Marriage, Children, and Labor Supply: Beliefs and Outcomes*

Yifan Gong[1], Todd Stinebrickner[1], and Ralph Stinebrickner[†2]

[1] *Department of Economics, University of Western Ontario, London, Ontario, Canada*
[2] *Mathematics Department, Berea College, 101 Chestnut St., Berea, KY 40404, USA*

April 15, 2020

Abstract

While a large literature is interested in the relationship between family and labor supply outcomes, little is known about the expectations of these objects at earlier stages. We examine these expectations, taking advantage of unique data from the Berea Panel Study. In addition to characterizing expectations, starting during college, the data details outcomes for ten years after graduation. Methodological contributions come from approaches to validate quality of survey expectations data and the recognition that expectations data, along with longitudinal data, can potentially help address endogeneity issues arising in the estimation of the causal effect of family on labor supply.

Keywords: gender differences in labor supply, children, marriage, expectations data

JEL: J13, J22

*This project was made possible by generous support from the Mellon Foundation, the Spencer Foundation, the National Science Foundation, and the Social Sciences and Humanities Research Council, and benefited from the insight of Peter Arcidiacono.

†Email addresses: ygong48@uwo.ca (Yifan Gong), trstineb@uwo.ca (Todd Stinebrickner), ralph_stinebrickner@berea.edu (Ralph Stinebrickner).

1 Introduction

Much research has documented the existence of gender differences in labor supply and how much these differences change over time. For example, Juhn and Potter (2006) find that the labor force participation rate of prime-age females with a college degree increased substantially between 1969 and 2004 (from 62.3% to 81.2%), while the labor force participation rate of prime-age males with a college degree remained high and stable over the same period (over 90%). Not surprisingly, descriptive evidence of relevance for understanding the reasons for these differences has a natural focus on the family. For example, evidence from the Current Population Survey indicates that married women were less likely than single women to participate in the labor force between 1984 and 2004 (Hoffman, 2009), and Eissa and Hoynes (2004) show that differences in labor force participation rates between married men and women were particularly large for those who had two or more children between 1985 and 1997.[1] Further, the policy relevance of gender differences in labor supply and the relationship of these differences to family outcomes has been well-recognized. Ruhm (1998), Schönberg and Ludsteck (2014), and Byker (2016) study the effect of parental leave policies on female labor supply. Gustafsson and Stafford (1992) and Bauernschuster and Schlotter (2015) analyze the importance of childcare provision. Blau and Kahn (2013) analyze the effects of several policies in a unified framework.

However, less studied is the reality that, if gender differences in labor supply are anticipated, men and women will tend to have different incentives to invest in human capital accumulation earlier in the lifecycle.[2] This paper is motivated by the reality that little is known about what males and females believe about future labor supply during the formal schooling period or during the early years in the labor force.[3] A traditional approach for characterizing beliefs about the probability of working at some point in the future would be to assume that individuals have Rational Expectations, e.g., that the perceived probability of working is given by the proportion of observationally similar individuals who are seen to be working at that point in the future. However, a recent "expectations" literature, which is built on the notion that beliefs are perhaps best viewed as data that can be collected directly using carefully written survey questions,

[1] See, e.g., Browning (1992) and Blundell and MaCurdy (1999) for surveys of early literature and Blau and Kahn (2007) for a more recent investigation.

[2] Gronau (1988), Blau and Ferber (1991), and Shaw and Shapiro (1987) discuss how gender differences in beliefs about future labor force attachment can lead to gender differences in the amount of human capital that is accumulated while working. College major represents a prominent example of a schooling decision that may depend on one's beliefs about future labor supply. There is a growing literature that has recognized the general usefulness of expectations data for understanding this decision (see, e.g., Zafar, 2013, Wiswall and Zafar, 2014, Stinebrickner and Stinebrickner, 2014a, Arcidiacono, Hotz and Kang, 2012, and Altonji, Arcidiacono, and Maurel, 2016).

[3] Wiswall and Zafar (2016) provide some of the only other direct evidence.

has suggested that this assumption is not necessarily appealing.[4] This paper contributes to this literature by taking advantage of beliefs about future labor supply that were collected for a group of individuals while they were still in college. The data come from the Berea Panel Study (BPS), perhaps the first extended longitudinal survey motivated very directly by the potential promise of expectations data.

Our belief data show that, on average, men and women have very similar perceived labor supply probabilities for age 28 under scenarios in which they are not married or are married without children. Indeed, in the sample, women indicate that, under these scenarios, they are slightly more likely to be working full-time and slightly less likely to not be working at all. However, we find a sizable gender gap in perceptions about the probability of working for the scenario where young children are present. A decomposition exercise indicates that this gender gap plays a more prominent role than the gender gap in beliefs about the timing of children in creating a substantial gender gap in beliefs about "unconditional" labor supply. Confidence in the ability of individuals to answer expectations questions comes both from the fact that: 1) two different approaches for characterizing beliefs about unconditional labor supply, which use two different survey questions, yield very similar results and 2) on average, male (female) beliefs about the future labor supply of females (males), both unconditional and conditional on family scenarios, correspond very closely to female (male) beliefs about their own future labor supply. These checks relate directly to research interested in providing evidence about measurement error that exists in responses to expectations questions on surveys (e.g., Ameriks et al., 2019, Manski and Molinari, 2010, Giustinelli, Manski, and Molinari, 2019). In particular, 1) relates closely to an approach for characterizing measurement error implemented in Gong, Stinebrickner and Stinebrickner (2019).[5][6]

The earlier discussion highlights that both gender differences in actual labor supply and gender differences in beliefs about future labor supply are of policy relevance. Then, our findings about beliefs raise a natural question: for a very recent cohort of graduates, are gender differences in actual labor supply present only for the scenario where young children are present? We examine this issue by taking advantage of the fact that the BPS followed respondents for approximately ten years after college graduation. We find that,

[4]It is worth noting that the empirical approaches in Blau and Ferber (1991) and Shaw and Shapiro (1987) are in the spirit of this literature in that they take advantage of views about future labor force attachment collected using surveys. Specifically, they examine whether expected or actual wage profiles (which presumably depend on human capital accumulation) are related to whether a person has "plans" to work in the future or to the number of years a worker thinks she will be out of the workforce before retirement.

[5]Intuitively, differences in a particular unconditional perceived probability computed using different sets of expectations questions are informative about the amount of measurement error present in the underlying survey questions. Later we discuss the assumptions necessary to employ this approach in the context of this paper, and show some results in an appendix.

[6]Under the assumption that the measurement error present in expectations data is classical, its magnitude can be estimated if researchers have repeated measurements (e.g., Drerup, Enke and von Gaudecker, 2014) or if researchers are willing to embed these measurement errors in fully specified structural models (e.g., Lochner, 2003). For a discussion and analysis of early concerns about expectations data in a development context, see Delavande, Gine, and Mckenzie (2011).

on average, beliefs are extremely accurate - if anything women work slightly more than men when married or married without children, but work very different amounts when children are present.[7] In addition, we find that gender differences in the actual timing of children and marriage are consistent with gender differences in beliefs about the timing of children and marriage, although both men and women are somewhat "optimistic" about how quickly these family outcomes will occur. More generally, this comparison of beliefs and actual outcomes represents a direct contribution to a small literature interested in using expectations data to examine the assumption of Rational Expectations.[8] While these types of comparisons are important because it is beliefs that affect decisions (Manski, 2004)[9], they are often not possible because much recent work using expectations data has involved researcher-led survey initiatives that, unlike the BPS, do not follow students for substantial periods of time.[10]

The important perceived role of children raises the question of whether the presence of young children has a causal effect on labor supply. Many researchers use an Instrumental Variables (IV) approach to deal with a well-recognized endogeneity problem that is present in this context: women who have young children at age 28 may differ from those who do not in ways that are related to their labor supply decisions. Previous studies have proposed several plausible IVs in this context. For example, Angrist and Evans (1998) constructs an IV based on whether the first two children in a family are of the same sex, exploiting the widely observed phenomenon of parental preferences for a mixed sibling-sex composition. Using administrative data from Denmark, Lundborg, Plug, and Rasmussen (2017) exploits the fertility variation among childless women induced by in vitro fertilization to estimate the effect of having children on females' careers.[11] However, given the nature of these IVs, these approaches will very often not be available to researchers using standard longitudinal data sources. An alternative strand of literature adopts an event-study framework to examine the effect of young children on female labor supply (see, e.g., Angelov, Johansson, and Lindahl, 2016, Kuziemko et al., 2018, Kleven

[7]Our investigation of whether the substantial "mommy effect" on labor supply was anticipated by females when making human capital investment decisions is closely related to Kuziemko et al. (2018), who examine whether women of relatively older cohorts (born in 1960s) in the UK and the US underestimated the employment effects of motherhood.

[8]We note that examining whether individuals have Rational Expectations is difficult since, for example, the distribution of a particular outcome may vary from earlier beliefs about that outcome because either beliefs are inaccurate or because aggregate shocks influence the outcome (D'Haultfoeuille, Gaillac, and Maurel, 2018).

[9]In support of this argument, Stinebrickner and Stinebrickner (2012) find that a simple theoretical implication related to college dropout - that the dropout decision should depend on both a student's cumulative GPA and beliefs about future GPA - is satisfied when beliefs are directly elicited through survey questions, but is not satisfied when beliefs are constructed under a version of Rational Expectations.

[10]A notable exception is Zafar and Wiswall (2016) who compare expectations about various future outcomes for undergraduate students at New York University in the year 2010 and realizations of these outcomes in the year 2016. Different from this paper, they do not have a focus on the comparison between perceived and actual conditional labor supply given various family outcomes. This is because the vast majority of the respondents were relatively young (average age 25) and were not married or did not have children in 2016.

[11]Other examples include Cruces and Galiani (2007), and Cristia (2008).

et al., 2019a/b). This literature uses the birth of the first child as the event and directly examines whether there is evidence of a divergence in labor supply trends in the years before childbirth.

We contribute to the literature by introducing the idea that, from a conceptual standpoint, expectations data have substantial promise for helping to address endogeneity concerns. We illustrate the promise of this idea by discussing the conceptual importance of including beliefs about future fertility in both cross-sectional and panel regressions of labor supply on children outcomes. Estimation is made possible by novel BPS data characterizing post-college beliefs about future fertility. Consistent with previous research using other methods, we find strong evidence that the actual birth of children has a large negative effect on the labor supply of college educated women, but not men.[12]

2 Berea Panel Study

The Berea Panel Study (BPS) is a longitudinal survey that was initiated by Todd and Ralph Stinebrickner to provide detailed information about the college and early post-college periods. Of particular importance here, the BPS was motivated by an explicit objective of exploring the potential promise of expectations data. Influenced by the early methodological work of, for example, Juster (1966), Dominitz (1998), and Dominitz and Manski (1996, 1997), the first BPS pilot took place in 1998.

Full cohorts of students who entered Berea College in the fall of 2000 and the fall of 2001 were surveyed approximately twelve times each year in college. In terms of in-school data, we primarily take advantage of expectations data that were collected at the halfway point of college - the beginning of the third year.[13] The details of these questions, which appear in Appendix A, will be discussed as they are encountered throughout the paper, but, briefly, the spirit of these questions is to allow individuals to express uncertainty about a labor market or family outcome (marriage or children) that may occur in the future.

