NBER WORKING PAPER SERIES

FERTILITY BELIEFS AND OUTCOMES: THE ROLE OF RELATIONSHIP STATUS AND ATTRACTIVENESS

Yifan Gong
Tyler Skura
Ralph Stinebrickner
Todd R. Stinebrickner

Working Paper 32578
http://www.nber.org/papers/w32578

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2024

This project was made possible by generous support from the Mellon Foundation, the Spencer Foundation, the National Science Foundation, the Social Sciences and Humanities Research Council, the Faculty of Social Sciences at University of Western Ontario, and the College of Business at University of Nebraska-Lincoln. We thank Pamela Giustinelli for valuable insight and suggestions, and data manager Lori Scafidi for invaluable research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Fertility Beliefs and Outcomes: The Role of Relationship Status and Attractiveness
Yifan Gong, Tyler Skura, Ralph Stinebrickner, and Todd R. Stinebrickner
NBER Working Paper No. 32578
June 2024
JEL No. J0,J12,J13

ABSTRACT

Unique data from the Berea Panel Study provides new evidence about fertility outcomes before age 30 and beliefs about these outcomes elicited soon after college graduation. Comparing outcomes and beliefs yields a measure of belief accuracy. Individuals who are unmarried and not in relationships at age 24 are extremely optimistic about the probability of having children, while married individuals have very accurate beliefs. Novel attractiveness measures are central for understanding fertility beliefs and outcomes for females but not for males. Marriage is a mechanism that is relevant for understanding differences in beliefs, outcomes, and misperceptions across relationship and attractiveness groups.

Yifan Gong
Department of Economics
College of Business
University of Nebraska-Lincoln
Lincoln, NE 68588
United States
ygong5@unl.edu

Tyler Skura
University of Western Ontario
tskura@uwo.ca

Ralph Stinebrickner
Berea College
Berea, KY 40403
and University of Western Ontario
ralphstineb@gmail.com

Todd R. Stinebrickner
Department of Economics
Western University
London, On N6A 5C2
and NBER
trstineb@uwo.ca

1 Introduction

Issues related to fertility are of prominent importance across economic subdisciplines.[1] This paper uses novel longitudinal data from the Berea Panel Study to provide new evidence about fertility outcomes and beliefs about these outcomes in the early post-college period. Fertility outcomes are of obvious importance because the presence of a child influences decisions that are made after the child is born. However, beliefs about these outcomes are also important because they are the relevant inputs for decisions that are made before child outcomes are realized. Naturally, it is of interest to examine the accuracy of these beliefs by comparing fertility outcomes and fertility beliefs.

The Berea Panel Study (BPS) is a longitudinal data collection project, which involved surveying students at a liberal arts college in central Kentucky—Berea College—from the time of college entrance, in 2000 and 2001, through 2014. The data from the BPS allow us to address two difficulties that have traditionally prevented a comprehensive examination of fertility outcomes and beliefs (and their comparison). The first difficulty is general in nature, although it is especially relevant in the fertility context: few datasets both characterize beliefs about outcomes that will be realized at a future time and provide information about the realizations of these outcomes. The use of surveys to directly elicit belief information has been embraced only relatively recently (Giustinelli, 2022). As a result, standard longitudinal data surveys, which provide outcome information over an extended period of time, have not traditionally elicited belief information.[2] More recent user-led surveys, which were initiated to explore the value of belief information, were not typically designed to provide long-term outcome information and tend to be short in duration, often cross-sectional.[3] Importantly, the BPS represents a rare marriage of these two different types of data sources. It was

[1] See, for example, Becker (1960); Becker et al. (1990); Boldrin et al. (2005); Black et al. (2013); Lovenheim and Mumford (2013); Olivetti and Petrongolo (2017); Boivin et al. (2022); Doepke et al. (2022).

[2] The Health and Retirement Study, which studies older Americans, is a notable exception

[3] Some examples of other user-led surveys that helped establish the value of expectations data can be seen in Wiswall and Zafar (2021), Kuziemko et al. (2018), and Boneva et al. (2022), which collect belief information related to fertility/family, and also Dominitz and Manski (1996), Arcidiacono et al. (2012), Armantier et al. (2015), Reuben et al. (2017), Arcidiacono et al. (2020), and Giustinelli and Shapiro (2024).

initiated as one of the very first user-led surveys with a central focus on collecting belief information, including beliefs about future fertility. However, unlike most other user-led surveys, it had an extended longitudinal design in which respondents were surveyed frequently throughout their time in college and annually for up to ten post-college years. This allows a unique opportunity to characterize both fertility beliefs and outcomes.[4]

The second difficulty is more context specific: starting as far back as Becker (1960), the challenge of finding robust predictors of fertility has been recognized. Given the obvious interplay between fertility and romantic relationships, of particular relevance is that measures of physical attractiveness (beauty) and personality are not typically observed in general-use surveys (Hamermesh and Biddle, 1994; Stinebrickner et al., 2019).[5] In the BPS, information about a broad set of personality traits is obtained from self-assessments. To characterize physical attractiveness, the BPS contains a rare photo-based measure. The new attractiveness (and personality) information can be used to obtain a better understanding of fertility outcomes. However, an entirely novel contribution comes from using unique belief information to provide the first evidence about whether individuals are perceptive about the role of a variable such as attractiveness.

In Section 2, we introduce the BPS and describe the data we use. Our primary sample includes recent college graduates who are childless at age 24. For this sample, we characterize the proportion of individuals who have a child born before age 30 and also the average perceived probability at age 24 of having a child born before age 30. We find that, on average, individuals at age 24 substantially overestimate the probability of having children in the future (average perceived probability of 0.576, outcome proportion of 0.361). This

[4]Our work is related to Wiswall and Zafar (2021), who elicited expectations of college students and then conducted a follow-up around age 25 to collect post-college outcomes. Given the age at follow-up, it was not possible to examine fertility, but they find evidence of overoptimism about marriage at age 25.

[5]The well-known dearth of attractiveness information can be seen in research examining the relationship between beauty and labor market outcomes. Hamermesh and Biddle (1994) find only three surveys in the U.S. and Canada (1977 QES, 1971 QLS, 1981 CQL) that contain labor market outcomes and (surveyor) evaluations of attractiveness. Scholz and Sicinski (2015) obtain attractiveness measures for the Wisconsin Longitudinal Survey (WLS) through a yearbook initiative. Empirically, past research has found evidence of factors, such as education-level and religion, that are related to fertility (Hazan and Zoabi, 2015; Götmark and Andersson, 2020).

indicates substantial ex-post belief inaccuracy (on average). Further, if there are no aggregate shocks, this ex-post inaccuracy implies that people do not have Rational Expectations, ex-ante.[6]

Intuitively, a person's marital situation is a strong predictor of future fertility. This motivated the BPS to pay close attention to the collection of information about a person's relationship status. Its characterization differentiates between being single, having a boyfriend/girlfriend, and being married. In Section 3, we obtain two main findings after stratifying by which of these groups a person is in at age 24. First, outcome data show that the proportion of individuals who have a child born before age 30 increases substantially across the three relationship stages at age 24, and belief data elicited at age 24 show that individuals correctly perceive that there is an increasing pattern. Second, the degree to which individuals overestimate the probability of having children before the age of 30 decreases dramatically across the three relationship stages at age 24, with the single group exhibiting an extremely large amount of overestimation, the boyfriend/girlfriend group exhibiting a smaller, but sizable amount of overestimation, and the married group exhibiting little overestimation. Our results are consistent with the notion that fertility is perhaps best thought of as the end result of a process in which a (typical) person moves sequentially through the three relationship stages and corrects misperceptions along the way. However, it is worth noting that, because our analysis is based on a cross-sectional comparison, the interpretation of our results could be influenced by the selection into relationship groups.

The single (at age 24) group has the most striking stratification results. They believe at 24 that they are quite likely to have a child before the age of 30 (average perceived probability of 0.428), but, in reality, are very unlikely to have a child before the age of 30 (outcome proportion of 0.131). We examine this group in more detail to understand whether this overoptimism leads to a delay or a permanent drop in fertility, and also to explore potential

[6]When expectations data are not available, researchers have traditionally used realized fertility outcomes at a point in time to characterize earlier beliefs about fertility for this time, under the assumptions of Rational Expectations and no aggregate shocks (Stinebrickner, 2001; Swann, 2005; Anstreicher and Venator, 2022). Ex-post belief accuracy is directly relevant for understanding the validity of this approach.

mechanisms underlying this overoptimism. With respect to the former, we find evidence that this group is also substantially overoptimistic at age 24 about total lifetime children. When this group observes at age 30 that they were much less likely to have a child during their twenties than expected, they revise beliefs about total lifetime children downward in a substantial fashion. With respect to the latter, we find that overoptimism about fertility before age 30 is predicted, in part but not entirely, by the fact that unmarried individuals are substantially overoptimistic about the probability of future marriage. This novel evidence about the prevalence of incorrect beliefs about fertility, as well as marriage, contributes to a growing, recent literature providing evidence that conceptually important beliefs may often be incorrect and that biases and information frictions may be more prevalent in some groups than others (Dominitz and Manski, 1996; Stinebrickner and Stinebrickner, 2012, 2014; Gong et al., 2020; de Bresser, 2021; D'Haultfoeuille et al., 2021; Crossley et al., 2024; Giustinelli and Shapiro, 2024).

