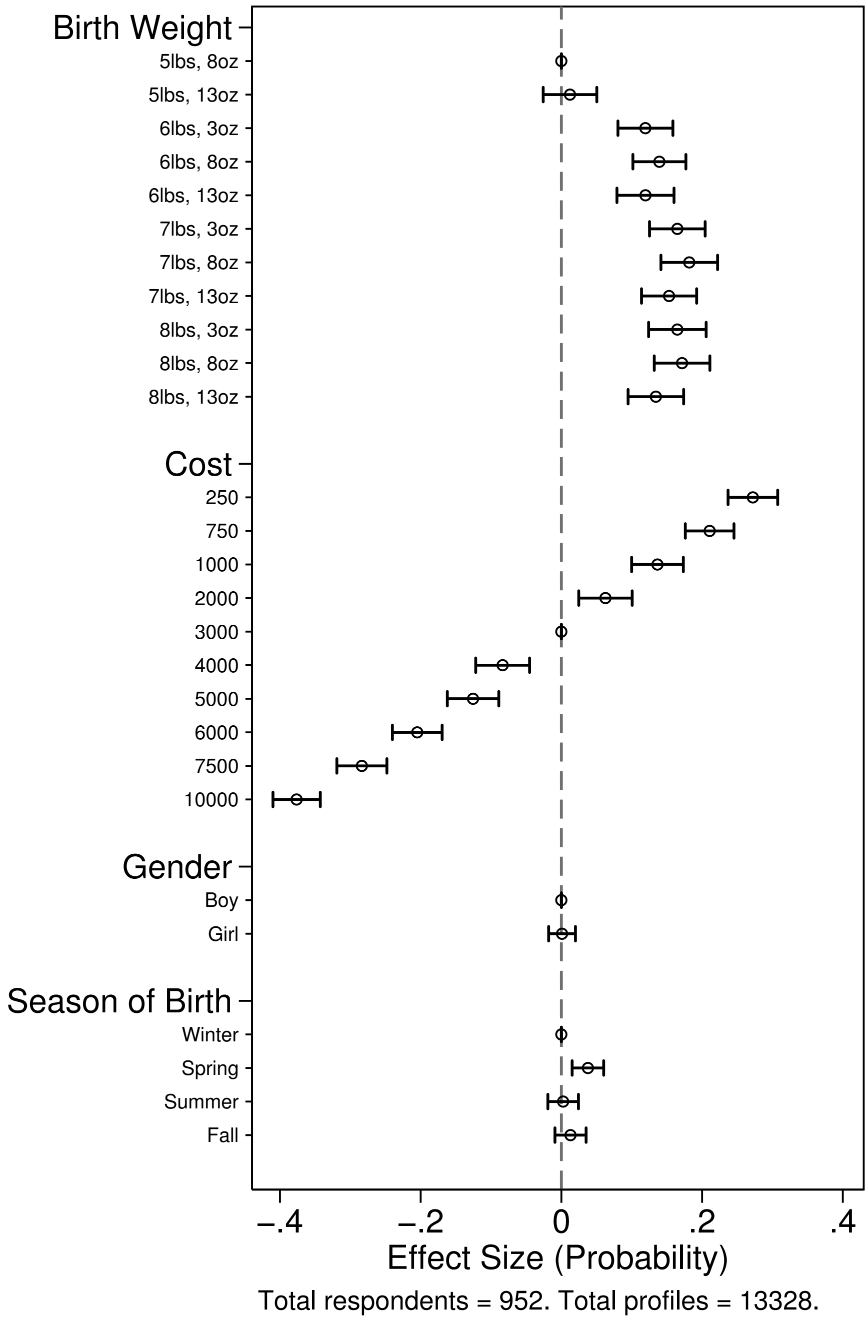
>>

Figure A4: Birth Weight from Administrative Data



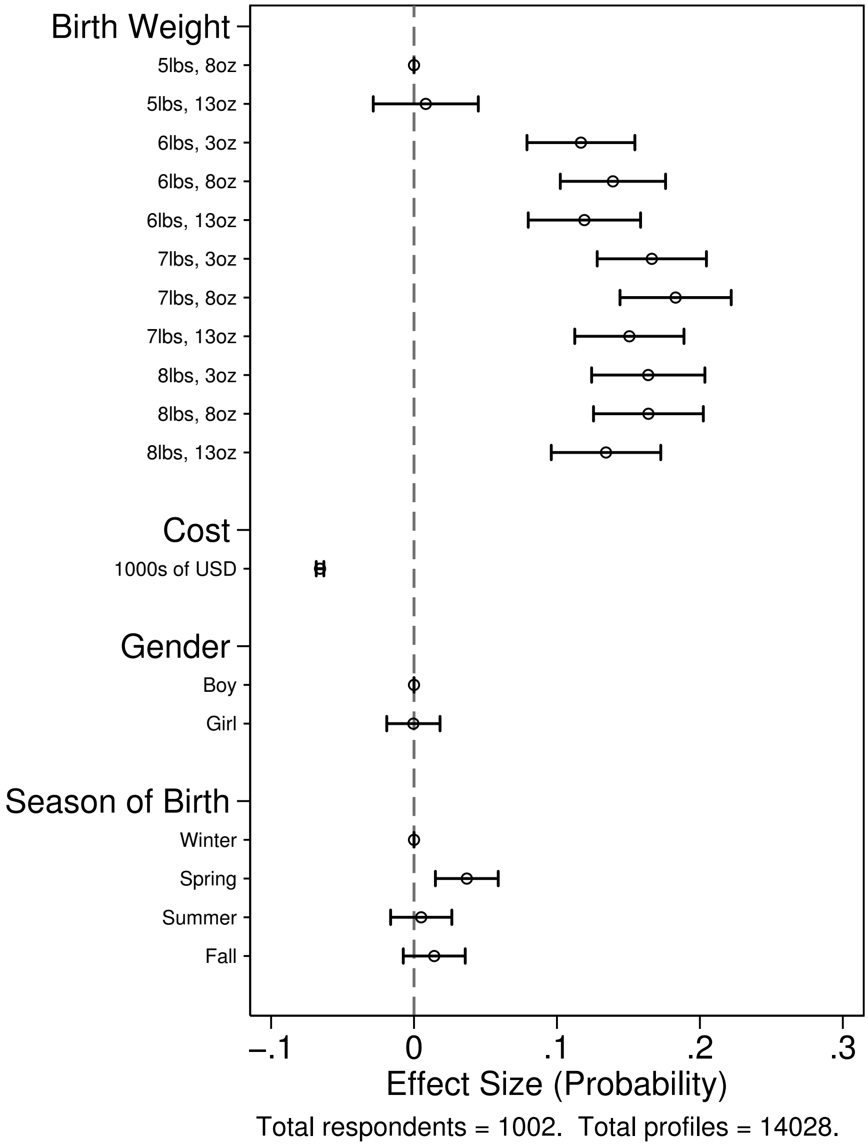
NOTES: Full birth weight distribution from all US births occurring in 2013 observed from NVSS birth certificate data (values below 500 grams or above 5000 grams are removed for display purposes). 84.09% of all births fall in the "normal" birth range of 2500 to 4000g. Of non-normal birth weights, 8.02% are low birth weight ($<$ 2,500 grams), and the remaining 7.89% were large ($>$ 4,000 grams).

Figure A5: Discrete Choice Experimental Results with Categorical Costs



NOTES: Refer to note to Figure 2. This figure is based on an identical sample, however now using a categorical, rather than a linear, measure of costs.

Figure A6: Discrete Choice Experimental Results (Full Sample)



NOTES: Refer to notes to figure 2. This figure is identical, however now also including the $\sim 5\%$ of the sample removed for failing consistency checks.