We provide some information about students at Berea College, by showing some descriptive evidence about the in-college sample, which consists of students who answered

[12]The negative effect of young children on labor supply for highly educated/skilled females has been consistently documented in a large literature (see, for example, Silles, 2016 and Aaronson, et al., 2017). In contrast, using PSID data, Lundberg and Rose (2002) find that males may adjust their hours upwards in response to the birth of a child.

[13]We choose to use expectations about future labor supply from the beginning of the third year because this is the first point in college when these expectations were collected for both of the BPS cohorts. Taking advantage of the fact that the 2001 cohort also answered these questions in their sophomore year, we find in online Appendix B (Table B.1 to Table B.8) that, on average, perceptions are similar in the second and third years. Examining perceptions in the third year is also useful for making the in-school and post-college samples we use comparable. Our post-college sample focuses on individuals who graduated from college - since individuals who left Berea and did not continue their education elsewhere were followed for a shorter amount of time after college than graduates. Students who answered the survey in the third year were likely to graduate since the large majority of dropout occurs before the start of the third year (Stinebrickner and Stinebrickner 2014b).

survey expectations questions at the beginning of the third year. 265 out of the 418 students in the sample are female. On average, these male and female students have similar college entrance exam scores; the average score on the combined American College Test (ACT) is 23.2 for male students and 23.7 for female students. Students at Berea College are typically from families of relatively low socioeconomic status; the average annual family income is roughly $24,000 at the time of entrance.

Graduates from the BPS were surveyed annually for approximately ten years after graduation. We take advantage of detailed information that was collected about labor market and family outcomes (marriage and children). The BPS post-college surveys also continued to recognize the usefulness of expectations data. Here we take advantage of beliefs about future family changes.

Berea College, which is located in central Kentucky, is unique in certain ways that have been explored in previous work using the BPS. For example, Berea has a focus on providing educational opportunities for students from low income backgrounds. As part of this mission, Berea provides a full tuition subsidy to all students (Stinebrickner and Stinebrickner, 2003a), which is made possible, in part, by all students participating in a work-study program (Stinebrickner and Stinebrickner, 2003b). As always, it is necessary to be appropriately cautious about the exact extent to which the results from Berea would generalize to other demographic groups or to other specific institutions. However, important for the notion that the basic lessons from our work are pertinent for thinking about what takes place elsewhere, Berea operates under a standard liberal arts curriculum and the students at Berea are similar in terms of college entrance exams to students at the surrounding flagship state universities (Stinebrickner and Stinebrickner, 2008).[14] In addition, in earlier work we found that schooling and post-schooling outcomes at Berea look generally similar to decisions made elsewhere. For example, dropout rates are similar to those at the University of Kentucky (for students from similar income backgrounds), patterns of major choice and major-switching at Berea are similar in spirit to those found in the NLSY by Arcidiacono (2004), and average wages in the early part of the lifecycle are similar to those seen for students of similar age in the NLSY-97 (Stinebrickner and Stinebrickner, 2008, 2012, 2014a, 2014b).

[14]The average ACT score of students at Berea College is somewhat higher than the average ACT score, 22.0, of students in the NELS-88 who attended four-year colleges nationally (National Center for Education Statistics, 98-105). The 25th percentile of ACT scores in the BPS sample is 21 at Berea, and was also 21 at The University of Kentucky and The University of Tennessee at the time. The 75th percentile of ACT scores in the BPS sample is 25 at Berea College, and was 26 at The University of Kentucky and The University of Tennessee at the time. Not surprisingly, these comparisons become slightly more favorable for students at Berea if we condition on students in the NELS-88 or students at the regional universities who come from similar family income backgrounds.

3 Beliefs about Labor Supply and Family Outcomes

This section characterizes beliefs about future labor supply. In Section 3.1, we begin by characterizing male and female beliefs about labor supply over the lifecycle unconditional on family outcomes. We note that these "unconditional beliefs" will be affected by both beliefs about the timing of future family changes and beliefs about future labor supply conditional on family outcomes. As such, in the latter portion of Section 3.1 we describe male and female beliefs about these objects, and in Section 3.2 we provide a method to formally assess the relative importance of these objects in determining gender differences in beliefs about unconditional future labor supply.

3.1 Descriptive Statistics

As discussed in Section 2, our analysis focuses on beliefs elicited halfway through college. Taking advantage of Survey Question A (Appendix A), Table 1 reports beliefs about labor supply at three different future ages. We note in advance that, while the BPS expectations questions elicit perceptions about the percent chance that a particular outcome will occur, for reasons of expositional ease we often refer to the elicited information as a perceived probability. This slight abuse of language implies that our perceived "probabilities" take on values of 0 to 100 (rather than 0 to 1).

Panel A indicates that, for the full sample of 418 students, beliefs exhibit a substantial lifecycle pattern. Perhaps most notably, the first row of Panel A shows that, on average, the perceived probability of working full-time increases from 62.0% at the age of 23 to 72.1% at the age of 28 to 79.6% at the age of 38, with this lifecycle increase being consistent with workers facing more job insecurity and having a stronger incentive to experiment by changing jobs early in careers (Topel and Ward, 1992 and Gervais, et al., 2016). The middle row of Panel A shows that the lifecycle increases in the average perceived probability of working full-time are accompanied by lifecycle decreases in the average perceived probability of working part-time. Combining the full-time and part-time results, the last row of Panel A shows that the average perceived probability of not working at all decreases somewhat over the lifecycle, from 7.6% to 5.7%.

We are particularly interested in whether, and how, beliefs differ by gender. The first column of Panel B and Panel C shows that beliefs about labor supply at the age of 23 are strikingly similar for males and females. The average perceived probability of working full-time is 62.3% for males and 61.8% for females, and the average perceived probability of not working at all is 7.7% for males and 7.5% for females. However, column 2 shows that students anticipate that substantial gender differences in labor supply will emerge by the age of 28. The average perceived probability of working full-time increases by 19.3% (to 81.6%) for males between the ages of 23 and 28 but by only 4.8% (to 66.6%) for females, and the average perceived probability of not working at all decreases by 4.05% (to 3.7%) for males between the ages of 23 and 28 but increases by 1.31% (to 8.8%)

Table 1: Beliefs about Future Labor Supply

Probability (%)	Age 23	Age 28	Age 38
Panel A: Full Sample, # of Obs. = 418			
Full-time	62.01	72.11	79.57
	(30.21)	(25.01)	(24.63)
Part-time	30.41	20.96	14.77
	(25.48)	(18.44)	(17.32)
Not Working	7.58	6.93	5.66
	(14.56)	(14.41)	(12.24)
Panel B: Male, # of Obs. = 153			
Full-time	62.30	81.62	87.74
	(31.15)	(20.78)	(18.88)
Part-time	30.00	14.73	9.09
	(26.13)	(16.19)	(13.41)
Not Working	7.70	3.65	3.17
	(16.33)	(8.86)	(7.87)
Panel C: Female, # of Obs. = 265			
Full-time	61.84	66.62	74.85
	(29.65)	(25.60)	(26.27)
Part-time	30.65	24.56	18.05
	(25.09)	(18.71)	(18.44)
Not Working	7.51	8.82	7.10
	(13.44)	(16.50)	(13.97)

Note: Standard deviations are in the parentheses.

for females. Comparing column 2 to column 3 reveals that students do not anticipate that further gender differences in labor supply will emerge between the ages of 28 and 38; gender differences in the average perceived probability of working full-time and the average perceived probability of not working at all are actually slightly smaller for the age of 38 than for the age of 28.

Thus, Table 1 shows that there do not exist gender differences in perceptions about labor supply at the age of 23, and gender differences in perceptions about how labor supply will change over time are isolated to the period between the age of 23 and the age of 28. This suggests that family factors, which may be expected to change between these ages, may play a central role in determining the gender differences in perceptions about labor supply at the age of 28. From the standpoint of understanding this role, the first important question is whether females do indeed tend to believe that important family changes will take place at or before age 28. If so, gender differences in perceptions about labor supply at age 28 could arise from two alternative family-related explanations: 1) males tend to believe that family changes will tend to occur later for them or 2) there exist gender differences in perceptions about the relationship between family changes and labor supply.

To examine the first question of whether females tend to believe that important family changes will take place at or before age 28, we take advantage of Question B, which elicits the probability that an individual's first child will be born at various ages, and Question C, which elicits the probability that an individual will be married at various ages. The third column of the Marriage panel of Table 2 indicates that women in our sample tend to believe that marriage will take place relatively early; on average, female respondents believe there is only a 16.5% chance of either never being married or being married after age 30 (7.7% never married, 8.8% after age 30). Similarly, the third column of the First Child panel of Table 2 indicates that women in our sample tend to believe that they will have children at a relatively early age; on average, female respondents believe there is only a 23.1% chance of either never having children or having a first child born after age 30 (11.0% never children, 12.1% first child after age 30).

Table 2: Beliefs about the Timing of Family Outcomes

Probability at Each Age (%)	Marriage			First Child		
	All	Male	Female	All	Male	Female
Before 23	22.51 (32.04)	17.26 (27.13)	25.54 (34.2)	8.47 (19.36)	5.63 (13.32)	10.12 (21.94)
24 to 25	22.03 (20.83)	20.95 (21.49)	22.65 (20.42)	15.07 (17.08)	13.30 (16.17)	16.09 (17.51)
26 to 27	19.96 (17.51)	19.89 (17.22)	20.01 (17.68)	26.68 (21.70)	23.08 (20.85)	28.77 (21.91)
28 to 29	16.30 (17.27)	18.02 (18.38)	15.30 (16.51)	22.68 (18.94)	23.98 (18.81)	21.94 (18.98)
After 30	10.22 (14.65)	12.74 (18.22)	8.77 (11.87)	15.00 (18.83)	20.05 (23.14)	12.08 (15.07)
Never	8.98 (19.84)	11.15 (23.58)	7.73 (17.20)	12.09 (24.28)	13.97 (26.19)	11.01 (23.04)
# of Obs.	418	153	265	418	153	265

Note: Standard deviations are in parentheses.

Turning to the first alternative explanation above for how gender differences in perceptions about labor supply at age 28 could arise, Table 2 provides evidence that males do indeed tend to believe that family changes will occur somewhat later for them. On average, male respondents believe that there is a 23.9% chance of either never being married or being married after age 30 (11.2% never married, 12.7% after age 30), and a test of the null hypothesis that the average perception of this probability for males is less than or equal to that for females is rejected at all traditional levels (t-statistic= 3.19). Similarly, on average, male respondents believe that there is a 34.0% chance of either never having children or having a first child born after age 30 (14.0% never children, 20.1% first child after age 30), and a test of the null hypothesis that the average perception of this probability for males is less than or equal to that for females is also rejected at all traditional levels (t-statistic= 4.22).