In Section 4, we examine the empirical relevance of attractiveness, using our novel attractiveness measure along with information about personality traits. We first detail a quantitatively large and highly significant positive association between attractiveness and fertility outcomes. Contributing to a general literature interested in understanding gender differences, we find that the role of attractiveness arises entirely through the female subsample (Lundberg, 2022). A one unit increase in attractiveness (on a 5-point scale) predicts an extremely large increase of 0.155 in the proportion of females who have a child before age 30 (or 38 percent of the 0.405 base proportion), with a t-stat of 4.07, but the predicted increase is very close to zero for males (-0.007). Our belief data provide striking evidence that individuals are perceptive about the role of attractiveness found for outcomes—women are aware that attractiveness matters strongly for them, while men are aware that attractiveness does not matter for them.

We are also able to provide evidence about the mechanisms through which attractiveness influences fertility. One conceptually important channel is marriage. Consistent with

the gender differences seen in the role of attractiveness for fertility outcomes and beliefs, attractiveness is positively related to both marriage outcomes and beliefs for females, but is unrelated to both marriage outcomes and beliefs for males. However, we also find striking evidence that attractiveness influences fertility through channels other than marriage. For example, the proportion of married females (at age 24) who have a child before age 30 is 0.328 higher (0.844 versus 0.516) for women who have above-median attractiveness than it is for women who have below median attractiveness. In short, it seems that attractiveness plays a fundamental role that permeates many aspects in the family context.

The attractiveness measure also allows a more refined examination of ex-post accuracy. Within the married group who tend to have quite accurate beliefs as a whole, those with below-median attractiveness substantially overestimate fertility outcomes (both the probability of having a child before age 30 and number of lifetime children) while those with above-median attractiveness have remarkably accurate beliefs.

Naturally, it is useful to be cautious about the interpretation of the attractiveness results. For example, one might be especially concerned that attractiveness might be related to other personality traits that also influence fertility outcomes and beliefs. However, our data allow us to examine this potential source of omitted variables bias. We find that the associations related to attractiveness remain strong even after we control for information about personality traits. As discussed in our conclusions, our results related to attractiveness contribute to both substantive and methodological literatures. Related to both of these contributions, existing evidence of the importance of attractiveness in economics comes largely from a small number of cases where researchers established a strong relationship between rare attractiveness evaluations and labor market outcomes (Hamermesh, 2011; Mobius and Rosenblat, 2006; Scholz and Sicinski, 2015). By showing the importance of attractiveness in a completely different but arguably equally important context, this paper bolsters the notion that it would be very valuable for this type of information to be available in general-use surveys.

2 Data and Aggregate Statistics

2.1 Berea Panel Study

The BPS was designed and administered by Todd and Ralph Stinebrickner to examine a wide range of issues during the college and early post-college periods.[7] Survey data from the years 2000-2014 was merged with detailed administrative data. Here we utilize portions of the administrative data, along with annual post-college data on fertility outcomes and beliefs for Berea graduates.

Berea College has some unique features that have been documented in past BPS research. Perhaps of particular note, the nature of the school's mission—providing an education to students of great promise but limited economic resources—leads to a student body that tends to be from lower income backgrounds (Gong et al., 2022). However, the academic preparation (e.g., college entrance exams) and the academic outcomes (e.g., major choices) of students at Berea are similar to students at nearby institutions like the University of Kentucky (Stinebrickner and Stinebrickner, 2008b). The study of one school requires awareness of generalizability issues, but can also have benefits. For example, intuitively appealing proxies for human capital at the start of the career such as cumulative college Grade Point Average (GPA) and college major, which are available in administrative data, are directly comparable across students from the same school.

One novel feature of the BPS is that it contains a measure of physical attractiveness. A (color) ID picture of each student was assessed by, on average, 47 evaluators using a five-point integer scale where 5="very attractive", 4="above average", 3="average", 2="below average", and 1="significantly below average."[8] The instructions that were given to each reviewer are shown in Appendix A. The goal of the survey approach was to elicit a measure that had a cardinal interpretation. In effect, the instructions asked each evaluator to report

[7]For a subset of previous BPS research, see Stinebrickner and Stinebrickner (2004, 2006, 2008a,b, 2012, 2014), Mehta et al. (2019), Agopsowicz et al. (2020), and Crossley et al. (2024).

[8]We chose college students to be the evaluators because their age made their attractiveness assessments most relevant for issues related to dating and marriage of our young sample.

the attractiveness quintile for each picture they evaluated, relative to the full set of 149 photos that they were randomly assigned to evaluate. Thus, an evaluation of 1 should mean that a picture was in the lowest 20% in terms of attractiveness and an evaluation of 5 should mean that a picture was in the top 20% in terms of attractiveness. For most of our analysis, our attractiveness measure for a person is the average of the (roughly) 47 evaluations of the person's picture. Later, we formally discuss the importance of having a large number of evaluators for the quality of this measure.[9] As discussed in more detail in Section 4, the data also contain self-assessments of attractiveness and personality traits.

Another important feature of the data is that the BPS characterizes relationship status using a rather comprehensive set of states. Whether a person is married is elicited in each year (Survey Question A.1, Appendix A). Further, the BPS differentiates between the possible states of being single and having a boyfriend or girlfriend (Survey Question A.2). Finally, the data allows us to characterize whether people are cohabiting with partners (Survey Question A.3).

The BPS allows a comparison of fertility beliefs and fertility outcomes because it contains: 1) expectations data characterizing beliefs about future fertility and 2) information about fertility outcomes over an extended period after college graduation. A natural fertility outcome of interest is the age at which a person's first child is born. With respect to 1), in each post-college year, the BPS used probabilistic expectations questions to elicit beliefs about this outcome. As an example, for an individual who is age 24 and childless, these beliefs are elicited using Survey Question 1.[10]

[9] Stinebrickner et al. (2019) and Scholz and Sicinski (2015) discuss the importance of recognizing sampling variation in assessments.

[10] A person's beliefs about whether to have a child and the timing of having a child are likely influenced by both the person's preferences for children and their physical ability to have children.

Question 1. We are interested in when you expect to have your first child. What is the percent chance that your first child will be born at each of the following ages or not at all?

Your Age	*Percent Chance your first child will be born when you are this age*
At age 24 or 25	_____
At age 26 or 27	_____
At age 28 or 29	_____
At or after Age 30	_____
Never have children	_____

We divide the perceived percent chances by 100 to obtain perceived probabilities.

With respect to 2), in each post-college year, the BPS elicited information about children using Survey Question A.4. Of particular relevance here, for each person the survey question elicits the birthdate of their oldest child. This allows us to compute the parent's age when the first child is born.

2.2 Sample Selection

Given our primary interest in understanding issues related to the birth of a first child, it is desirable to examine beliefs soon after college graduation. We use age 24 when Berea graduates will have typically completed at least one post-college survey. The BPS attained very high response rates in both the college and post-college periods. Ninety-three percent of graduates in the two cohorts completed at least one annual post-college survey. Belief information is available at age 24 for 469 graduates. Our broadest sample is the 454 of these students for whom observable characteristics, including attractiveness, are available. However, because beliefs about the timing of one's first child are not relevant for people who already have a child at age 24, our main focus is on the 391 of these individuals who are childless at age 24.

In terms of outcomes, we examine births before the age of 30, since a considerable number of respondents were only slightly older than 30 in the last BPS survey year. Fertility and marital outcome information is available at age 30 for the 355 of the 391 students (91 percent)

who answered the annual BPS survey at age 30 or older. These 355 students represent our primary sample. In terms of descriptive statistics, 68% of this sample are female, and the mean (std. deviation) of attractiveness and GPA are 2.58 (0.773) and 3.20 (0.456), respectively.

Turning to relationship status, at age 24, 29% of the sample are married, 34% have a boyfriend or girlfriend, and 37% are single. Dividing the group who have a boyfriend or girlfriend further, we find that 24% of this group are cohabiting with their partner and 76% are not.[11] We note that the number of observations after stratifying by cohabitation is quite small. Thus, for our primary analyses in subsequent sections we examine differences across the married, boyfriend/girlfriend, and single groups, but also note that further stratification to take into account cohabitation tends to further strengthen our main results.

2.3 Fertility Outcomes, Beliefs, and Inaccuracy

The perceived probability of having a child before age 30 is obtained by adding the responses in the first three rows of Question 1. The first row of Table 1 reveals substantial inaccuracy in beliefs for the primary sample as a whole. The average perceived probability of having a child before age 30 is 0.576, while the proportion of individuals who have a child before age 30 is only 0.361.[12] While we find that, as expected, males have somewhat lower fertility outcomes and beliefs in the sample, we do not find meaningful gender differences in inaccuracy.[13] Sections 3 and 4 explore other predictors of fertility beliefs, outcomes, and inaccuracy, with a focus on relationship status and attractiveness.

[11] As expected we find that virtually all married people (99%) report that they are cohabiting.