Table A2: Birth Characteristics and WTP for Birth Weight Re-weighting by State Population

	(1) Continuous	(2) Categorical
Birth Weight (in 1000s of grams)	0.090***	
	[0.010]	
Cost (in 1000s of dollars)	-0.062***	-0.062***
	[0.001]	[0.001]
5lbs, 13oz		0.019
		[0.021]
6lbs, 3oz		0.108***
		[0.021]
6lbs, 8oz		0.131***
		[0.021]
6lbs, 13oz		0.121***
		[0.023]
7lbs, 3oz		0.157***
		[0.022]
7lbs, 8oz		0.172***
		[0.023]
7lbs, 13oz		0.150***
		[0.022]
8lbs, 3oz		0.166***
		[0.023]
8lbs, 8oz		0.167***
		[0.022]
8lbs, 13oz		0.131***
		[0.020]
Girl	-0.003	-0.004
	[0.011]	[0.011]
Spring	0.033***	0.032***
	[0.012]	[0.012]
Summer	-0.005	-0.005
	[0.012]	[0.012]
Fall	0.010	0.010
	[0.012]	[0.012]
WTP for Birth Weight (1000 grams)	1452.2	
95% CI	[1111.6;1792.8]	
Observations	13328	13328

Refer to Table 2 for full notes. This table replicates these results assigning probability weights to respondents based on their state of residence so that the likelihood a particular respondent is included in the survey is the same as their state's portion of the national population.