While these gender differences in perceptions about the timing of marriage and children are quantitatively and statistically significant, a back-of-the-envelope calculation suggests that they are probably not, by themselves, large enough to account for the gender differences in beliefs about labor supply at age 28 that were seen in Panels B and C of Table 1 (middle column).[15] To examine the relevance of the second alternative explanation above for how gender differences in perceptions about labor supply could arise at age 28 - that there exist gender differences in perceptions about the relationship between family changes and labor supply - we take advantage of Survey Question D, which elicited perceptions about the probabilities of working full-time, working part-time, and not working at age 28 under the scenarios in which an individual is not married, married without a child, married with a youngest child between zero and two years old, and married with a youngest child between three and five years old.[16]

Table 3: Beliefs about Conditional Labor Supply at Age 28

Probability (%)	Unmarried	Own - Married			Spousal - Married		
		No Kids	Age 0-2	Age 3-5	No Kids	Age 0-2	Age 3-5
Panel A: Full Sample, # of Obs. = 418							
Full-time	85.52	83.69	60.02	67.34	80.23	68.19	71.27
	(18.79)	(19.09)	(32.48)	(30.78)	(21.84)	(30.37)	(29.24)
Part-time	11.99	13.49	25.43	20.73	15.39	20.90	19.13
	(15.53)	(15.78)	(21.20)	(18.57)	(17.09)	(20.07)	(18.56)
Not Working	2.49	2.82	14.56	11.93	4.38	10.91	9.60
	(6.43)	(6.60)	(23.82)	(21.45)	(9.22)	(20.56)	(18.41)
Panel B: Male, # of Obs. = 153							
Full-time	82.20	83.13	81.33	83.59	68.03	46.80	49.55
	(21.74)	(20.47)	(22.07)	(19.93)	(24.60)	(32.26)	(31.14)
Part-time	14.30	13.40	15.13	12.65	23.70	29.07	29.18
	(16.96)	(15.91)	(17.81)	(14.88)	(18.73)	(21.12)	(18.90)
Not Working	3.50	3.47	3.54	3.76	8.27	24.14	21.27
	(8.56)	(8.15)	(8.06)	(8.34)	(12.96)	(28.10)	(25.42)
Panel C: Female, # of Obs. = 265							
Full-time	87.44	84.01	47.71	57.95	87.27	80.54	83.82
	(16.55)	(18.24)	(31.14)	(32.01)	(16.35)	(20.92)	(18.93)
Part-time	10.65	13.54	31.37	25.40	10.60	16.18	13.32
	(14.47)	(15.70)	(20.72)	(18.88)	(13.97)	(17.80)	(15.64)
Not Working	1.91	2.45	20.92	16.64	2.13	3.27	2.86
	(4.68)	(5.47)	(27.33)	(24.99)	(4.82)	(7.17)	(6.14)

Note: Standard deviations are in parentheses.

Turning our attention first to the role of marriage, the first column of Table 3 shows

[15] For example, as reported in Table 2, the average perceived probability of having a first child before age 28 is about 10 percentage points smaller for males than for females. If males and females have the same perceptions about the how family changes will influence labor supply, then, even in the most extreme scenario where having a first child before age 28 leads to a 100 percentage points decrease in the probability of working full-time, this ten percentage point difference in beliefs about having a first child before age 28 would only produce a 10 percentage point difference in perceptions about the probability of working full-time at age 28. As shown in Table 1, this is less than the observed perceived difference of 15 percentage points.

[16] Our decision to not differentiate between unmarried with a child and unmarried without a child is driven by a practical consideration related to survey length. We note that only 6.2 percent of respondents in our sample had children but were not married at age 28.

that perceptions about labor supply at the age of 28 are very similar for men and women under the unmarried scenario. For example, for women, the sample average perceived probability of working full-time, 87.4%, is actually slightly higher than the corresponding sample average perceived probability for men, 82.2%. Further, comparing the results in the first column to the Married-No-Kids results in the second column, we see that both males and females believe that, when no kids are present, marriage will have virtually no relationship with the probability of working full-time or not working at all (or working part-time). Thus, our results strongly indicate that individuals do not believe that marriage per se will contribute to gender differences in labor supply.

However, perceptions about labor supply change dramatically when respondents are asked to consider the presence of a child. For example, looking across the first row of Table 3, the third column shows that, for the full sample, the average perceived probability of full-time work decreases from 83.7% to 60.0% when an individual has a youngest child between the ages of zero and two at age 28 and the fourth column shows that, for the full sample, the average perceived probability of full-time work decreases from 83.7% to 67.3% when an individual has a youngest child between the ages of three and five at age 28. Of particular note, this change arises almost entirely because of the perceived relationship between children and labor supply for women. The second, third, and fourth columns of Panel B show that males believe children will have virtually no relationship with their labor supply. In contrast, the second, third, and fourth columns of Panel C reveal that females believe that there is a strong negative relationship between the presence of a young child and labor supply. Specifically, a comparison of the second and third columns show that the average perceived probability of full-time work for females decreases from 84.0% to 47.7% when a female has a youngest child between the ages of zero and two at age 28, and a comparison of the second and fourth columns show that the average perceived probability of full-time work for females decreases from 84.0% to 58.0% when an individual has a youngest child between the ages of three and five at age 28.

Our survey questions also elicited perceptions about the probability of a future spouse working full-time, working part-time, and not working at age 28 under the scenarios in which he/she is married without a child, married with a youngest child between zero and two years old, and married with a youngest child between three and five years old. This information is important because, from a conceptual standpoint, students' investment decisions depend on both their beliefs about own labor supply and their beliefs about spousal labor supply (Lundberg, 1988, Chiappori 1992, Chiappori and Donni, 2011, and Erosa et al., 2017). Comparing the Spousal - Married columns in Panel C to the Own -Married columns in Panel B reveals that females' beliefs about spousal labor supply are strikingly similar to males' beliefs about their own labor supply, which, as described earlier, indicate that males anticipate being highly likely to work under any family situation. For example, on average, females believe that their spouses have an 80.5% chance of working full-time when they have a youngest child less than or equal to two years old, while, on average, males believe that they have an 81.3% chance of working full-time under this scenario. Similarly, comparing the Spousal - Married columns in Panel B to the Own - Married columns in Panel C shows that there are strong similarities between males' beliefs about spousal labor supply and females' beliefs about own labor supply, which, as described earlier, indicate that females anticipate being much less likely to work when they have young children. For example, on average, males believe that their spouses have a 46.8% chance of working full-time when they have a youngest child two years old or younger, while, on average, females believe that they have a 47.7% chance of working full-time under this scenario. These findings of a strong agreement between the beliefs of males and females serve as a validation of the notion that students' responses to the survey questions are meaningful.[17]

3.2 An Alternative Approach for Characterizing the Beliefs about Future Labor Supply

The middle column of Table 1 shows beliefs about labor supply at age 28 elicited directly using survey Question A.2 in Appendix A. In Section 3.2.1, we develop an "alternative" method to compute these beliefs, taking advantage of BPS information characterizing beliefs about labor supply under various family scenarios and BPS information allowing the characterization of beliefs about the probability of these scenarios occurring. We have two primary motivations for developing this method. As discussed in Section 3.2.2, the first primary motivation is that it allows us to quantify the individual contributions of the two alternative family-related explanations for gender differences in beliefs about labor supply that were raised in Section 3.1: 1) relative to males, females tend to believe that family changes will occur earlier and 2) relative to males, females tend to believe that they are less likely to work when family changes occur. A second motivation is that a comparison between beliefs constructed under this alternative method and the directly elicited beliefs in Table 1 can be considered a validation of the quality of our expectations data, with this validation exercise related to a formal method for addressing measureemnt error in Gong, Stinebrickner, and Stinebrickner (2019).

3.2.1 An Alternative Method for Computing Beliefs about Labor Supply

We let P_i^j denote student i's perception at the halfway mark of college about the probability of having work status j at age 28, where $j = F$, P, and N correspond to "Full-time Work", "Part-time Work," and "Not Working," respectively. Our alternative method for

[17] Roughly speaking, under the assumption that men expect to marry women that are similar to the women in our sample and vice versa, these strong similarities between males' reported beliefs about spousal labor supply and females' reported beliefs about own labor supply also provide suggestive evidence that males and females interpret these survey questions similarly.

computing P_i^j notes that P_i^j can be written as a function of beliefs about future family outcomes and beliefs about future work status j conditional on these outcomes, and takes advantage of BPS data of relevance for characterizing these beliefs. Given strong evidence in Section 3.1 that the labor supply perceptions of individuals are not influenced by marriage per se, we focus on children as the family outcome of interest. Recognizing that the influence of children on labor supply may depend on the age of the children, we characterize a person's children situation using the age of the person's youngest child. Denoting the age of the youngest child of person i at age 28 as a_i, where $a_i = 0$ if i does not have a child, and letting A_i represent the random variable describing i's subjective beliefs at the beginning of the junior year about a_i, P_i^j is given by:

$$P_i^j = \int (P_i^j | A_i = a_i) dF_{A_i}(a_i), j \in \{F, P, N\}, \tag{1}$$

where $F_{A_i}(a_i)$ and $P_i^j | A_i = a_i$, respectively, denote the cdf of A_i and student i's perception at the beginning of the junior year about the probability of work status j given that the realization of A_i is a_i.

Given $F_{A_i}(a_i)$ and $P_i^j | A_i = a_i$ it is straightforward to approximate the integral using, for example, standard simulation methods.[18] What is necessary is to describe how we characterize $F_{A_i}(a_i)$ and $P_i^j | A_i = a_i$ given the unique expectations data available in the BPS.

Beginning with the characterization of $P_i^j | A_i = a_i$, as discussed in Section 3.1, survey Question D (Appendix A) provides the relevant BPS information. The question elicits the perceived conditional probability of having work status j at age 28 given a particular family scenario. The set of scenarios includes being not married, married without a child, married with a youngest child between the ages of 0 and 2, and married with a youngest child between the ages of 3 and 5, and we denote the conditional probabilities associated with these scenarios as $P_i^{j,NM}$, $P_i^{j,NK}$, $P_i^{j,02}$, $P_i^{j,35}$, respectively.

What is needed is to characterize $P_i^j | A_i = a_i$ for all possible realizations a_i. We start by characterizing this conditional probability for realizations of A_i at the extremes. In practice, a person could have zero kids ($A_i = 0$) if either he/she is unmarried or is married but has no children. Given our earlier finding that perceptions about labor supply are virtually identical for these two scenarios, we approximate $P_i^j | A_i = 0$ by the average of $P_i^{j,NM}$ and $P_i^{j,NK}$. We denote this average $P_i^{j,N}$. Considering the other extreme, because it has been widely recognized that the effect of children on labor supply tends to decrease substantially when children attend school, we assume that children equal to or older than 6 years old do not affect labor supply. That is, $(P_i^j | A_i \geq 6) = (P_i^j | A_i = 0) = P_i^{j,N}$.[19]

[18] E.g., this integral can be approximated by computing the average value of $P_i^j | A_i = a_i$ for large number of random realizations a_i drawn from the distribution of A_i.

[19] For example, consistent with this assumption, Blau and Winkler (2018) document that, in year 2015, the labor participation rate of women with children under age 6 is roughly 10 percentage points lower than that of women with children between the ages of 6 and 17.