[12] A general, but largely unstudied, concern in surveys of this type is that the process of eliciting expectations about a possible event may influence the likelihood of the outcome occurring. This might be the case, for example, if expectations data are used to explore beliefs about the costs and benefits of a particular decision. A loosely related concern when studying accuracy over time is that improved accuracy may be a result of higher quality responses over time, rather than changes in actual accuracy. In the absence of a project design in which there exists a control group that was not exposed to survey questions, one might think about examining how outcomes and beliefs vary with the number of past surveys that a person failed to complete at a point in time. However, this approach would not seem particularly promising in the context

Table 1: Cumulative fertility and marriage outcomes before age 30 and their beliefs by current (age 24) relationship status.

	Fertility		Marriage	
	Outcomes mean/sd/n	Beliefs mean/sd/n	Outcomes mean/sd/n	Beliefs mean/sd/n
Total	0.361	0.576	0.434	0.635
	(0.481)	(0.343)	(0.497)	(0.332)
	[355]	[355]	[251]	[251]
Married	0.692	0.760		
	(0.464)	(0.325)		
	[104]	[104]		
Boyfriend/Girlfriend	0.322	0.576	0.669	0.784
	(0.469)	(0.330)	(0.472)	(0.272)
	[121]	[121]	[121]	[121]
Cohabiting	0.429	0.520	0.750	0.754
	(0.504)	(0.385)	(0.441)	(0.320)
	[28]	[28]	[28]	[28]
Non-Cohabiting	0.303	0.599	0.663	0.799
	(0.462)	(0.305)	(0.475)	(0.256)
	[89]	[89]	[89]	[89]
Single	0.131	0.428	0.215	0.496
	(0.338)	(0.297)	(0.413)	(0.323)
	[130]	[130]	[130]	[130]

Notes: Column 1 reports the proportion of individuals who have a child before age 30 for the full sample of individuals who are childless at age 24 (Row 1), and for this full sample stratified by relationship status at age 24 (subsequent rows). Column 2 reports the average perceived probability (elicited at age 24) of having a child before age 30 for the same groups as in Column 1. Column 3 reports the proportion of individuals who get married before age 30 for the subset of the full childless sample who are not married at age 24, and for this subset stratified by relationship status (other than married) at age 24 (Rows 3-4). Column 4 reports the average perceived probability (elicited at age 24) of getting married before age 30 for the same groups as in Column 3.

Standard deviations are shown in parentheses while number of observations are shown in square brackets.

Sample sizes among Non-Cohabiting and Cohabiting sub-groups do not sum to 121 since 4 individuals with a boyfriend/girlfriend did not report cohabitation information.

3 Fertility and Relationship Status

3.1 Relationship Status as a Predictor

The Role of Relationship Status

We examine fertility outcomes and beliefs after stratifying by each person's relationship status—married, boyfriend/girlfriend, single—at age 24. The first column of Table 1 shows that the proportion of individuals who have a child before age 30 is 0.131 for the single group, 0.322 for the boyfriend/girlfriend group, and 0.692 for the married group. The second column shows that the average perceived probability of having a child before age 30 is 0.428, 0.576, and 0.760 for the three groups. Thus, our first main finding is that: 1) the proportion of individuals who have a child born before age 30 increases substantially across the relationship stages at age 24, and individuals correctly perceive that there is an increasing pattern. Looking across the columns in a particular row provides evidence about the inaccuracy of fertility beliefs. The factor by which individuals overestimate the probability of having a child by age 30 is 3.27 (0.428/0.131) for the single group, 1.79 for the boyfriend/girlfriend group, and only 1.10 for the married group. Thus, our second main finding is that: 2) ex-post inaccuracy varies dramatically with relationship status at age 24, with the substantial inaccuracy for the full sample coming almost entirely from unmarried individuals, especially single individuals whose beliefs are extremely inaccurate.

As noted earlier, our subsamples become quite small when we further stratify by the cohabitation status of individuals with a boyfriend or girlfriend. Nonetheless, as seen in the fourth and fifth rows of Table 1, this further stratification, if anything, strengthens the finding that ex-post inaccuracy is substantially higher for individuals who are not as far along in their relationship stage. Specifically, the factor by which individuals overestimate the probability of having a child by age 30 is 1.977 (0.599/0.303) for those in the boyfriend/girlfriend group

studied here - since survey non-response in a particular year may be strongly related to fertility outcomes.

[13]The difference between the average perceived probability and the outcome proportion is $0.258 = 0.523 - 0.265$ for males and $0.196 = 0.601 - 0.405$ for females.

who are not cohabiting, but is 1.215 (0.520/0.428) for those in the boyfriend/girlfriend group who are cohabiting. Thus, in terms of accuracy, those who are cohabiting have factors that are quite similar to the factors for the married group. Relative to these cohabiting individuals, those who are not cohabiting have factors that are more similar to the factors for the single group.[14]

The two findings above are consistent with the notion that the birth of children is best viewed as the end result of a process in which individuals first move through the relationship stages of being single, having a boyfriend/girlfriend, and getting married. As an individual moves closer to the end of this process, uncertainty is resolved and misperceptions about earlier stages are corrected (or become irrelevant).

Panel A of Table 2 shows results from regressions of fertility outcomes (a child before age 30) and fertility beliefs (perceived probability of a child before age 30) on not only relationship information, but also on other potential non-relationship predictors in the BPS. The two findings from Table 1 remain. With respect to the first finding, we see in both Column 4 (outcomes) and Column 1 (beliefs) that, relative to the omitted category, Single, both the coefficients on Married and Boyfriend/Girlfriend are significantly positive, with the Married coefficient substantially larger than the Boyfriend/Girlfriend coefficient. With respect to the second finding, compared to a very substantial positive misperception (average perceived probability − outcome proportion) for the Single group, the misperception predicted by Boyfriend/Girlfriend is smaller by 0.046 (Boyfriend/Girlfriend in Column 4 − Boyfriend/Girlfriend in Column 1), and the misperception predicted by Married is smaller by 0.218 (Married in Column 4 − Married in Column 1).

An important question is whether the substantial overoptimism about children before age

[14]Given the small number of cohabiting individuals, our results in the remainder of the paper abstract away from issues related to cohabitation. At first glance, these accuracy results suggest that perhaps the cohabiting individuals in the boyfriend/girlfriend group should be combined with the married group rather than the non-cohabiting subset of the boyfriend/girlfriend group. However, making this less appealing, the cohabiting group is more similar to the non-cohabiting group than the married group in terms of, for example, fertility outcomes. Regardless, because the number of individuals in the cohabiting group is small, the subsequent results in the paper are not influenced substantially by this decision.

30 tends to imply (ex-post) inaccurate beliefs about the total number of lifetime children, against the alternative that overoptimistic individuals simply tend to delay fertility. We focus on the single group because they were seen in Table 1 to have a 30 percentage point overestimation of the probability of having a child before age 30. While completed fertility is not observed, the BPS does contain an annual survey question that elicits beliefs about total number of lifetime children (Survey Question A.5). Table B3 shows that the expected total number of children falls for this group by 0.238 between age 24 and age 30 (a drop of 14.3 percent, from 1.68 to 1.44), with a test of no change yielding a t-statistic of 3.45. This drop is very substantial, implying (roughly speaking) that one less child than expected born before age 30 translates to a reduction of $\frac{0.238}{0.3} = 0.793$ (between ages 24 and 30) in the expected number of lifetime children. Individuals in the single group are seemingly substantially overoptimistic about how many children they will have in their lifetimes.

Other Predictors

Next, we turn to examining potential non-relationship predictors of fertility beliefs and outcomes in the BPS. Traditional measures of human capital, such as educational attainment, are strong predictors of fertility because, for example, they are related to the opportunity cost of having children. We examine whether, within a group of individuals with the same level of education, it is possible to find further predictors of fertility that are related to human capital. Row 5 of Table 2 shows results for cumulative college GPA at the time of graduation. Looking across the columns of Row 5, Panel A, there is little statistical evidence that this measure is related to either beliefs or outcomes.[15]

However, a very different story emerges in Row 4, Panel A of Table 2, which examines the more-novel attractiveness measure described in Section 2.1—the average attractiveness evaluation of a person's picture across all evaluators. Looking across this row, all coefficients on Attractiveness in both the beliefs and outcomes specifications are positive and significant

[15]For some of the specifications, e.g., Column 3 of Panel B, the point estimate of the coefficient on GPA is moderately large. However, because the standard error is large (partly due to the small sample size), we cannot reject the null of no linear association.