Table A3: Comparison of Confidence Intervals for WTP from Delta Method and Block Bootstrap

Estimation Sample	WTP Point Estimate	Delta Method 95% CI	Block Bootstrap 95% CI	Full Results
Main Sample	$1438.3	[$1119.4;$1757.1]	[$1107.2;$1769.3]	Table 2
Parents	$1718.4	[$1232.4;$2204.4]	[$1233.9;$2202.8]	Table 3
Non-Parents	$1172.4	[$753.0;$1591.8]	[$757.5;$1587.3]	Table 3
Non-Parent Planners	$1376.5	[$760.6;$1992.4]	[$766.3;$1986.7]	Table 3
Non-Parent Non-Planners	$1002.5	[$433.6;$1571.3]	[$445.4;$1559.5]	Table 3
Women	$1275.7	[$856.4;$1695.0]	[$864.9;$1686.6]	Table A4
Men	$1663.8	[$1173.7;$2153.8]	[$1173.3;$2154.2]	Table A4
Mothers	$1652.1	[$1070.2;$2234.1]	[$1079.5;$2224.8]	Table A4
Fathers	$1911.3	[$1023.7;$2789.9]	[$1013.5;$2809.2]	Table A4

NOTES: WTP point estimates are all calculated as the ratio of the coefficient on birth weight (in 1000s of grams) to the coefficient on costs in dollars. The delta method for the 95% confidence interval is displayed in tables throughout the paper and is calculated directly from regression coefficients and the maximum likelihood function, while the block bootstrap confidence interval is based on re-sampling with replacement over survey respondents (not over individual profiles) in order to maintain the correct dependence structure within survey respondents. In all cases, 1,000 bootstrap samples are performed, and the 95% confidence interval is taken using the WTP in each of these 1,000 replications. Additional discussion of the relative merits of each method is available in Hole (2007*a*).

Table A4: Birth Characteristics and Willingness to Pay for Birth Weight by Gender

	All	All Respondents		Parents Only	
	(1)	(2)	(3)	(4)	(5)
		Female	Male	Mother	Father
Birth Weight (in 1000s of grams)	0.091***	0.082***	0.103***	0.103***	0.116***
	[0.010]	[0.013]	[0.014]	[0.017]	[0.024]
Cost (in 1000s of dollars)	-0.063***	-0.064***	-0.062***	-0.062***	-0.061***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.004]
Girl	0.001	0.025*	-0.030**	0.024	-0.031
	[0.010]	[0.013]	[0.015]	[0.017]	[0.024]
Spring	0.039***	0.061***	0.011	0.075***	0.010
	[0.011]	[0.015]	[0.017]	[0.019]	[0.031]
Summer	0.002	-0.001	0.010	-0.002	-0.006
	[0.011]	[0.015]	[0.016]	[0.019]	[0.029]
Fall	0.015	0.029*	-0.003	0.035*	0.005
	[0.011]	[0.015]	[0.017]	[0.019]	[0.026]
WTP for Birth Weight (1000 grams)	1438.3	1275.7	1663.8	1652.1	1911.3
95% CI	[1119.4;1757.1]	[856.4;1695.0]	[1173.7;2153.8]	[1070.2;2234.1]	[1023.7;2798.9]
Observations	13328	7448	5880	4662	2100

Average marginal effects from a logit regression are displayed. All columns include option order fixed effects and round fixed effects. Standard errors are clustered by respondent. Willingness to pay and its 95% confidence interval is estimated based on the ratio of costs to the probability of choosing a particular birth weight. The 95% confidence interval is calculated using the delta method for the ratio. Male and Female and Mother and Father refer to characteristics of experimental respondents.