For characterizing the conditional probability $P_i^j|A_i = a_i$ for values of a_i that correspond to having a young child, the relevant observed information is: $P_i^{j,02}$ and $P_i^{j,35}$. The former represents the expected value of $P_i^j|A_i = a_i$ over the child ages in the set $(0, 3)$. The latter represents the expected value of $P_i^j|A_i = a_i$ over the child ages in the set $[3, 6)$. Then, under the simplifying assumption that $P_i^j|A_i = a_i$ is constant over $(0, 3)$ and is constant over $[3, 6)$, the perceived unconditional probability of work status j, P_i^j, is given by:

$$\begin{aligned} P_i^j &= \text{Prob}(A_i \in \{0\} \cup (6, \infty))P_i^{j,N} + \text{Prob}(A_i \in (0, 3))P_i^{j,02} + \text{Prob}(A_i \in [3, 6))P_i^{j,35} \\ &\equiv \sum_{k \in \{N,02,35\}} \pi_i^{A,k} P_i^{j,k}, j \in \{F, P, N\}, \end{aligned} \tag{2}$$

where $\pi_i^{A,N} \equiv \text{Prob}(A_i \in \{0\} \cup (6, \infty))$, $\pi_i^{A,02} \equiv \text{Prob}(A_i \in (0, 3))$, and $\pi_i^{A,35} \equiv \text{Prob}(A_i \in [3, 6))$.

In terms of characterizing $F_{A_i}(a_i)$, if we were to assume that each student expects to have at most one child, the distribution of A_i would come directly from a student's beliefs about the age of having the first child reported in Section 3.1. Relaxing this assumption requires that we take into consideration each person's beliefs about the age of having a second child or subsequent children. While it was not feasible for the BPS to collect this additional information directly, our approach can take advantage of survey question E, which elicited beliefs about the total number of children a person will have in his/her lifetime. Table 4 summarizes responses to this survey question. Both male and female students believe that, on average, they have more than a 60% chance of having more than one child in their lifetime. Comparing the second column to the third column shows that, relative to men, women believe they are more likely to have two or more children and less likely to have only one child or no child at all.

Table 4: Beliefs about the Number of Children

Probability (%)	All	Male	Female
0 Child	12.97	15.06	11.76
	(24.04)	(25.57)	(23.03)
1 Child	20.83	21.78	20.29
	(17.99)	(17.69)	(18.14)
2 Children	34.38	32.56	35.44
	(22.11)	(21.17)	(22.58)
3 Children	21.31	20.84	21.58
	(17.81)	(17.68)	(17.88)
$\geq$ 4 Children	10.51	9.77	10.94
	(16.85)	(14.65)	(17.98)
# of Obs.	418	153	265

Note: Standard deviations are in parentheses.

Formally, let $g_{i,q}$ denote student i's age when the qth child is born, and $G_{i,q}$ denote the random variable describing student i's beliefs about $g_{i,q}$ at the beginning of the junior year. As discussed earlier, $G_{i,1}$ can be directly elicited from survey Question B. To take advantage of survey Question E to estimate $G_{i,q}$, for $q \geq 2$, we begin by assuming that student i believes that, net of the 10 months ($\frac{5}{6}$ year) necessary for pregnancy, the age gap between having two consecutive children follows an exponential distribution with mean $\mu_{i,q}$. Formally, we have:

$$G_{i,q+1} - G_{i,q} - \frac{5}{6} \sim \text{Exp}(\mu_{i,q+1}). \tag{3}$$

We assume that students believe they will have no more than four children, and that children will not be born after age 40.[20] The value of $\mu_{i,q}$ can be computed from student i's beliefs about the number of children he/she will have and information on $G_{i,1}$. Our approach is detailed in online Appendix C. In terms of the intuition underlying the approach, if student i believes, for example, that he/she will have the first child relatively early but that he/she is not likely to have more than one child, the value of $\mu_{i,2}$ would have to be large so that the second child is unlikely to arrive before the age of 40. Similarly, if student i believes that he/she will have the first child relatively late but still expects to have more than one child with high probability, the value of $\mu_{i,2}$ would have to be small. The same intuition can be applied to the computation of $\mu_{i,3}$ and $\mu_{i,4}$.

We compute P_i^j using Equation (2). With the distributions of $G_{i,1}$ and $G_{i,q+1} - G_{i,q}$, $q = 1, 2, 3$, computed using the method described above, we employ a simulation-based method to approximate the terms involving A_i in Equation (2). Specifically, for each student, we simulate his/her fertility history a large number of times and use these simulated histories to approximate the perceived probability of having the youngest child in given age ranges at age 28.

3.2.2 Decomposing the Gender Difference in Beliefs about Labor Supply

Table 1 showed that, relative to males, females believe they are less likely to be working full-time at age 28 and more likely to not be working at all at age 28. The descriptive statistics in Section 3.1 indicated that two child-related explanations are likely relevant for understanding this gender difference: 1) females believe they are more likely to have a young child at age 28 and 2) females believe they are less likely to work when they have a young child. Here we perform a decomposition, which takes advantage of the method we developed in Section 3.2.1 to quantify the relative importance of these two explanations.

For ease of illustration, we focus on a representative male student ($i = M$) and a representative female student ($i = F$). The beliefs of a representative student about a particular object of interest are found by averaging the beliefs of all students of the same gender about that object.

We first use Equation (2) to compute beliefs of the two representative students $i =$

[20] Using 35 or 45 instead yields very similar results.

M, F about unconditional labor supply, P_i^j, for $j = F, P, N$. This computation requires beliefs about conditional labor supply, $P_i^{j,N}$, $P_i^{j,02}$, and $P_i^{j,35}$, and beliefs about the age of the youngest child at age 28, A_i. Beliefs about conditional labor supply for the representative students are given by the averages in Panel B and Panel C of Table 3. Beliefs about the age of the youngest child at age 28 can be computed using the method described in the latter part of Section 3.2.1, which utilizes information in Table 2 and Table 4. We find that the representative male student believes that, at age 28, he has 37.46% chance of having a youngest child between the ages of 0 and 2 ($A_M \in (0, 3)$), a 10.56% chance of having a youngest child between the ages of 3 and 5 ($A_M \in [3, 6)$), and a 51.98% chance of either not having a child ($A_M = 0$) or having a youngest child of age 6 or older ($A_M \in (6, \infty)$). Similarly, we find that the representative female student believes that, at age 28, she has a 45.25% chance of having a youngest child between the ages of 0 and 2 ($A_F \in (0, 3)$), a 15.27% chance of having a youngest child between the ages of 3 and 5 ($A_F \in [3, 6)$), and a 39.48% chance of either not having a child ($A_F = 0$) or having a youngest child of age 6 or older ($A_F \in (6, \infty)$).

The results of our computation using Equation (2) show that the (subjective) probabilities of working full-time, working part-time, and not working at age 28 for the representative male student are 82.02%, 14.44%, and 3.54%, respectively, while these probabilities for the representative female student are 64.96%, 22.28%, and 12.76%, respectively. Overall, these numbers are quite similar to the average directly elicited (subjective) probabilities of working full-time, working part-time, and not working for males and females seen in the second column of Panel B and Panel C of Table 1. Then, the numbers show that the computation method is capable of producing the large gender difference in beliefs about unconditional labor supply at age 28 that is observed in the directly elicited expectations data. Thus, our results represent a joint validation of the survey questions and the assumptions that are necessary for the alternative method used in the decomposition, with this validation exercise relating directly to the analysis in Giustinelli and Shapiro (2019).[21] The spirit of this validation also relates closely to a formal method for addressing measurement error proposed and implemented in Gong, Stinebrickner, and Stinebrickner (2019). When beliefs about the same object of interest can be computed using two different sets of expectations questions, differences in the computed beliefs reflect measurement error in the underlying expectations questions. Then, the amount of measurement error can be quantified given assumptions about the

[21] Giustinelli and Shapiro (2019) construct measures of beliefs about unconditional labor supply from the beliefs about labor supply conditional on health and beliefs about health. They compare these computed unconditional beliefs to the self-reported unconditional beliefs about labor supply.

manner in which measurement error enters the expectations questions.[22]

The numbers in the previous paragraph indicate that the representative female student believes she is 17.1% less likely than a representative male student to be working full-time at the age of 28. When we recompute the probability of working full-time for the representative female after replacing her beliefs about the age of the youngest child at age 28, A_F, by the beliefs of the male student, A_M, the gender difference is reduced to 12.6%. Thus, gender differences in beliefs about the timing of children can explain roughly 26% of the gender difference in beliefs about the probability of working at age 28. The remaining 74% is explained by gender differences in beliefs about working for the different possible child scenarios. We find similar results when we decompose gender differences in the probability of not working at all; gender differences in beliefs about the timing of family outcomes can explain roughly 24% of the gender difference of 9.2% in beliefs about the probability of not working at all at age 28, with the remaining 76% explained by gender differences in beliefs about working for the different possible child scenarios.

4 Actual Labor Supply and Family Outcomes at Age 28 and Comparison with Beliefs

Because individual decisions are based on beliefs at the time of decision-making, whether beliefs tend to be accurate is of importance for a variety of policy reasons. Using annual post-college surveys, the BPS collected outcomes related to labor supply and family for roughly ten years after graduation, past the age of 30. In Section 4.1, we compare the average reported probability of having a particular outcome to the fraction of individuals that have that outcome. A particular focus of this section is to examine whether gender differences in outcomes are consistent with the gender differences in beliefs uncovered in Section 3. In Section 4.2, we note the additional benefits of examining whether individual-level perceptions are strong predictors of individual-level outcomes, and take advantage of the fact that the BPS is relatively rare in allowing this type of examination.

[22] Online Appendix D shows results for this context obtained using a similar approach. Gong, Stinebrickner, and Stinebrickner, who study beliefs about future income, make an assumption that measurement error is classical. A literature on rounding in responses to survey expectations questions is relevant for considering the appeal of this assumption (see, e.g., Manski and Molinari, 2010, Giustinelli, Manski, and Molinari, 2019). We provide evidence that some rounding is present in students' responses to the survey questions used in this paper. For example, around 30% of the respondents reported a perceived full-time work probability that is either 0%, 50%, or 100%. Similarly, around 27% of the respondents reported a perceived part-time work probability that is either 0%, 50%, or 100%. Thus, it is worthwhile to view our measurement error results with appropriate caution.

4.1 Average Beliefs and Outcomes

Our in-school survey elicited beliefs about labor supply, marriage, and children at the age of 28 (among other ages). Our annual post-college survey allows us to characterize actual labor supply, marriage, and children outcomes at this same age. In terms of characterizing labor supply outcomes, students report whether they are currently working (Question G in Appendix A) and the number of hours they work (Question H). We assume that a student is working full-time if he/she is currently working 35 or more hours per week, and is working part-time if he/she is working less than 35 hours per week. Marital status comes directly from Question I. The age of a respondent's youngest child comes from questions asking whether a respondent currently has at least one child, and if so, the age at which the youngest child was born (Question J). While in earlier sections we distinguish between children who are between 0 and 2 years of age and children who are between 3 and 5 years of age, in this section, for reasons related to sample size, we combine these categories.