Table 2: Cumulative fertility outcomes by age 30 regressed on observable characteristics

	Panel A: Pooled Sample						**Panel B: Male Sub-Sample**						**Panel C: Female Sub-Sample**					
	Beliefs			Outcomes			Beliefs			Outcomes			Beliefs			Outcomes		
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Male	-0.072	-0.075	-0.061	-0.110	-0.119	-0.108												
	(0.037)	(0.039)	(0.036)	(0.048)	(0.054)	(0.048)												
Married	0.310		0.149	0.528		0.476	0.308		0.227	0.658		0.341	0.312		0.112	0.479		0.510
	(0.041)		(0.050)	(0.056)		(0.100)	(0.077)		(0.085)	(0.088)		(0.163)	(0.050)		(0.060)	(0.070)		(0.112)
Boyfriend/Girlfriend	0.137		0.058	0.183		0.185	0.201		0.169	0.280		0.293	0.114		0.030	0.145		0.145
	(0.041)		(0.040)	(0.053)		(0.054)	(0.067)		(0.071)	(0.073)		(0.075)	(0.053)		(0.051)	(0.073)		(0.073)
Attractiveness	0.064	0.084	0.065	0.083	0.116	0.082	-0.004	0.005	0.013	-0.025	-0.007	-0.018	0.076	0.100	0.072	0.120	0.155	0.121
	(0.022)	(0.023)	(0.021)	(0.030)	(0.032)	(0.030)	(0.042)	(0.047)	(0.043)	(0.052)	(0.061)	(0.052)	(0.026)	(0.026)	(0.025)	(0.037)	(0.038)	(0.037)
College GPA	0.004	0.043	-0.018	0.047	0.113	0.047	-0.056	-0.033	-0.093	0.073	0.119	0.087	0.055	0.104	0.035	0.023	0.101	0.024
	(0.039)	(0.042)	(0.039)	(0.054)	(0.058)	(0.054)	(0.068)	(0.076)	(0.070)	(0.086)	(0.100)	(0.086)	(0.051)	(0.054)	(0.051)	(0.075)	(0.078)	(0.075)
Age 30 marriage beliefs			0.401						0.226						0.484			
			(0.063)						(0.101)						(0.076)			
Age 30 marriage						0.063						0.366						-0.039
						(0.095)						(0.164)						(0.104)
Constant	0.257	0.233	0.133	-0.271	-0.333	-0.277	0.487	0.583	0.454	-0.225	-0.069	-0.335	0.086	-0.007	-0.065	-0.273	-0.437	-0.277
	(0.154)	(0.159)	(0.152)	(0.184)	(0.200)	(0.185)	(0.235)	(0.241)	(0.249)	(0.283)	(0.336)	(0.280)	(0.206)	(0.205)	(0.194)	(0.255)	(0.260)	(0.256)
College major controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	355	355	345	355	355	355	113	113	109	113	113	113	242	242	236	242	242	242

Notes: Standard errors are in parentheses.

Regressions in Panel A are based on the pooled sample. Regressions in Panel B and C are based on the male and female sub-samples, respectively. The dependent variable in the first three columns (beliefs) is the perceived probability of having a child before age 30. The dependent variable in the Outcomes columns in each panel is a binary variable that takes the value of 1 if a child is born before age 30 and 0 otherwise.

at all traditional levels. Given that our ability to characterize attractiveness is quite unique, we return to an in-depth analysis of its role and importance in Section 4.

3.2 Misperceptions About Marriage as a Mechanism

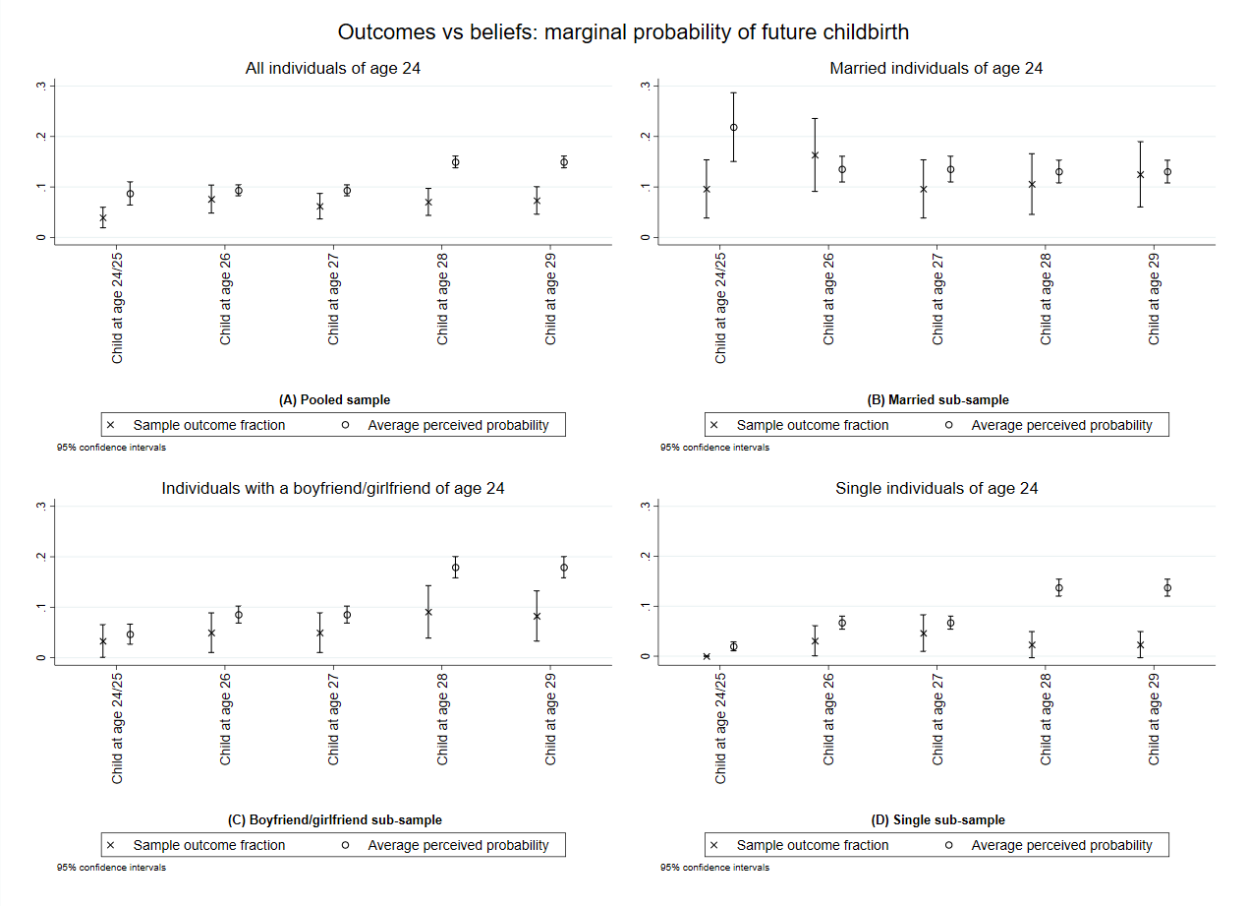
We explore the potential mechanisms for overoptimism about fertility, focusing on the strongly overoptimistic single group. As a starting point, we examine age patterns in this overoptimism by taking advantage of the disaggregated (marginal) belief probabilities in Question 1, along with annual outcome data. Panel D of Figure 1, shows that, for the single group, the size of the overestimation about fertility grows substantially as the twenty-four year old respondent looks further into the future. At each of the future ages 24/25, 26, and 27, the overestimation (at age 24) is small and the 95 percent confidence intervals for beliefs and outcomes tend to overlap. However, at each of the future ages 28 and 29, the overestimation (at age 24) has become very substantial—the sample average perceived probability is over twice as large as the sample outcome proportion—and the confidence intervals no longer overlap.

Given that an overwhelming majority of births in our sample occur within marriages, this increasing pattern in overoptimism suggests a potential mechanism: single individuals, who are seemingly likely to have a good sense of whether they will be married in the near future, might be substantially overoptimistic about getting married in the longer term.[16] For this mechanism to be relevant, it is necessary that: 1) beliefs about marriage are overoptimistic and 2) beliefs about marriage are predictors of beliefs about fertility. These conditions can be examined empirically because marital outcomes were collected for the full sample period and Survey Question A.2 elicits beliefs about future marriage using a format similar to Question 1.

Continuing to examine the single group, we find evidence in support of both of these two conditions. With respect to the first condition, the last row of Table 1 shows that the

[16] Among all the children born to sample members, 2.76% were born to single individuals, 4.83% were born to individuals with a boyfriend/girlfriend, and 92.41% were born to married individuals.

Figure 1: Outcomes vs beliefs: marginal probability of future childbirth



Note: For age 24 to age 29, Panel A plots the proportion of individuals whose first child was born (blue) at each age and the average perceived probability (elicited at age 24) of having the first child at each age. Panel B, C, and D plot these numbers for subsamples stratified by relationship status.

The BPS outcome data characterizes the exact age at which a person's first child arrives, while the belief information is elicited using two year future-age bins. To make outcomes and beliefs comparable, we obtain beliefs about individual future ages by assuming that the perceived probability associated with each of the two ages within a bin is equal. A slight complication exists for the first age bin because some individuals may have recently turned 24 when they completed the survey while others may have been close to turning 25. Thus, for both outcomes and beliefs, we use a combined 24/25 age bin, which contains any remaining part of age 24 along with age 25.

average perceived probability of being married before age 30 is 0.496, while the proportion of individuals who are married before age 30 is only 0.215. To provide evidence about the second condition, we regress the perceived probability of having a child before age 30 on the perceived probability of becoming married before age 30. We estimate a slope parameter of 0.491, with a t-statistic of 4.95.