Table A5: Birth Characteristics and Willingness to Pay for Birth Weight

	(1) Parents v Non-Parents	(2) Parents v Planners	(3) Parents v Non-Planners	(4) Planners v Non-Planners
Birth Weight (in 1000s of grams)	0.075***	0.090***	0.063***	0.063***
	[0.013]	[0.020]	[0.018]	[0.018]
Cost (in 1000s of dollars)	-0.063***	-0.063***	-0.062***	-0.064***
	[0.001]	[0.001]	[0.001]	[0.002]
Birth Weight × Parent	0.031	0.016	0.043*	
	[0.020]	[0.025]	[0.023]	
Birth Weight × Planning Children				0.027
				[0.027]
Girl	0.001	0.002	0.004	-0.006
	[0.010]	[0.011]	[0.011]	[0.014]
Spring	0.039***	0.044***	0.045***	0.023
	[0.011]	[0.013]	[0.013]	[0.016]
Summer	0.002	0.006	-0.005	0.009
	[0.011]	[0.013]	[0.013]	[0.015]
Fall	0.015	0.021	0.016	0.003
	[0.011]	[0.013]	[0.013]	[0.016]
WTP for Birth Weight (1000 grams)	1189.8	1424.8	1005.3	979.3
95% CI (Birth Weight)	[769.7;1609.8]	[795.4;2054.2]	[443.3;1567.3]	[427.3;1531.4]
WTP for Interation	491.2389	258.3548	694.8903	417.7294
95% CI (Interaction)	[-118.8;1101.3]	[-512.0;1028.7]	[-20.7;1410.5]	[-403.0;1238.4]
Observations	13328	9716	10374	6566

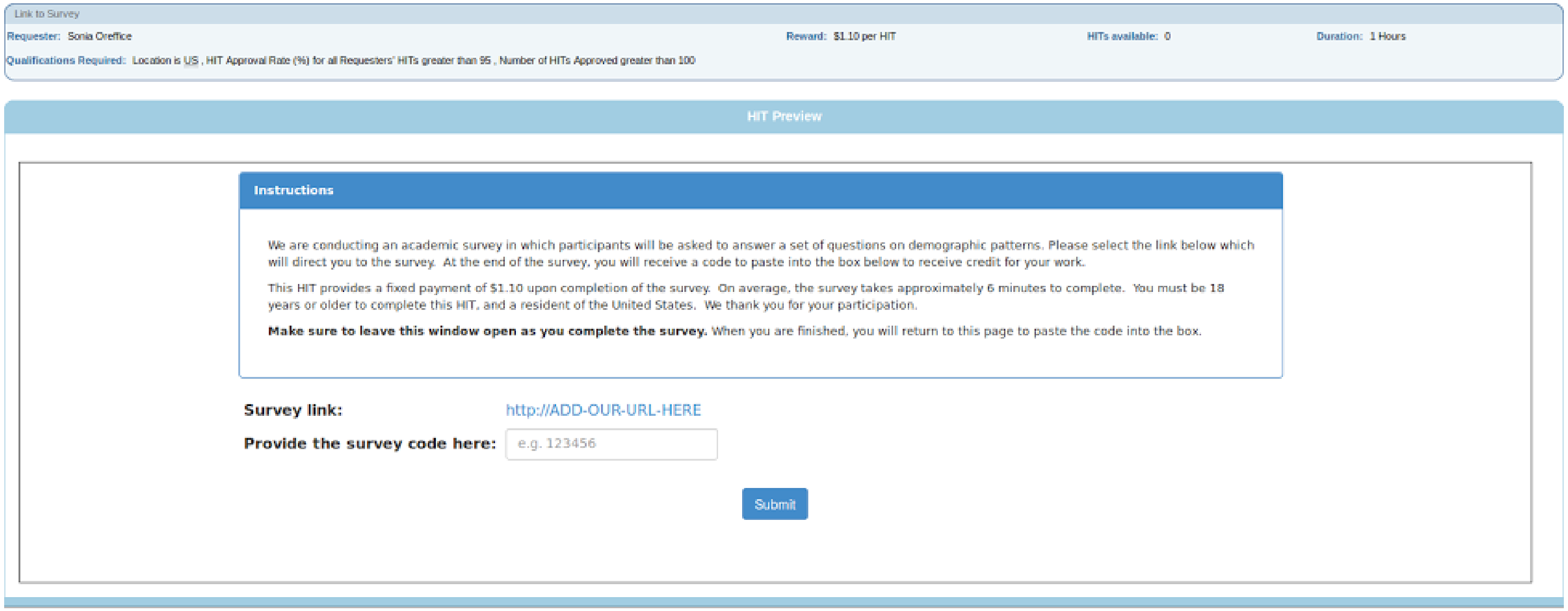
Refer to Table 2 for full notes. Each specification interacts birth weight with a dummy in order to estimate the differential importance of birth weight, as well as WTP. Values for WTP of the baseline group are displayed first in the footer, followed by the *differential* WTP for the interaction group. Each model also includes the uninteracted dummy as a control. Column 1 consists of all observations, so the interaction is interpreted as the difference between all parents and all non parents. Column 2 consists of all parents and all non parents who plan to have children (non parents who do not plan to have children are removed from the sample) so the interaction is interpreted as the difference between all parents and non parents who plan to have children. Column 3 consists of all parents and non parents who *don't* plan to have children, and column 4 consists of non-parents only, where the interaction is interpreted as the difference between those who plan to have children and those who do not.