As seen in Table 1, 418 (153 male, 265 female) students answered the labor supply, marriage, and children expectations questions that we utilize from the halfway point of college. The first column of Table 5 shows the average perceived probabilities associated with a variety of outcomes for this sample. This information is generated using the same survey questions as in Section 3, with, in some cases, the information in Table 5 being repeated from earlier tables to ease comparisons. 460 respondents (158 male, 302 female) answered the labor supply, marriage and children questions characterizing their outcomes at age 28.[23] The second column of Table 5 shows the actual fraction of this sample that has each particular outcome. To explore the potential concern that the samples in Column 1 and 2 might not be entirely comparable due to selection issues, Columns 3 and 4 repeat Columns 1 and 2 for the 317 individuals that appear in both of the samples. A comparison of the first column with the third column shows that the sample average perceived probabilities are almost identical in the two samples, and for none of the outcomes is it possible to reject at a 5% level the null hypothesis that the average perceived probabilities are the same for the two columns. Similarly, a comparison of the second column with the fourth column shows that the sample fractions are almost identical in the two samples, and for none of the outcomes is it possible to reject at a 5% level the null hypothesis that the fractions are the same in the populations associated for the two columns. Given these results, in the remainder of this section we exploit the benefits of using as many observations as possible by performing comparisons based on the samples present in Columns 1 and 2.

In Table 5, comparing the last entry in the first column of Panel B to the last entry in the first column of Panel C shows that, as seen earlier in Table 1, the average perceived

[23]The post-college sample is larger, in part, because participation on the BPS baseline survey was a necessary condition for participation in subsequent in-school surveys, but was not a necessary condition for participation in post-college surveys.

probability of working full-time at age 28 is 66.6% for women and 81.6% for men. Of primary interest in this section is whether there actually exists a substantial gender difference in the full-time outcome. Comparing the last entry in the second column of Panel B to the last entry in the second column of Panel C shows the fraction of respondents working full-time at age 28 is 72.0% for women and 81.0% for men. Thus, on average, both men and women have quite accurate beliefs about full-time work, and, as a result, the gender difference in the fractions of men and women working full-time at age 28, $9.0\% = 81.0\% - 72.0\%$, is similar in spirit to the gender difference in the perceived probabilities of working full-time at age 28, $15\% = 81.6\% - 66.6\%$.[24] A generally similar result is obtained when we examine the outcome of working at all at age 28 (full-time or part-time). There exists a gender difference of 5.5% (female 84.4%, male 89.9%) in the fractions of men and women working at age 28, while there exists a gender difference of 5.2% (female 91.2%, male 96.4%) in the average perceived probabilities of working at age 28.

Turning to the family variables, we find that beliefs about marriage and children are not as accurate as beliefs about labor supply. For example, consistent with some evidence in Wiswall and Zafar (2016), the first row of Table 5 shows that the combined male-female sample is considerably optimistic about the probability of being married at age 28 (average perceived probability 72.7%, actual fraction married 52.4%) and the second row shows that the combined male-female sample is optimistic about the probability of having a youngest child five years old or younger at age 28 (average perceived probability 55.7%, actual fraction with child five years old or younger 29.4%).[25] Our results are broadly consistent with the findings in Giustinelli and Shapiro (2019) that respondents are better at predicting labor supply than the conditioning variable (health).

However, Panels B and C reveal that there exist important gender differences in the actual timing of children, which are in line with gender differences in perceptions about the timing of children. Specifically, women are 9.0% more likely than men to have a young child at age 28 (32.5% female, 23.4% male), and, on average, believe they are 13.1% more likely than men to have a young child at age 28 (60.5% female, 47.4% male). The gender difference in the fractions of men and women that are married at age 28 (52.7% female, 51.9% male) and the gender difference in the perceived probabilities of being married at age 28 (female 75.9%, male 67.1%) are both smaller than their children counterparts.

Given our finding that women believe that a young child is associated with substan-

[24] We cannot reject the null that the average gender difference in the actual fraction of men and women working full-time is the same as the average gender difference in the perceived probabilities of working full-time at a 10% level.

[25] In the study of NYU students by Wiswall and Zafar (2016), a complication would arise when comparing perceptions about family outcomes to actual family outcomes because the age at which perceptions are elicited is not the same as the age 25, for which family outcomes are observed. However, they are able to establish that respondents are too optimistic about marriage because they find that the fraction of their sample that is married (or has children) at age 25 is extremely low, and, therefore, lower than the average perceived probability of being married at a younger age.

tially lower labor supply, an open question is why the overoptimism about the timing of children seen in Table 5 does not lead women to substantially understate the actual probability of working full-time or working at all. One possibility is that the probability of working conditional on having a young child is smaller than women anticipated. We find some evidence that this is the case. Specifically, the last row of Table 6 shows that the average perceived probability of working conditional on having a young child is 74.5% for females, while the fraction of women working with young children is 65.4%. Moreover, looking across the other family scenarios in Panel C of Table 6 reveals that the optimism about the probability of working is not isolated to the young-child scenario, but, rather, is seen for all family scenarios. The results for full-time work reported in the Full-time columns of Panel C of Table 6 are generally similar with regard to optimism - on average, women have quite accurate beliefs about the probability of working full-time in the "Married with Young Children" scenario, but over-estimate the probability of working full-time in the "Unmarried" and "Married without Children" scenarios.

4.2 Individual Beliefs and Outcomes

The previous subsection compared the sample average perceived probability of a particular outcome occurring in the future to the fraction of respondents in the sample that have that outcome occur in the future. While Manski (2004) suggests the value of this type of full-sample comparison for characterizing the accuracy of beliefs, the relatively rare ability of the BPS to examine whether an individual's expectations are predictive of his/her own future outcomes is of additional usefulness. Some recent research has noted that such an examination is valuable because it provides evidence about whether expectations questions can indeed be successful in eliciting useful individual-level information about beliefs. However, given that the value of expectations data now seems to be widely accepted, here we focus on the benefit that, under the assumption that beliefs can be correctly characterized, the examination of the predictive ability of individual-level expectations allows a different, stronger test of Rational Expectations than is possible by the comparison between sample-level perceptions and sample-level outcomes.

Specifically, consider a regression in which the dependent variable is an indicator variable that takes the value one if a particular outcome is observed to occur for person i at some time t and the independent variable is i's perceived probability (during school) of the outcome occurring at the future time t.[26] If there is no aggregate shock that affects the actual outcome, the existence of Rational Expectations would imply that the constant in this regression would have a value of zero and the coefficient on the perceived probability would have a value of one. From an intuitive standpoint, this is the case because it corresponds, roughly speaking, to a situation where, for any subgroup of the

[26]The error term in this regression represents outcome-influencing factors that were not observed by student i when perceived probabilities were elicited. By construction, they are uncorrelated with perceived probabilities.

Table 5: Comparing Average Beliefs with Average Actual Outcomes at Age 28

Probability (%)	All Observations		Same Sample	
	Beliefs	Outcomes	Beliefs	Outcomes
Panel A: Full Sample				
Married	72.65	52.39	72.94	51.42
	(1.33)	(2.33)	(1.53)	(2.81)
Having a child between ages of 0 and 5	55.67	29.35	55.14	29.02
	(1.50)	(2.12)	(1.73)	(2.55)
Working	93.07	86.30	92.33	85.80
	(0.71)	(1.60)	(0.89)	(1.96)
Full-time	72.11	75.10	72.10	74.05
	(1.23)	(1.89)	(1.42)	(2.31)
# of Obs	418	460	317	317
Panel B: Male				
Married	67.11***	51.90	65.66***	48.11
	(2.40)	(3.97)	(3.00)	(4.85)
Having a child between ages of 0 and 5	47.39***	23.42**	46.14***	22.64*
	(2.53)	(3.37)	(3.04)	(4.06)
Working	96.35***	89.87*	96.03***	89.62
	(0.72)	(2.40)	(0.98)	(2.96)
Full-time	81.62***	81.01**	82.32***	79.25*
	(1.69)	(2.97)	(2.17)	(3.77)
# of Obs	153	158	106	106
Panel C: Female				
Married	75.85***	52.65	76.60***	53.08
	(1.54)	(2.87)	(1.68)	(3.44)
Having a child between ages of 0 and 5	60.46***	32.45**	59.66***	32.23*
	(1.80)	(2.69)	(2.03)	(3.22)
Working	91.18***	84.44*	90.48***	83.89
	(1.02)	(2.09)	(1.23)	(2.53)
Full-time	66.62***	71.99**	66.55***	71.42*
	(1.58)	(2.41)	(1.74)	(2.89)
# of Obs	265	302	211	211

Note: Standard errors are in parentheses. *** indicates that a test of the null hypothesis that a particular average for males (in Panel B) is the same as the corresponding average for females (in Panel C) is rejected at a 0.01 level of significance. ** and * indicate significance at 0.05 and 0.1, respectively.

Table 6: Average Perceived and Actual Labor Supply Probabilities

Probability (%)	# of Obs.	Outcomes		# of Obs.	Beliefs	
		Working	Full-time		Working	Full-time
Panel A: Full Sample						
Unmarried	219	89.50	76.58	154	97.01	83.65
		(2.07)	(2.76)		(0.46)	(1.50)
Married without Children	130	91.54	82.31	85	98.18	86.48
		(2.44)	(3.23)		(0.61)	(1.97)
Married with Young Children	111	73.87	63.72	78	80.85	63.31
		(4.17)	(3.90)		(2.66)	(3.47)
Panel B: Male						
Unmarried	76	88.16	76.32	55	95.47**	78.42**
		(3.71)	(4.64)		(1.09)	(2.97)
Married without Children	49	89.80	81.63	30	98.00	85.43
		(4.32)	(5.26)		(1.32)	(3.79)
Married with Young Children	33	93.94***	90.91***	21	98.12***	88.80***
		(4.15)	(4.81)		(1.74)	(4.47)
Panel C: Female						
Unmarried	143	90.21	76.71	99	97.87**	86.57**
		(2.49)	(3.43)		(0.38)	(1.63)
Married without Children	81	92.59	82.72	55	98.27	87.02
		(2.91)	(4.09)		(0.62)	(2.25)
Married with Young Children	78	65.38***	52.04***	57	74.48***	49.54***
		(5.39)	(4.70)		(3.54)	(3.94)

Note: The second and the third columns report the fraction of students who were working at all and working full-time at age 28 given each of the family outcomes. Similarly, the fifth and sixth columns report the average perceived conditional probability of working at all and working full-time at age 28 given a certain family outcome for students who actually had that family outcome at age 28. The sample sizes are reported in the first and fourth columns. Standard errors are in the parenthesis. *** indicates that a test of the null hypothesis that a particular average for males (in Panel B) is the same as the corresponding average for females (in Panel C) is rejected at a 0.01 level of significance. ** and * indicate significance at 0.05 and 0.1, respectively.

sample that has the same perceived probability, the fraction in the subgroup for which the outcome occurs is equal to the perceived probability. For a variety of reasons, including the concern that labor supply and family outcomes might be affected by aggregate shocks to some extent, we do not wish to take this exercise too literally, but, instead note that the closer the slope coefficient is to one (and the closer the constant is to zero), the more correct the students' beliefs tend to be in a Rational Expectations sense.