Given the relevance of this mechanism, we quantify the role of misperceptions about marriage (for misperceptions about fertility) by calculating how much of the misperceptions about fertility would remain if people had "correct" perceptions about the probability of having a child before the age of 30 conditional on marital status. A rough back-of-the-envelope calculation suggests that overoptimism about marriage before age 30 for the single group accounts for approximately 28 percent of the overoptimism about fertility before the age of 30.[17] Remaining overoptimism, relating to the probability of having a child conditional on being married, would come from, for example, overoptimism about how quickly marriage will take place or about how quickly children will arrive after marriage. Evidence that these non-marriage channels of misperceptions are also relevant comes from the boyfriend/girlfriend group who are seen in Table 1 to exhibit overoptimistic fertility beliefs (average perceived probability = 0.576, outcome proportion = 0.322) despite having relatively accurate marriage beliefs (average perceived probability = 0.784, outcome proportion = 0.669). For this group, the average perceived probability of having a child before age 30 conditional on being married before age 30 is $0.735 = 0.576/0.784$, while the actual probability of having a child before age 30 conditional on being married before age 30 is only $0.481 = 0.322/0.669$. Similarly, for the single group, these two numbers are $0.962 = 0.428/0.496$ and $0.609 = 0.131/0.215$.

[17] In the single group, 0.36 of those who were married at age 30 have children before age 30, while 0.069 of those who were not married at age 30 have children before age 30. Combining these conditional probabilities with the perceived probabilities of getting married before age 30 yields an unconditional probability of having a child before age 30 of 0.213 ($0.496 \times 0.36 + 0.504 \times 0.069 \approx 0.213$). This implies an overestimation of 0.082 ($0.213 - 0.131$), which is 28 percent of the total overestimation of 0.297 ($0.428 - 0.131$).

4 Fertility and Attractiveness

4.1 The Role of Attractiveness

Combined Male-Female Sample

In Section 3.1 we found that attractiveness is an important predictor of fertility beliefs and outcomes. Here we examine the role of this trait further. Noting that some of the effect of attractiveness may operate through a relationship status channel, to quantify the "total" predictive ability of attractiveness, Columns 2 and 5, Panel A of Table 2 show results for the full-sample from belief and outcome specifications that include the attractiveness measure but not relationship status. Column 5 shows that a one unit increase in attractiveness (1.29 standard deviations) predicts a 0.116 increase in the probability of a child before age 30, with a t-stat of 3.59. To stress the quantitative importance of this estimate, this predicted increase is 32 percent of the observed proportion of sample members who have a child before age 30 (0.361, Table 1). Column 2 provides striking evidence that individuals are perceptive about the role of attractiveness. Consistent with the result in Column 5, Column 2 shows that a one-unit increase in attractiveness is related to an almost 0.084 increase in the perceived probability of having a child before age 30 (t-stat = 3.65). Our paper not only provides rare evidence about the role of attractiveness for predicting an outcome other than earnings, but also provides the first evidence about whether individuals understand the role of attractiveness.[18]

To examine the relevance of relationship status as a mechanism for our attractiveness results, in Column 3 (beliefs) we add, to Column 2, relationship status at age 24 and the perceived probability of being married before age 30. Similarly, in Column 6 (outcomes) we add, to Column 5, relationship status at age 24 and marital status at age 29. In both of these specifications, controlling for the relationship status variables reduces the estimated Attractiveness coefficient, but the coefficient remains quantitatively large and statistically

[18] See, for example, Henderson and Anglin (2003); Jokela (2009) for investigations of non-earnings outcomes in disciplines other than economics.

significant (Column 3: estimate = 0.065, t-stat = 3.10, Column 6: estimate = 0.082, t-stat = 2.73).[19] Thus, the way that attractiveness influences fertility is seemingly complex and broad, operating not only through marriage per se but also largely through other channels, perhaps related in some way to marital satisfaction.[20]

Male-Female Differences

Past literature indicates that males and females value attractiveness (and other personality traits) differently when considering a person as a potential partner (Fisman et al., 2006). Contributing to this literature, we find striking evidence that the role of attractiveness seen in the full sample for both beliefs and outcomes is due entirely to its importance for females. Panels B and C of Table 2 show results after stratifying the sample by gender. The last three columns of Row 3, Panel C reveal that fertility outcomes are even more strongly related to attractiveness for females than was seen for the sample as a whole, while the similarly strong connection between beliefs and attractiveness seen in the first three columns shows that females are perceptive about this relationship. In contrast, the last three columns of Row 3, Panel B indicate that fertility outcomes are not related to attractiveness for males, while the absence of a connection between beliefs and attractiveness seen in the first three columns shows that males are perceptive about this lack of a relationship.

Row 2 of Table 3 provides intriguing evidence that at least some of the differential effect of attractiveness on fertility by gender operates through a differential effect of attractiveness on marriage by gender. Columns 5 and 6 show that attractiveness has a strong relationship with marriage outcomes at age 29 for females (t-stat = 3.02), but this relationship is not present for males (the coefficient is negative and not significant). Columns 2 and 3 show a remarkably similar pattern for beliefs. Again, as was found for fertility, these marriage results show that males and females are perceptive about the role of attractiveness for marriage.

To examine whether the attractiveness results help us further understand issues related to

[19]The Attractiveness coefficients remain virtually unchanged when we add indicators of marriage at each age 25-29 or add perceived probabilities of marriage at each age 25-29.

[20]Jokela (2009) discuss non-marriage avenues through which attractiveness could influence fertility. Meltzer et al. (2014) discuss the role of physical attractiveness on the trajectory of marital satisfaction.

ex-post accuracy, we reconstruct Table 1 for females after stratifying by whether a person's attractiveness is above or below the sample median. Our findings from Section 3.1 about how fertility beliefs, fertility outcomes, and ex-post accuracy vary across relationship stages are largely present for the stratified results. However, attractiveness allows us to provide more refined evidence about ex-post accuracy, with this being particularly relevant for what takes place within the married group who, as a whole, were seen in Table 1 to exhibit quite accurate beliefs. As seen in Tables B1 and B4 (Appendix B), married females with above median attractiveness have extremely accurate beliefs about having a child before age 30 (average perceived probability = 0.859, outcome proportion = 0.844) and, on average, have beliefs about the expected lifetime number of children that remain constant between age 24 and 29 (age 24 = 2.17, age 29 = 2.17). In contrast, as seen in Tables B2 and B5, married females with below median attractiveness substantially overestimate the probability of having a child before the age of 30 (average perceived probability = 0.701 and outcome proportion = 0.516), and, on average, have a large downward revision of expected lifetime number of children between age 24 and age 29 (age 24 = 1.909 and age 29 = 1.574). Both the overestimation and the downward revision are significant at a 5% level. These differences in accuracy by attractiveness are mainly driven by substantially lower fertility outcomes of the below-median attractiveness group. Specifically, the proportion of married females who have a child before age 30 is 0.328 (0.844 − 0.516) higher for the above-median group than the below-median group, and this difference is significant at a 1% level. Attractiveness has a powerful presence even within marriage.

Table 3: 5-year-ahead marriage beliefs and outcomes regressed on observable characteristics

	Beliefs			Outcomes		
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled	Female	Male	Pooled	Female	Male
Male	0.012			-0.017		
	(0.038)			(0.059)		
Attractiveness	0.036	0.055	-0.049	0.078	0.111	-0.040
	(0.021)	(0.023)	(0.047)	(0.033)	(0.037)	(0.071)
College GPA	0.105	0.104	0.154	0.166	0.235	0.029
	(0.039)	(0.047)	(0.075)	(0.060)	(0.074)	(0.105)
College major controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	345	236	109	355	242	113

Notes: Standard errors are in parentheses.
The dependent variable in the first three columns (beliefs) is the perceived probability (elicited at age 24) of being married five years from today (age 29). We elicit marriage beliefs using Survey Question A.6 from the BPS since this question is answered by all survey participants, including those who are currently married. The marriage beliefs used here and those used in table 1 are found to be highly correlated (correlation = 0.7) among unmarried individuals. Note that they should not be perfectly correlated since Survey Question A.6 also considers the possibility of divorce. The dependent variable in the last three columns (outcomes) is a binary variable that takes the value of 1 if the individual is married at age 29 and 0 otherwise.

4.2 Quality of the Attractiveness Measure

Our observed attractiveness measure for person i is the average **reported** attractiveness evaluation of i's picture **across the** J_i **evaluators** who were randomly assigned the picture. Abstracting for the time being away from issues related to the quality of pictures, the most salient question related to the quality of our observed measure is how it relates to the conceptual object of interest - the average **true** attractiveness evaluation of i's picture across **all possible** (relevant-aged) people in the population.[21] The simple statistical model below is helpful for thinking about the amount of measurement error that is present in our observed attractiveness measure, and how this might bias coefficients in our regressions.