Table A6: Birth Characteristics and Willingness to Pay for Birth Weight

	All	Parent		Non-Parents	
	(1)	(2) Yes	(3) No	(4) Planning	(5) Not-Planning
Cost (in 1000s of dollars)	-0.063***	-0.062***	-0.064***	-0.065***	-0.063***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]
5lbs, 13oz	0.012	-0.006	0.030	0.073*	-0.009
	[0.020]	[0.029]	[0.028]	[0.038]	[0.040]
6lbs, 3oz	0.119***	0.121***	0.121***	0.144***	0.100***
	[0.020]	[0.030]	[0.027]	[0.038]	[0.037]
6lbs, 8oz	0.141***	0.145***	0.138***	0.173***	0.108***
	[0.019]	[0.028]	[0.027]	[0.040]	[0.037]
6lbs, 13oz	0.119***	0.151***	0.087***	0.144***	0.039
	[0.021]	[0.029]	[0.029]	[0.043]	[0.039]
7lbs, 3oz	0.165***	0.192***	0.139***	0.161***	0.121***
	[0.020]	[0.029]	[0.028]	[0.043]	[0.038]
7lbs, 8oz	0.181***	0.218***	0.142***	0.186***	0.106***
	[0.021]	[0.030]	[0.028]	[0.041]	[0.039]
7lbs, 13oz	0.154***	0.161***	0.147***	0.160***	0.137***
	[0.020]	[0.028]	[0.028]	[0.042]	[0.038]
8lbs, 3oz	0.164***	0.195***	0.131***	0.204***	0.071*
	[0.021]	[0.030]	[0.029]	[0.041]	[0.039]
8lbs, 8oz	0.170***	0.196***	0.146***	0.201***	0.098**
	[0.020]	[0.028]	[0.029]	[0.043]	[0.039]
8lbs, 13oz	0.133***	0.132***	0.137***	0.160***	0.119***
	[0.020]	[0.030]	[0.028]	[0.041]	[0.037]
Girl	0.000	0.006	-0.006	-0.011	-0.000
	[0.010]	[0.014]	[0.014]	[0.020]	[0.019]
Spring	0.038***	0.054***	0.021	0.014	0.025
	[0.011]	[0.016]	[0.016]	[0.024]	[0.022]
Summer	0.003	-0.004	0.010	0.030	-0.005
	[0.011]	[0.016]	[0.015]	[0.023]	[0.021]
Fall	0.015	0.025*	0.003	0.010	-0.002
	[0.011]	[0.015]	[0.016]	[0.024]	[0.023]
Observations	13328	6762	6566	2954	3612

Average marginal effects from a logit regression are displayed. All columns include option order fixed effects and round fixed effects. Standard errors are clustered by respondent. Willingness to pay and its 95% confidence interval is estimated based on the ratio of costs to the probability of choosing a particular birth weight. The 95% confidence interval is calculated using the delta method for the ratio. No WTP figures are displayed in the table footer as each birth weight category is associated with its own WTP. These values are all displayed in Figure 4, or are displayed for the linear specification in Table 3.

Figure A7: Mechanical Turk Front Page



NOTES: Respondents first see the survey front page on MTurk before being redirected to the survey located Qualtrics (as displayed in figures A1, A2 and A3). A description of the process followed by respondents is provided in appendix B.

B Survey Response Procedure

Below we describe the survey response procedure as seen by survey respondents.

1. All respondents meeting survey criteria ($> 95\%$ approval rating, > 100 completed MTurk tasks, US based, and non-participants in the pilot) were able to see the Mechanical Turk HIT with the title "Link to Survey" along with the description displayed in Appendix Figure A7. Respondents are instructed that payment is conditional upon completing the survey and providing a randomized code which is displayed at the end of the survey.

2. Respondents accept participation and are directed to the discrete choice experiment on the Qualtrics survey platform.

3. Respondents must complete each question in order to move forward, and after completing the survey the randomized code is displayed.

4. Respondents return to the MTurk front page, enter their unique completed survey code and receive payment.