The first panel of Table 7 shows results obtained by regressing an indicator for whether a respondent is married at the age of 28 on the respondent's perceived probability (during school) of being married at the age of 28. Similarly, the second panel of Table 7 shows the results from regressing an indicator for whether a respondent has a young child at age 28 (5 years old or less) on the respondent's perceived probability of having a young child at age 28.[27] The results in the second row of these two panels indicate that perceptions are strong predictors of actual family outcomes. A one percentage point increase in the perceived probability of being married and having a young child at age 28, respectively, is associated with a 0.66 and 0.56 percentage point increase in the actual probability of being married and having a young child at age 28, respectively. These estimates are significant at a 1% level, and, as seen in the third row of the first two panels (Table 7) lead to correlations of 0.36 and 0.38. Further, as shown in the first row of Table 7, we cannot reject the null hypothesis that the constant term in either of the regressions is 0, even at a 10% significance level. Thus, although Section 4.1 shows that students are, on average, incorrect in their beliefs about marriage and children, we find that their misperceptions are likely to take a relatively simple form; the actual probability of being married at age 28 is only about two-thirds of what each student believes and the actual probability of having a young child at age 28 is only a little more than one-half what each student believes.

Table 7: Regression of Actual Outcomes at Age 28 on Perceived Probabilities of Outcomes

	Married	Young Child	Working	Full-time
Constant	0.035	-0.018	0.634	0.631
	(0.075)	(0.049)	(0.112)	(0.070)
Probability	0.66	0.559	0.242	0.156
	(0.097)	(0.077)	(0.120)	(0.093)
Correlation	0.3554	0.3784	0.1117	0.0932
# of Obs	328	328	327	325

Note: Standard errors are in parentheses. Probability is the perceived (during school) percent chance (0 to 100) of the outcome in a particular column occurring at age 28.

Turning to results related to labor supply in the last two panels of Table 7, the first column of each of these panels shows that a one percentage point increase in the perceived

[27]The perceived probability of being married at age 28 is constructed using a student's reported probability of getting married at each age under the assumption that divorce does not happen before age 28 and that the probability of getting married at age 28 is the same as that at age 29. The perceived probability of having a young child is computed using the method detailed in Section 3.2.1.

probability of working at all and working full-time, respectively, is associated with a 0.24 and 0.16 percentage point increase in the actual probability of working at all and working full-time at age 28, respectively. These estimates are significant at a 5% level and a 10% level, respectively, and, as seen in the last row of Table 7, lead to correlations of 0.11 and 0.09.

The fact that the correlation between perceptions and outcomes is lower for the labor supply outcomes than for the family outcomes may be a bit surprising at first glance given that Section 4.1 found that, on average, perceptions about labor supply are somewhat more accurate than perceptions about the family outcomes.[28] But, of course, there is nothing internally consistent about this finding. Generally speaking, this correlation pattern will tend to occur if there exist relevant factors in the determination of labor supply that are not accounted for by the respondents at the time that perceptions are elicited, while these types of unaccounted-for factors are not as prevalent in case of the family outcomes. For example, differences in beliefs about children (or marriage) might be determined primarily by differences in preferences about how much a person enjoys being around children, while differences in beliefs about being out of the workforce (which are strongly related to children) might be determined primarily by differences in feelings about the extent to which young children benefit from spending time with a parent. Then, the correlation pattern would arise if, for example, people with preferences to not have kids are able to control this entirely while people who wish to not work (when they have kids) find that this is often not financially feasible given, for example, the labor market outcomes of their spouses.[29]

5 Estimating the Effect of Young Children on Labor Supply

A large traditional literature has been interested in quantifying the causal effect of young children on the labor supply of women, and, to a lesser extent, on men (see e.g., Nakamura and Nakamura, 1992 for a survey of early literature and Angrist and Evans, 1998, Cruces and Galiani, 2007, Cristia, 2008, and Lundborg, Plug, and Rasmussen, 2017 for more recent investigations). Of relevance for thinking about this issue for women, our results

[28] We have also performed these regression-based analysis separately by gender. For both males and females, we find a generally similar pattern that the association between beliefs about family outcomes and actual family outcomes is stronger than the association between beliefs about labor supply and actual labor supply. Looking across genders, we also find that females' expectations are more "rational" in the sense that, in the regressions for almost all the outcomes, the estimate of the intercept parameter is closer to zero, and the estimate of the slope parameter is closer to one for females.

[29] We find some informal evidence in support of this general notion that individuals may have substantially less "control" than they think about labor market outcomes, relative to family outcomes. Among individuals who have a strong belief that they will not be working full-time at age 28 (less than a 20% chance of working full-time), only 35% are actually not working full-time at age 28. However, among individuals who have a strong belief that they will not have children at age 28 (less than a 20% chance of having children), 86% do not have a child at age 28.

in previous sections indicated that, consistent with what was expected by students when they were in college, females tend to work much less at age 28 when they have a young child (while males do not). However, the interpretation of this relationship is complicated by a potential, well-recognized endogeneity problem: women who have young children at age 28 may differ from those who do not in ways that are related to their labor supply decisions. In theory, this endogeneity issue can be addressed by a simple cross-sectional regression if all relevant differences between women with and without young children can be controlled for by observable characteristics. Unfortunately, it may be difficult to find observable characteristics that are able to credibly control for differences in, for example, unobserved preferences for leisure or for spending time with young children, which would tend to be correlated with both fertility status and labor supply.

Here we introduce the idea that, from a conceptual standpoint, expectations data have substantial promise for helping to address endogeneity concerns in the estimation of the actual effect of children on labor supply.[30] To begin to illustrate this promise, we first focus entirely on one well-recognized reason that women with children at age 28 may be different than women without children - that they may invest differently in human capital at earlier stages if they anticipate that they are more likely to have children in the future and if they believe that children are related to time away from the labor force. Our results in Section 3.1, which show that women believe that labor supply is related to children, and our results in Section 4.2, which show that women are able to predict in advance whether they will have children, provide some of the first direct evidence that this type of endogeneity may indeed be of relevance. The empirical difficulty arises, in practice, because it is virtually impossible to perfectly measure differences in human capital investments between women with and without children. In addition to the obvious problem of fully characterizing all formal training and schooling that a worker receives while in the workforce, the learning-by-doing that takes place on a job is likely to be a complicated process that depends on the exact set of tasks being performed.[31]

Our insight is that, under our illustrative scenario, it is sufficient to control for the beliefs about children (rather than direct measures of earlier human capital investments) because it is differences in these underlying beliefs that generate differences in human capital investments. The BPS makes this approach possible because, in each post-college year, the survey elicited beliefs about the probability of children being born at each future age.

Table 8 shows, separately for females and males, a regression of actual labor supply outcomes at age 28 on indicator variables for marriage and the presence of children younger than age 5, measures of human capital at the time of college graduation (ACT

[30]Our analysis is loosely related to a growing set of papers using survey expectations to characterize and estimate subjective ex ante causal/treatment effects (see, e.g., Wiswall and Zafar, 2016, Arcidiacono et al., 2019, and Giustinelli and Shapiro, 2019).

[31]For example, Sullivan, Stinebrickner and Stinebrickner (2019) find that tasks performed in the past are stronger predictors of current wage earnings, which are measures of human capital levels.

Table 8: Cross-sectional Regression

	Female		Male	
	Working	Full-time	Working	Full-time
Marriage: Outcome	0.025	0.065	-0.010	0.013
	(0.056)	(0.071)	(0.068)	(0.090)
Marriage: Beliefs	0.046	-0.042	0.005	0.026
	(0.115)	(0.146)	(0.121)	(0.160)
Young Children: Outcome	-0.194	-0.282	-0.061	0.070
	(0.056)	(0.071)	(0.073)	(0.097)
Young Children: Beliefs	-0.260	-0.082	0.273	0.210
	(0.108)	(0.138)	(0.118)	(0.157)
# of Obs.	134	134	278	275

Note: Standard errors are in parentheses.

score, high school GPA, and college GPA), post-college beliefs about the probability of being married and having young children at age 28, and demographic characteristics (gender and race). We examine both the outcome of working at age 28 (either part-time or full-time) and the the outcome of working full-time at age 28. In this specification, we use average perceived probabilities from the first two post-college years as measures of post-college beliefs.[32]

The results for females are in the first two columns of Table 8. Starting with the post-college belief variables, the last row of the table shows that, all else equal, women who, early in their career, anticipate having children at age 28 are significantly less likely to work at age 28. This is consistent with, among other things, the motivating notion that women who expect to have children choose to invest less in human capital early in their careers, and, therefore, have less incentive to work at age 28. Turning to the variables of primary interest, we find that actual marital status at age 28 does not have a significant effect on actual labor supply outcomes at age 28. However, the presence of young children has substantial effects. For example, the third row of Column 1 shows that having a young child at age 28 decreases the probability of working by 19.4 percentage points (t-stat 3.47) and the third row of Column 2 shows that having a young child at age 28 decreases the probability of working full-time by 28.2 percentage points (t-stat 3.95). Results for males in last two columns of Table 8 indicate that men who anticipate having children are significantly more likely to work at age 28, but that neither marriage or children affect labor supply in a significant manner.

We note that, because there likely exist other sources of endogeneity in this context, we believe that the preceding paragraphs are best viewed as an illustrative discussion of the potential benefits of expectations for addressing endogeneity concerns. For example,

[32] Ideally, it would be desirable to include perceived probabilities in all post-college years. However, this is empirically difficult because the number of respondents who appear in every wave of the post-college surveys is small. We have considered alternative specifications where we use average perceived probabilities from the first post-college year and from the first three post-college years. The main results remain qualitatively similar.

one obvious concern is that women who tend to have a strong preference for leisure (or a strong dislike for work) may be more likely to have children. To the extent that these types of preferences are permanent in nature, it is natural to exploit within-person variation using a Fixed Effects estimator. This type of specification would no longer require the type of across-person variation in beliefs that are utilized in Table 8. However, it does highlight another endogeneity concern that can also potentially be addressed using expectations information. If individuals in period t anticipate having a child in period $t+1$ and believe that this will influence their labor supply in period $t+1$, they may have an incentive to change their labor supply in period t, to, for example, make up for lost income in $t+1$. This implies that changes in labor supply tend to be related to changes in beliefs about future family outcomes, which are typically correlated with actual family outcomes. Then, an endogeneity issue might be present when beliefs are unobserved.

This endogeneity concern can be addressed using expectations information if, as in the BPS, beliefs at time t about $t+1$ can be included in the Fixed Effects specification. Formally, we consider a specification in which current labor supply outcomes (i.e., working or working full-time) depend on current family outcomes (i.e., marriage and children), beliefs about family outcomes in the next year (i.e., getting married and having children), other observed characteristics (age, gender, race), and both unobserved permanent (fixed effect) and transitory factors. We note that it is probably most prudent to continue to view our results as illustrative. For example, an endogeneity concern would remain if people tend to have children if they know that they might lose their job in the future or if they anticipate other reasons that their job may not seem as rewarding in the future (e.g., changes in compensation or working conditions).