Let A_{ij} denote the actual evaluation of person i's picture by evaluator j. Then, our conceptual object of interest is the mean of A_{ij} across all possible evaluators j in the relevant population, which we denote $\mu_{A,i}$. We denote the variance of A_{ij} across j as $\sigma^2_{A,i}$. Letting R_{ij} denote evaluator j's reported attractiveness evaluation for person i, our observed attractiveness measure is the sample mean of R_{ij}, denoted $\bar{R}_i \equiv \frac{1}{J_i}\sum_j^{J_i} R_{ij}$. Thus, the quality issue at hand is how $\bar{R}_i$ relates to $\mu_{A,i}$.[22]

One reason that $\bar{R}_i$ may differ from $\mu_{A,i}$ is that the reported attractiveness evaluation of evaluator j might not accurately reflect the true attractiveness evaluation of evaluator j if, for example, j is not fully focused when evaluating i's picture. We allow for the possibility of measurement error by writing

$$R_{ij} = A_{ij} + \epsilon_{ij}, \tag{1}$$

where ϵ_{ij} represents idiosyncratic measurement error.

A second reason that $\bar{R}_i$ may differ from $\mu_{A,i}$ arises because i's picture is assigned to only a

[21] One might wonder whether, for example, a picture at a single point in time can effectively represent a person's attractiveness at that time or at later times. This issue is discussed in Scholz and Sicinski (2015) and also mentioned in our conclusions.

[22] Stinebrickner and Stinebrickner (2004) studies a similar measurement issue in the context of college study time.

sample of size J_i from the full population of possible evaluators. This possibility is relevant if there exists heterogeneity across j in the true attractiveness evaluations of i's picture. We allow for the potential for this type of heterogeneity by denoting $U_{ij} \equiv A_{ij} - \mu_{A,i}$ as the deviation of evaluator j's true attractiveness evaluation of i from the mean true attractiveness evaluation of i.

Given this notation, the manner in which the two potential reasons influence the relationship between $\bar{R}_i$ and $\mu_{A,i}$ can be seen by writing:

$$\begin{aligned}
\bar{R}_i &= \frac{1}{J_i}\sum_{j}^{J_i}(A_{ij} + \epsilon_{ij}) \\
&= \mu_{A,i} + \frac{1}{J_i}\sum_{j}^{J_i}[(A_{ij} - \mu_{A,i}) + \epsilon_{ij}] \\
&\equiv \mu_{A,i} + \frac{1}{J_i}\sum_{j}^{J_i}(U_{ij} + \epsilon_{ij}), \\
&\equiv \mu_{A,i} + \eta_i. \qquad (2)
\end{aligned}$$

That is, the report is the sum of the object of interest and a composite error $\eta_i \equiv \frac{1}{J_i}\sum_{j}^{J_i}(U_{ij} + \epsilon_{ij})$, which contains the objects of relevance for the two reasons.

For each i, all evaluations j are randomly sampled from the same population. Hence, by construction, U_{ij} has a mean of zero and a variance of $\sigma^2_{A,i}$. We assume the measurement error ϵ_{ij} is classical for each picture i, i.e., $E_j[\epsilon_{ij}] = 0$ and $U_{ij} \perp \epsilon_{ij}$, and let $\sigma^2_{\epsilon,i}$ denote $var_j[\epsilon_{ij}]$.[23] As a result, for each i, the composite error term η_i has a mean of zero and a variance of $\frac{1}{J_i}(\sigma^2_{A,i} + \sigma^2_{\epsilon,i})$.

Because all the regression analyses conducted in this paper are at the individual (student) level, what is relevant for our analysis is the statistical property of the distribution of η_i across pictures of these students. Using the law of iterated expectations and the law of total

[23] $U_{ij} \perp \epsilon_{ij}$ for all i is not necessary for estimating the magnitude of the bias but is useful for understanding the heterogeneity in A_{ij} (U_{ij}). Given the format of our question, measurement error could also arise because of rounding (Manski and Molinari, 2010; Giustinelli et al., 2022).

variance, it is easy to show that $E_i[\eta_i] = 0$, $E_i[\mu_{A,i}\eta_i] = 0$, and $var_i[\eta_i] = E_i[\frac{1}{J_i}(\sigma^2_{A,i} + \sigma^2_{\epsilon,i})] = E_i[\frac{1}{J_i} var_j[R_{ij}]]$. This means η_i has the property of a (potentially) heteroscedastic classical measurement error.

It is well-known that the existence of such type of classical measurement error in the independent variable could lead to attenuation bias. For example, consider a simple OLS regression with attractiveness as the only regressor. Using $\bar{R}_i$ instead of the desired $\mu_{A,i}$ yields a bias in the OLS estimator given by the equation below:

$$\beta^{OLS} \underset{p}{\rightarrow} \beta \cdot \left(1 - \frac{var_i[\eta_i]}{var_i[\bar{R}_i]}\right), \tag{3}$$

where β^{OLS} is the OLS estimator using $\bar{R}_i$ as the only regressor and β is the true coefficient on the attractiveness measure.

It is clear that what matters for the magnitude of the bias is the amount of variation due to measurement error ($var_i[\eta_i]$) relative to the total variation in the attractiveness measure ($var_i[\bar{R}_i]$). The former can be consistently estimated by $\widehat{var}_i[\eta_i] \equiv \frac{1}{n}\sum_i^n [\frac{1}{J_i}(\widehat{var}_j[R_{ij}])]$, where $\widehat{var}_j[R_{ij}]$ is the sample variance of R_{ij} for person i. The latter can be consistently estimated by $\widehat{var}_i[\bar{R}_i]$, the sample variance of $\bar{R}_i$. For the sample of 355 people in our main sample, we have $\widehat{var}_i[\eta_i] = 0.016$ and $\widehat{var}_i[\bar{R}_i] = 0.602$. Thus, Equation 3 implies that β^{OLS} is biased downwards by a factor of only 2.64%. Note that the large number of evaluations plays a non-trivial role in the quality of our measure. For example, if we had ten evaluations for each picture (i.e., $J_i = 10$ for all i, instead of roughly $J_i = 47$), then this bias would be approximately 11.15%. If we had only four evaluations for each picture, the bias would be approximately 23.89%. Further, if as often has been the case in previous studies of attractiveness, there is only one evaluator for each picture, the bias would be approximately 55.66%.[24]

It is interesting to think about whether we can predict some of the variation in U_{ij}

[24]The single evaluator scenario is common when, for example, attractiveness is evaluated by a person conducting some type of survey.

across evaluators. For each evaluator, we have information about their gender and their self-evaluated attractiveness. Using a regression with R_{ij} as the dependent variable and a picture fixed effect to control for $\mu_{A,i}$, we find that 1) male photos receive similar ratings from male and female evaluators, but female photos receive higher ratings from female evaluators than male evaluators and 2) photos tend to receive higher ratings when evaluators self-report that they are more attractive.[25]

4.3 Interpretation and Robustness

When recognizing why it is necessary to be cautious about the interpretation of the attractiveness results, one might be especially concerned that attractiveness might be related to other personality traits that also influence fertility outcomes such as a person's communication ability (Feingold, 1992). To examine this, we utilize self-reported personality measures (described in Table 4), which are available for all but three sample members. We focus on females, because the importance of attractiveness comes entirely from this group, and a specification without relationship variables, because this specification roughly captures the "total" importance of attractiveness.

Columns 1 (beliefs) and 4 (outcomes) of Table 4 repeat the specifications in Columns 2 (beliefs) and 5 (outcomes) of Table 2. It is important to consider measurement error in self-reported personality traits. We can provide some information about the quality of the self-reported measures because one of these measures is self-reported attractiveness, for which we also observe our peer-assessed measure. Columns 2 and 4 of Table 4 replace peer-assessed attractiveness with self-reported attractiveness. These two measures take on somewhat different values, so the coefficients associated with the measures have slightly different interpretations. However, consistent with our findings for the peer-assessed measure, the self-reported estimates are quantitatively important (coefficients of 0.082 and 0.098 for the beliefs and outcomes regressions, respectively) with large t-statistics (3.04 and 2.72,

[25] See Table B6 in Appendix B for detailed results.

respectively). Thus, the evidence suggests that the self-reported measures contain important content.

Columns 3 and 6 of Table 4 add a full set of self-reported personality traits to the specifications in Columns 1 and 4. Importantly, we do not find evidence of an omitted variables problem related to personality traits—there is little change in the attractiveness coefficients. We find little evidence that personality traits are related to fertility. There is some evidence that women who have outgoing personalities believe they are more likely to have children before age 30 and that women who are physically strong have fewer children before age 30. The latter result further bolsters the notion that attractiveness plays an important role in fertility, because physical strength may represent a somewhat different dimension of attractiveness (e.g., a person's size/weight) than our (facial) attractiveness measure.

5 Conclusions

This paper provides the first comparison of beliefs about fertility and actual fertility outcomes. We find substantial heterogeneity in the ex-post accuracy of beliefs. For example, married individuals with above-average attractiveness have extremely accurate beliefs, while single individuals vastly overestimate the likelihood of fertility outcomes. This accuracy of beliefs is fundamentally important because many decisions of interest to policymakers depend on fertility. For example, previous literature has documented a strong negative effect of young children on female labor supply (Eissa and Hoynes, 2004; Hoffman, 2009; Gong et al., 2022), so that beliefs about future fertility may affect, for example, the anticipated return to current human capital investments or the trade-off between current consumption and savings.

Generally, policies that incentivize human capital investments (e.g., tuition subsidies for graduate school or other types of training programs) can be particularly beneficial for young

Table 4: Cumulative fertility outcomes before age 30 and their beliefs regressed on observable characteristics - female

	Beliefs			Outcomes		
	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness	0.100		0.102	0.155		0.149
	(0.026)		(0.026)	(0.038)		(0.040)
phys_attract		0.082			0.098	
		(0.027)			(0.036)	
College GPA	0.104	0.116	0.113	0.101	0.122	0.096
	(0.054)	(0.055)	(0.056)	(0.078)	(0.078)	(0.081)
Self-reported personality traits						
Physical strength top 25%			−0.069			−0.201
			(0.088)			(0.116)
Physical coordination top 25%			−0.015			0.129
			(0.063)			(0.092)
Raw intelligence top 25%			−0.024			−0.013
			(0.050)			(0.073)
Communication skills top 25%			0.058			−0.061
			(0.054)			(0.072)
Outgoing personality top 25%			0.098			0.087
			(0.052)			(0.072)
Creativity top 25%			−0.000			0.065
			(0.050)			(0.069)
Relate to others top 25%			0.050			−0.007
			(0.056)			(0.076)
College major controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	242	239	239	242	239	239

Notes: Standard errors are in parentheses.
All regressions are based on the female sub-sample. The dependent variable in the first three columns (beliefs) is the perceived probability of having a child before age 30. The dependent variable in the last three columns (outcomes) is a binary variable that takes the value of 1 if a child is born before age 30 and 0 otherwise.

For each personality trait, individuals report whether they believe they are in the first, second, third, or fourth quartile, with the comparison group being other students in their college cohort. We find, roughly speaking, that around 75 percent of respondents believe they are above the median with respect to a particular personality trait. As a result, for all personality traits other than self-reported attractiveness, we use a binary variable which equals 1 if the individual believes they are within the top 25% of their cohort, and zero otherwise.

workers who have many years to receive returns. Our results about belief accuracy suggest a second reason that incentivizing early human capital investments is beneficial—without these policies, young and mostly single workers may tend to under-invest in their human capital due to misperceptions about future fertility. This second reason becomes much less relevant as workers become older and are more likely to be married.

The heterogeneity by relationship status and attractiveness that we document has important implications for data collection. With respect to relationship status, our findings suggest the importance of differentiating between the relationship status of unmarried individuals for survey collection. With respect to attractiveness, our results further highlight the substantial value of including this type of information in general-use surveys, and suggest that its collection need not be overly burdensome. Indeed, our primary measure comes from simple ID-type pictures. One set of quality issues relates to how best to construct measures from a particular picture. We find that, because there is non-trivial heterogeneity in the ratings of a particular picture across evaluators, it is important to have a fairly large number of evaluators. Another set of quality issues relates to the qualities of the pictures themselves. Conceptually, ID pictures have the appealing quality that they are standardized so that variation in attractiveness evaluations is not due to differences in picture quality. We note that it is an open question whether we might obtain stronger results if we moved away from a scenario of using a single picture taken before the beginning of the sample period we analyze. However, it is reassuring that we obtain qualitatively similar results using a self-reported measure of attractiveness elicited during the sample period.

An ongoing challenge in the burgeoning expectations literature is to find ways to provide confidence in the quality of expectations data. One concern is that expectations data might be noisy, and therefore have little content. The sharpness of many of our results is relevant in this respect. For example, our finding that fertility (and marriage) beliefs and outcomes are strongly related to attractiveness for females, but not for males requires not only that 1) attractiveness is an important attribute and that 2) the attractiveness measure is of

high quality, but also that 3) the belief data is of high quality. Additional faith from our survey questions comes from our findings that 1) the accuracy of beliefs improves over the relationship stages, single, boyfriend/girlfriend, and married; and 2) overestimation about fertility outcomes is followed by a downward adjustment of expected number of lifetime children.

References

Agopsowicz, A., C. Robinson, R. Stinebrickner, and T. Stinebrickner (2020). Careers and mismatch for college graduates. *Journal of Human Resources 55*(4), 1194–1221.

Anstreicher, G. and J. Venator (2022, September). To Grandmother's House We Go: Childcare Time Transfers and Female Labor Mobility. Boston College Working Papers in Economics 1051, Boston College Department of Economics.

Arcidiacono, P., V. J. Hotz, and S. Kang (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics 166*(1), 3–16. Annals Issue on "Identification and Decisions", in Honor of Chuck Manski's 60th Birthday.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020). Ex ante returns and occupational choice. *Journal of Political Economy 128*(12), 4475 – 4522.

Armantier, O., W. Bruine de Bruin, G. Topa, W. van der Klaauw, and B. Zafar (2015). Inflation expectations and behavior: Do survey respondents act on their beliefs? *International Economic Review 56*(2), 505–536.

Becker, G. S. (1960). An Economic Analysis of Fertility. In *Demographic and Economic Change in Developed Countries*, NBER Chapters, pp. 209–240. National Bureau of Economic Research, Inc.

Becker, G. S., K. M. Murphy, and R. Tamura (1990). Human capital, fertility, and economic growth. *Journal of Political Economy 98*(5), S12–S37.

Black, D. A., N. Kolesnikova, S. G. Sanders, and L. J. Taylor (2013, 03). Are Children "Normal"? *The Review of Economics and Statistics 95*(1), 21–33.

Boivin, J., R. Vassena, M. Costa, E. Vegni, M. Dixon, B. Collura, M. Markert, C. Samuelsen, J. Guiglotto, E. Roitmann, and A. Domar (2022). Tailored support may reduce mental and

relational impact of infertility on infertile patients and partners. *Reproductive BioMedicine Online 44*(6), 1045–1054.

Boldrin, M., M. De Nardi, and L. E. Jones (2005, February). Fertility and social security. Working Paper 11146, National Bureau of Economic Research.

Boneva, T., M. Golin, K. Kaufmann, C. Rauh, and K. M. Kaufmann (2022). Beliefs about maternal labor supply. CESifo Working Paper Series 10148, CESifo.

Crossley, T., Y. Gong, T. R. Stinebrickner, and R. Stinebrickner (2024). Examining Income Expectations in the College and Early Post-college Periods: New Distributional Tests of Rational Expectations. *Journal of the European Economic Association* (forthcoming).

de Bresser, J. (2021). Evaluating the accuracy of counterfactuals the role of heterogeneous expectations in life cycle models. Discussion Paper 2021–034, Tilburg University, Center for Economic Research.

D'Haultfoeuille, X., C. Gaillac, and A. Maurel (2021, May). Rationalizing rational expectations: Characterizations and tests. TSE Working Papers 21-1211, Toulouse School of Economics (TSE).

Doepke, M., A. Hannusch, F. Kindermann, and M. Tertilt (2022, April). The economics of fertility: A new era. Working Paper 29948, National Bureau of Economic Research.

Dominitz, J. and C. F. Manski (1996). Eliciting Student Expectations Of The Returns To Schooling. Econometrics 9411002, University Library of Munich, Germany.

Eissa, N. and H. W. Hoynes (2004). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of Public Economics 88*(9), 1931–1958.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin 111*(2), 304–341.

Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics 121*(2), 673–697.

Giustinelli, P. (2022, August). Expectations in Education: Framework, Elicitation, and Evidence. Working Papers 2022-026, Human Capital and Economic Opportunity Working Group.

Giustinelli, P., C. F. Manski, and F. Molinari (2022). Tail and center rounding of probabilistic expectations in the health and retirement study. *Journal of econometrics 231*(1), 265–281.

Giustinelli, P. and M. D. Shapiro (2024, April). Seate: Subjective ex ante treatment effect of health on retirement. *American Economic Journal: Applied Economics 16*(2), 278–317.

Gong, Y., R. Stinebrickner, and T. Stinebrickner (2022). Marriage, children, and labor supply: Beliefs and outcomes. *Journal of Econometrics 231*(1), 148–164.

Gong, Y., T. Stinebrickner, and R. Stinebrickner (2020). Perceived and actual option values of college enrollment. *Journal of Applied Econometrics 35*(7), 940–959.

Gong, Y., T. Stinebrickner, R. Stinebrickner, and Y. Yao (2022). The role of non-pecuniary considerations: Locations decisions of college graduates from low income backgrounds. University of Western Ontario, Centre for Human Capital and Productivity (CHCP) Working Papers 20221, University of Western Ontario, Centre for Human Capital and Productivity (CHCP).

Götmark, F. and M. Andersson (2020). Human fertility in relation to education, economy, religion, contraception, and family planning programs. *BMC Public Health 20*(265).

Hamermesh, D. S. (2011). *Beauty Pays: Why Attractive People Are More Successful*. Princeton University Press.

Hamermesh, D. S. and J. E. Biddle (1994). Beauty and the labor market. *The American Economic Review 84*(5), 1174–1194.

Hazan, M. and H. Zoabi (2015, September). Do highly educated women choose smaller families? *Economic Journal 125*(587), 1191–1226.

Henderson, J. J. and J. M. Anglin (2003). Facial attractiveness predicts longevity. *Evolution and Human Behavior 24*(5), 351–356.