Table 9: Fixed Effects Estimates

	Female		Male	
	Working	Full-time	Working	Full-time
Marriage: Outcome	-0.043	-0.013	-0.041	-0.060
	(0.032)	(0.038)	(0.036)	(0.045)
Marriage: Beliefs	-0.035	-0.066	0.005	-0.082
	(0.061)	(0.074)	(0.072)	(0.091)
Young Children: Outcome	-0.189	-0.262	-0.003	-0.030
	(0.025)	(0.031)	(0.032)	(0.040)
Young Children: Beliefs	0.088	0.019	-0.006	0.008
	(0.043)	(0.052)	(0.053)	(0.066)
# of Obs.	2,102	2,094	1,093	1,089

Note: Standard errors are in parentheses.

The Fixed Effects estimator is implemented using annual survey data collected from 2007 to 2013. The first two columns of Table 9 report the estimation results for females. Beginning again with beliefs, the last row in the Working column indicates that a one percentage point increase in the probability of having a child in the next year increases the probability of working at all by 0.09 percentage point, with the t-statistic having a value of 2.05. Thus, there is evidence that women "expecting" children may wish to be in the workforce in the current period to compensate for the future decrease in labor supply that often accompanies children. Looking at the estimated effect of beliefs about children in the Full-time column, we find that beliefs about future child outcomes do not have a statistically significant effect on the probability of working full-time. Thus, the work adjustment seemingly operates through the addition of part-time jobs. An additional regression using part-time work as the dependent variable provides evidence that this is case.[33] The second row shows that beliefs about marriage do not have a significant effect on the labor supply of women.

Turning again to the family variables of primary interest, the first and third rows of Table 9 show results that are very similar to the cross-sectional results in Table 8. Actual marital status at age 28 does not have a significant effect on actual labor supply outcomes at age 28, but the presence of a young child decreases the probability of a woman working (either full-time or part-time) by 18.9 percentage points (t-stat 7.46) and decreases the probability of working full-time by 26.2 percentage points (t-stat 8.50).

The results for males are shown in the last two columns of Table 9. Broadly consistent with the other results in the paper, we find that labor supply outcomes are not affected by either beliefs about marriage and children or actual marriage and children outcomes.

6 Conclusion

From a conceptual standpoint, human capital investment decisions in college, such as those related to study effort, dropout, and major, are influenced by students' beliefs about future labor market attachment. This paper provides a comprehensive analysis of these beliefs, with a focus on gender differences, taking advantage of expectations data from the Berea Panel Study.

Our results suggest that, on average, both men and women are quite well-informed about their future labor supply. This implies that the difference between the average perceived probabilities of working for men and women is similar to the difference between the fractions of men and women observed to be working in the post-college data. We employ a decomposition to investigate why women believe they are less likely to be working at age 28 than males. While the fact that women (correctly) believe that they will have children earlier than men plays some role, the large majority of the gender difference arises because women (correctly) believe that they will be less likely to be working when they have young children.

The paper makes two primary methodological contributions to the expectations literature. First, it explores approaches to validate the quality of expectations data. Related

[33] In this specification, a one percentage point increase in the probability of having a child in the next year increases the probability of working part-time in the current period by 0.07 percentage points. This coefficient is significant at a 10% level.

to work in Gong, Stinebrickner and Stinebrickner (2019), the first approach is built on a comparison of the unconditional probabilities of working at a particular future age computed using two different sets of survey questions. The second approach takes advantage of the fact that the BPS elicits both beliefs about own labor supply and beliefs about spousal labor supply.

Second, the paper suggests a novel use for expectations data. The most obvious use of expectations data is to capture beliefs that theory suggests are relevant for decision-making. The design of the BPS, with its first pilot in 1998, also recognized a second use of expectations data - allowing individuals to express uncertainty about outcomes that would occur in the future.[34] Further exploiting the unique design of the BPS, this paper examines a third use for expectations data - that the direct elicitation of beliefs can help address endogeneity concerns, when these concerns are present because variations in independent variables are caused by variations in beliefs. In an illustrative example, we utilize this approach to quantify the effect that children have on the labor supply of women. In addition to finding that children have a substantial effect on female labor supply, we also find that labor supply at a particular point in time is related to beliefs, at earlier stages, about future fertility.

[34]The BPS data of this type has been used in papers such as Stinebrickner and Stinebrickner (2014a) to study college major and Stinebrickner and Stinebrickner (2012, 2014b) to study dropout. For other early research recognizing this use see, e.g., Blass, Lach, and Manski (2010), van der Klaauw and Wolpin (2008), and van der Klaauw (2012). More recent work has recognized that this type of measurement, when used in an experimental setting, can allow one to examine how beliefs about outcomes change in response to changes in beliefs about factors that influence decisions (Zafar and Wiswall, 2014, Delavande and Zafar, forthcoming).

References

[1] Daniel Aaronson, Rajeev Dehejia, Andrew Jordan, Cristian Pop-Eleches, Cyrus Samii, and Karl Schulze. The effect of fertility on mothers' labor supply over the last two centuries. Technical report, National Bureau of Economic Research, 2017.

[2] Joseph G Altonji, Peter Arcidiacono, and Arnaud Maurel. The analysis of field choice in college and graduate school. *Handbook of the Economics of Education*, 5:305–396, 2016.

[3] John Ameriks, Gábor Kézdi, Minjoon Lee, and Matthew D Shapiro. Heterogeneity in expectations, risk tolerance, and household stock shares: The attenuation puzzle. *Journal of Business & Economic Statistics*, pages 1–27, 2019.

[4] Nikolay Angelov, Per Johansson, and Erica Lindahl. Parenthood and the gender gap in pay. *Journal of Labor Economics*, 34(3):545–579, 2016.

[5] Joshua D Angrist and Wlliam N Evans. Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3):450–477, 1998.

[6] Peter Arcidiacono. Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1):343–375, 2004.

[7] Peter Arcidiacono, V Joseph Hotz, and Songman Kang. Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16, 2012.

[8] Peter Arcidiacono, V Joseph Hotz, Arnaud Maurel, and Teresa Romano. Ex ante returns and occupational choice. *Working Paper*, 2019.

[9] Stefan Bauernschuster and Martin Schlotter. Public child care and mothers' labor supply—evidence from two quasi-experiments. *Journal of Public Economics*, 123:1–16, 2015.

[10] Asher A Blass, Saul Lach, and Charles F Manski. Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. *International Economic Review*, 51(2):421–440, 2010.

[11] Francine D Blau and Marianne A Ferber. Career plans and expectations of young women and men: The earnings gap and labor force participation. *Journal of Human Resources*, pages 581–607, 1991.

[12] Francine D Blau and Lawrence M Kahn. Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor Economics*, 25(3):393–438, 2007.

[13] Francine D Blau and Lawrence M Kahn. Female labor supply: Why is the united states falling behind? *American Economic Review*, 103(3):251–56, 2013.

[14] Francine D Blau and Anne E Winkler. *The economics of women, men, and work*. Oxford University Press, 2018.

[15] Richard Blundell and Thomas MaCurdy. Labor supply: A review of alternative approaches. In *Handbook of labor economics*, volume 3, pages 1559–1695. Elsevier, 1999.

[16] Martin Browning. Children and household economic behavior. *Journal of Economic Literature*, 30(3):1434–1475, 1992.

[17] Tanya S Byker. Paid parental leave laws in the united states: does short-duration leave affect women's labor-force attachment? *American Economic Review*, 106(5):242–46, 2016.

[18] Pierre-André Chiappori. Collective labor supply and welfare. *Journal of political Economy*, 100(3):437–467, 1992.

[19] Pierre-André Chiappori and Olivier Donni. Nonunitary models of household behavior: a survey of the literature. In *Household economic behaviors*, pages 1–40. Springer, 2011.

[20] Julian P Cristia. The effect of a first child on female labor supply evidence from women seeking fertility services. *Journal of Human Resources*, 43(3):487–510, 2008.

[21] Guillermo Cruces and Sebastian Galiani. Fertility and female labor supply in latin america: New causal evidence. *Labour Economics*, 14(3):565–573, 2007.

[22] Adeline Delavande, Xavier Giné, and David McKenzie. Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of development economics*, 94(2):151–163, 2011.

[23] Adeline Delavande and Basit Zafar. Gender discrimination and social identity: experimental evidence from urban pakistan. *Journal of Political Economy*, forthcoming.

[24] Xavier d'Haultfoeuille, Christophe Gaillac, and Arnaud Maurel. Rationalizing rational expectations? tests and deviations. *NBER Working Paper*, (w25274), 2018.

[25] Jeff Dominitz. Earnings expectations, revisions, and realizations. *Review of Economics and statistics*, 80(3):374–388, 1998.

[26] Jeff Dominitz and Charles F Manski. Eliciting student expectations of the returns to schooling. *Journal of Human Resources*, 31(1), 1996.

[27] Jeff Dominitz and Charles F Manski. Using expectations data to study subjective income expectations. *Journal of the American statistical Association*, 92(439):855–867, 1997.

[28] Tilman Drerup, Benjamin Enke, and Hans-Martin von Gaudecker. Measurement error in subjective expectations and the empirical content of economic models. 2014.

[29] Nada Eissa and Hilary Williamson Hoynes. Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of public Economics*, 88(9-10):1931–1958, 2004.

[30] Andrés Erosa, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson. Hours, occupations, and gender differences in labor market outcomes. *NBER Working Paper*, (w23636), 2017.

[31] Martin Gervais, Nir Jaimovich, Henry E Siu, and Yaniv Yedid-Levi. What should i be when i grow up? occupations and unemployment over the life cycle. *Journal of Monetary Economics*, 83:54–70, 2016.

[32] Pamela Giustinelli, Charles F Manski, and Francesca Molinari. Precise or imprecise probabilities? evidence from survey response on late-onset dementia. *NBER Working Paper*, (w23636), 2019.

[33] Pamela Giustinelli and Matthew D Shapiro. Seate: Subjective ex ante treatment effect of health on retirement. *NBER Working Paper*, (w26087), 2019.

[34] Yifan Gong, Todd Stinebrickner, and Ralph Stinebrickner. Uncertainty about future income: Initial beliefs and resolution during college. *Quantitative Economics*, 2019.

[35] Reuben Gronau. Sex-related wage differentials and women's interrupted labor careers-the chicken or the egg. *Journal of labor Economics*, 6(3):277–301, 1988.

[36] Siv Gustafsson and Frank Stafford. Child care subsidies and labor supply in sweden. *Journal of Human resources*, pages 204–230, 1992.

[37] Saul D Hoffman. The changing impact of marriage and children on women's labor force participation. *Monthly Lab. Rev.*, 132:3, 2009.

[38] Chinhui Juhn and Simon Potter. Changes in labor force participation in the united states. *Journal of Economic Perspectives*, 20(3):27–46, 2006.

[39] F Thomas Juster. Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association*, 61(315):658–696, 1966.