Hoffman, S. (2009). The changing impact of marriage and children on women's labor force participation. *Mon. Lab. Rev. 132*(3).

Jokela, M. (2009). Physical attractiveness and reproductive success in humans: Evidence from the late 20 century. *Evol Hum Behav 30*(5), 342–350.

Kuziemko, I., J. Pan, J. Shen, and E. Washington (2018). The mommy effect: Do women anticipate the employment effects of motherhood? NBER Working Papers 24740, National Bureau of Economic Research, Inc.

Lovenheim, M. F. and K. J. Mumford (2013, 05). Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market. *The Review of Economics and Statistics 95*(2), 464–475.

Lundberg, S. (2022, April). Gender Economics: Dead-Ends and New Opportunities. IZA Discussion Papers 15217, Institute of Labor Economics (IZA).

Manski, C. F. and F. Molinari (2010). Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics 28*(2), 219–231.

Mehta, N., R. Stinebrickner, and T. Stinebrickner (2019, January). Time-Use And Academic Peer Effects In College. *Economic Inquiry 57*(1), 162–171.

Meltzer, A. L., J. K. McNulty, G. L. Jackson, and B. R. Karney (2014). Sex differences in the implications of partner physical attractiveness for the trajectory of marital satisfaction. *Journal of personality and social psychology 106*(3), 418—428.

Mobius, M. M. and T. S. Rosenblat (2006, March). Why beauty matters. *American Economic Review 96*(1), 222–235.

Olivetti, C. and B. Petrongolo (2017, Winter). The economic consequences of family policies: Lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives 31*(1), 205–230.

Reuben, E., M. Wiswall, and B. Zafar (2017, 01). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal 127*(604), 2153–2186.

Scholz, J. K. and K. Sicinski (2015, March). Facial Attractiveness and Lifetime Earnings: Evidence from a Cohort Study. *The Review of Economics and Statistics 97*(1), 14–28.

Stinebrickner, R. and T. Stinebrickner (2004). Time-use and college outcomes. *Journal of Econometrics 121*(1-2), 243–269.

Stinebrickner, R. and T. Stinebrickner (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics 90*(8-9), 1435–1454.

Stinebrickner, R. and T. Stinebrickner (2008a). The causal effect of studying on academic performance. *The B.E. Journal of Economic Analysis & Policy 8*(1).

Stinebrickner, R. and T. Stinebrickner (2008b). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review 98*(5), 2163–84.

Stinebrickner, R. and T. Stinebrickner (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics 32*(3), 601–644.

Stinebrickner, R., T. Stinebrickner, and P. Sullivan (2019). Beauty, Job Tasks, and Wages: A New Conclusion about Employer Taste-Based Discrimination. *The Review of Economics and Statistics 101*(4), 602–615.

Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics 30*(4), 707–748.

Stinebrickner, T. R. (2001). Compensation policies and teacher decisions. *International Economic Review 42*(3), 751–780.

Swann, C. (2005). Welfare reform when recipients are forward-looking. *Journal of Human Resources 40*(1).

Wiswall, M. and B. Zafar (2021). Human capital investments and expectations about career and family. *Journal of Political Economy 129*(5), 1361–1424.

A Survey Questions

Question A.1 Are you currently married? YES NO
Maiden name of spouse _____ AGE of spouse _____ Religion of spouse _____
When did you get married to current spouse? Month _____ Year _____

Question A.2 (answered by unmarried) Do you have a boyfriend/girlfriend? YES NO
Name of boyfriend/girlfriend _____ What is his/her AGE? _____
What is his/her RELIGION? _____

We are interested in whether you think you will get married and when you think you will get married. What is the percent chance that your next marriage will take place at each of the following ages or not at all? **Note: Each number should be between 0 and 100 and the numbers should sum to 100.**

Your Age	*Percent Chance of next marriage taking place at this age*
At or before Age 23	_____
At age 24 or 25	_____
At age 26 or 27	_____
At age 28 or 29	_____
At or after Age 30	_____
Never get married	_____

Question A.3
Do you currently live with one or more parents, grandparents or an aunt or uncle? YES NO
Do you currently live with a spouse or boyfriend/girlfriend? YES NO

Question A.4 (answered by everyone)
What is your current AGE? _____ What is your Religion? _____
How many children do you currently have? 0 1 2 3 4 5 or more
If you have children, when was **your oldest child** born? Month _____ Year _____
If you have more than 1 children, when was **your youngest child** born?
Month _____ Year _____

Question A.5 (answered by everyone) What is the percent chance that you will have the following total number of children during your lifetime? **Note: Each number should be between 0 and 100 and the numbers should app up to 100.**

Number of children	*Percent Chance of this number of children*
0	_____
1	_____
2	_____
3	_____
4 or more	_____

Question A.6 (answered by everyone) What is the percent change you will be married five years from today? _____ Note: Number should be between 0 and 100. Take into account the possibility that you will get married and probability you will get divorced.

Attractiveness Survey:

1. What is your gender? Circle one Male Female

2. Take a minute or two to look through the three pages of pictures that are stapled to this page.

3. We would like you to tell us how attractive you find each person on the three pages. Please write a number between 1 and 5 on each of the picture. The number you write should represent how attractive you think the person in the picture is **relative to the other people on the three pages.**

Use the following scale where higher numbers correspond to people being more attractive.
5 = very attractive
4 = above average
3 = average
2 = below average
1 = significantly below average

Note: Because you are rating the attractiveness of each person relative to other pictures on the three pages, you should be writing roughly the same amount of each number. But, there is no need for there to be exactly the same number of each type.

4. Using the scale above, how would you rate your own attractiveness?_____

B Additional Figures and Tables

Table B1: Cumulative age 30 fertility and marriage outcomes and beliefs by current (age 24) relationship status—above median-attractiveness females.

	Fertility		Marriage	
	Outcomes mean/sd/n	Beliefs mean/sd/n	Outcomes mean/sd/n	Beliefs mean/sd/n
Total	0.512	0.685	0.539	0.709
	(0.502)	(0.314)	(0.502)	(0.301)
	[121]	[121]	[76]	[76]
Married	0.844	0.859		
	(0.367)	(0.235)		
	[45]	[45]		
Boyfriend/Girlfriend	0.378	0.643	0.689	0.809
	(0.490)	(0.310)	(0.468)	(0.253)
	[45]	[45]	[45]	[45]
Single	0.226	0.494	0.323	0.564
	(0.425)	(0.294)	(0.475)	(0.310)
	[31]	[31]	[31]	[31]

Table B2: Cumulative age 30 fertility and marriage outcomes and beliefs by current (age 24) relationship status—below median-attractiveness females.

	Fertility		Marriage	
	Outcomes mean/sd/n	Beliefs mean/sd/n	Outcomes mean/sd/n	Beliefs mean/sd/n
Total	0.298	0.516	0.367	0.571
	(0.459)	(0.355)	(0.485)	(0.347)
	[121]	[121]	[90]	[90]
Married	0.516	0.701		
	(0.508)	(0.370)		
	[31]	[31]		
Boyfriend/Girlfriend	0.344	0.534	0.625	0.740
	(0.483)	(0.363)	(0.492)	(0.310)
	[32]	[32]	[32]	[32]
Single	0.155	0.407	0.224	0.478
	(0.365)	(0.302)	(0.421)	(0.333)
	[58]	[58]	[58]	[58]

Table B3: Expected number of lifetime children at age 24 and 29 by current (age 24) relationship status—pooled.

	Age 24 mean/sd/n	Age 29 mean/sd/n
Total	1.819 (0.814) [331]	1.663 (0.853) [332]
Married	2.048 (0.854) [98]	1.955 (0.862) [99]
Boyfriend/Girlfriend	1.769 (0.766) [112]	1.644 (0.815) [112]
Single	1.680 (0.792) [121]	1.442 (0.815) [121]

Table B4: Expected number of lifetime children at age 24 and 29 by current (age 24) relationship status—female above median attractiveness.

	Age 24 mean/sd/n	Age 29 mean/sd/n
Total	1.996 (0.704) [113]	1.882 (0.780) [113]
Married	2.167 (0.679) [42]	2.172 (0.771) [42]
Boyfriend/Girlfriend	1.898 (0.716) [43]	1.783 (0.744) [43]
Single	1.892 (0.696) [28]	1.598 (0.727) [28]

Table B5: Expected number of lifetime children at age 24 and 29 by current (age 24) relationship status—female below median attractiveness.

	Age 24 mean/sd/n	Age 29 mean/sd/n
Total	1.658 (0.882) [113]	1.498 (0.821) [113]
Married	1.909 (0.931) [30]	1.574 (0.843) [30]
Boyfriend/Girlfriend	1.602 (0.806) [29]	1.577 (0.849) [29]
Single	1.548 (0.882) [54]	1.414 (0.802) [54]

Table B6: Attractiveness ratings regressed on evaluator characteristics

	(1)
Evaluator_Male	-0.112
	(0.018)
Male * Evaluator_Male	0.134
	(0.032)
Evaluator Self-Rating	0.033
	(0.010)
Constant	-0.113
	(0.140)
Picture Fixed Effects	Yes
Observations	14127

Standard errors in parentheses.