[40] Henrik Kleven, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, volume 109, pages 122–26, 2019a.

[41] Henrik Kleven, Camille Landais, and Jakob Egholt Søgaard. Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, 11(4):181–209, 2019b.

[42] Ilyana Kuziemko, Jessica Pan, Jenny Shen, and Ebonya Washington. The mommy effect: Do women anticipate the employment effects of motherhood? *NBER Working Paper*, (w24740), 2018.

[43] Lance Lochner. Individual perceptions of the criminal justice system. *NBER Working Paper*, (w9474), 2003.

[44] Shelly Lundberg. Labor supply of husbands and wives: A simultaneous equations approach. *The Review of Economics and Statistics*, pages 224–235, 1988.

[45] Shelly Lundberg and Elaina Rose. The effects of sons and daughters on men's labor supply and wages. *Review of Economics and Statistics*, 84(2):251–268, 2002.

[46] Petter Lundborg, Erik Plug, and Astrid Würtz Rasmussen. Can women have children and a career? iv evidence from ivf treatments. *American Economic Review*, 107(6):1611–37, 2017.

[47] Charles F Manski. Measuring expectations. *Econometrica*, 72(5):1329–1376, 2004.

[48] Charles F Manski and Francesca Molinari. Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics*, 28(2):219–231, 2010.

[49] Alice Nakamura and Masao Nakamura. The econometrics of female labor supply and children. *Econometric Reviews*, 11(1):1–71, 1992.

[50] Christopher J Ruhm. The economic consequences of parental leave mandates: Lessons from europe. *The quarterly journal of economics*, 113(1):285–317, 1998.

[51] Uta Schönberg and Johannes Ludsteck. Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. *Journal of Labor Economics*, 32(3):469–505, 2014.

[52] Lois B Shaw and David Shapiro. Women's work plans: Contrasting expectations and actual work experience. *Monthly Lab. Rev.*, 110:7, 1987.

[53] Mary A Silles et al. The impact of children on women's labour supply and earnings in the uk: evidence using twin births. *Oxford Economic Papers*, 68(1):197–216, 2016.

[54] Ralph Stinebrickner and Todd Stinebrickner. The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *The American economic review*, 98(5):2163–2184, 2008.

[55] Ralph Stinebrickner and Todd Stinebrickner. Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707–748, 2012.

[56] Ralph Stinebrickner and Todd Stinebrickner. A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, 81(1):426–472, 2014a.

[57] Ralph Stinebrickner and Todd Stinebrickner. Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, 32(3):601–644, 2014b.

[58] Ralph Stinebrickner, Todd Stinebrickner, and Paul Sullivan. Job tasks, time allocation, and wages. *Journal of Labor Economics*, 37(2):399–433, 2019.

[59] Ralph Stinebrickner and Todd R Stinebrickner. Understanding educational outcomes of students from low-income families evidence from a liberal arts college with a full tuition subsidy program. *Journal of Human Resources*, 38(3):591–617, 2003a.

[60] Ralph Stinebrickner and Todd R Stinebrickner. Working during school and academic performance. *Journal of labor Economics*, 21(2):473–491, 2003b.

[61] Robert H Topel and Michael P Ward. Job mobility and the careers of young men. *The Quarterly Journal of Economics*, 107(2):439–479, 1992.

[62] Wilbert Van der Klaauw. On the use of expectations data in estimating structural dynamic choice models. *Journal of Labor Economics*, 30(3):521–554, 2012.

[63] Wilbert Van der Klaauw and Kenneth I Wolpin. Social security and the retirement and savings behavior of low-income households. *Journal of Econometrics*, 145(1-2):21–42, 2008.

[64] Matthew Wiswall and Basit Zafar. Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824, 2014.

[65] Matthew Wiswall and Basit Zafar. Human capital investments and expectations about career and family. *NBER Working Paper*, (w22543), 2016.

[66] Basit Zafar. College major choice and the gender gap. *Journal of Human Resources*, 48(3):545–595, 2013.

Appendices

A Survey Questions

Question A.1. Your work status in the future may or may not depend on whether you are married and/or whether you have children. Taking into account the chances you might have children and the chances you might be married, what is the percent chance that you will be working full-time, part-time, or not working **at age 23** (or first year out of college if you will be older than 23 at graduation). Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance full-time at age 23 ____________
% Chance part-time at age 23 ____________
% Chance not working at age 23 ____________

Question A.2. Your work status in the future may or may not depend on whether you are married and/or whether you have children. Taking into account the chances you might have children and the chances you might be married, what is the percent chance that you will be working full-time, part-time, or not working **at age 28** (or five years out of college if you will be older than 23 at graduation). Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance full-time at age 28 ____________
% Chance part-time at age 28 ____________
% Chance not working at age 28 ____________

Question A.3. Your work status in the future may or may not depend on whether you are married and/or whether you have children. Taking into account the chances you might have children and the chances you might be married, what is the percent chance that you will be working full-time, part-time, or not working **at age 38** (or fifteen years out of college if you will be older than 23 at graduation). Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance full-time at age 38 ____________
% Chance part-time at age 38 ____________
% Chance not working at age 38 ____________

Question B. We are interested in whether you expect to have at least one child and when you expect to start having children. What is the percent chance that your first child will be born when you are each of the following ages? For example, the number on the first line is the percent chance that your first child will be born at or before age 23. On the last line enter the percent chance that you never have children. **Numbers should be between 0 and 100 and the numbers should sum to 100.**

Your Age	Percent Chance of first 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 C. 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 first 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 first 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 D.1. Assume at age 28 **you are not married**. What is the percent chance that **you** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance you work full-time	____________
% Chance you work part-time	____________
% Chance you are not working	____________

Question D.2. Assume at age 28 **you are married but have no children**. What is the percent chance that **you** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance you work full-time ____________
% Chance you work part-time ____________
% Chance you are not working ____________

Assume at age 28 **you are married but have no children**. What is the percent chance that **your spouse** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance spouse works full-time ____________
% Chance spouse works part-time ____________
% Chance spouse not working ____________

Question D.3. Assume at age 28 you are married and have a **child that is less than two years of age**.[35] What is the percent chance that **you** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance you work full-time ____________
% Chance you work part-time ____________
% Chance you are not working ____________

Assume at age 28 you are married and have a **child that is less than two years of age**. What is the percent chance that **your spouse** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance spouse works full-time ____________
% Chance spouse works part-time ____________
% Chance spouse not working ____________

[35]Given that the next portion of Question D asks the respondent to report her beliefs about future labor supply under the scenario where she has a youngest child between ages of 3 and 5, we interpret this portion of the Question as eliciting the respondent's beliefs about future labor supply under the scenario where she has a youngest child between the ages of 0 and 2.

Question D.4. Assume at age 28 you are married and have a **child that is between 3 and 5 years of age**. What is the percent chance that **you** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance you work full-time ____________
% Chance you work part-time ____________
% Chance you are not working ____________

Assume at age 28 you are married and have a **child that is between 3 and 5 years of age**. What is the percent chance that **your spouse** will be working full-time, part-time, or not working at all at age 28? Note: Each number should be between 0 and 100 and the three numbers should add up to 100.

% Chance spouse works full-time ____________
% Chance spouse works part-time ____________
% Chance spouse not working ____________

Question E. 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 add up to 100.**

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

Question F. What is your current AGE?______

Question G. Are you currently working in a job for pay? YES NO

Question H. How many jobs do you currently have?______
Note:If you have more than one job, please refer to the job in which you earn the most money per week as JOB1 and the job in which you earn the second most money per week as JOB2.
How many hours do you typically work **each week** in your job(s)?
Hours JOB1______ Hours JOB2______

Question I. Are you currently married? YES NO

Question J. 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 one child, when was **your youngest child** born?
Month______ Year______

Online Appendices

B Beliefs at the Beginning of Sophomore Year

Table B.1: Beliefs about Future Labor Supply, Cohort 2001, Year 2

Probability (%)	Age 23	Age 28	Age 38
Panel A: Full Sample, # of Obs. = 254			
Full-time	56.88	71.36	76.88
	(29.02)	(24.73)	(24.53)
Part-time	32.56	20.29	16.77
	(24.21)	(16.97)	(17.55)
Not Working	10.56	8.35	6.35
	(17.51)	(14.64)	(13.33)
Panel B: Male, # of Obs. = 103			
Full-time	57.72	78.45	84.39
	(30.91)	(21.39)	(19.61)
Part-time	31.75	15.68	12.06
	(25.86)	(14.86)	(15.63)
Not Working	10.53	5.87	3.55
	(18.46)	(14.03)	(10.92)
Panel C: Female, # of Obs. = 151			
Full-time	56.31	66.52	71.76
	(27.65)	(25.67)	(26.18)
Part-time	33.12	23.44	19.98
	(23.00)	(17.59)	(18.06)
Not Working	10.57	10.04	8.27
	(16.83)	(14.81)	(14.44)

Note: Standard deviations are in the parenthesis.

Table B.2: Beliefs about Future Labor Supply, Cohort 2001, Year 3

Probability (%)	Age 23	Age 28	Age 38
Panel A: Full Sample, # of Obs. = 218			
Full-time	61.17	71.92	77.40
	(24.73)	(17.94)	(18.05)
Part-time	30.48	19.62	15.29
	(24.21)	(16.97)	(17.55)
Not Working	8.36	8.46	7.30
	(15.88)	(16.82)	(15.17)
Panel B: Male, # of Obs. = 84			
Full-time	61.80	81.01	84.96
	(30.17)	(21.46)	(22.23)
Part-time	30.88	14.58	10.43
	(25.64)	(16.47)	(15.06)
Not Working	7.32	4.40	4.61
	(15.08)	(9.86)	(9.84)
Panel C: Female, # of Obs. = 134			
Full-time	60.77	66.22	72.66
	(30.10)	(27.64)	(28.51)
Part-time	30.22	22.78	18.34
	(24.14)	(18.10)	(19.07)
Not Working	9.01	11.00	8.99
	(16.33)	(19.56)	(17.51)

Note: Standard deviations are in the parenthesis.

Table B.3: Beliefs about the Timing of Family Outcomes, Cohort 2001, Year 2

Probability at Each Age (%)	Marriage			First Child		
	All	Male	Female	All	Male	Female
Before 23	17.07	14.12	19.08	7.24	5.47	8.44
	(24.41)	(20.47)	(26.58)	(14.95)	(10.39)	(17.29)
24 to 25	24.33	21.97	25.93	18.80	16.11	20.64
	(20.70)	(20.36)	(20.78)	(20.67)	(20.04)	(20.89)
26 to 27	22.42	21.78	22.86	25.81	22.72	27.91
	(16.88)	(17.16)	(16.67)	(19.88)	(18.30)	(20.62)
28 to 29	15.71	17.36	14.58	20.13	21.79	18.99
	(15.35)	(15.33)	(15.27)	(16.77)	(16.01)	(17.18)
After 30	11.20	15.20	8.48	16.88	22.85	12.81
	(15.28)	(19.35)	(10.90)	(21.08)	(24.90)	(16.84)
Never	9.28	9.57	9.08	11.14	11.06	11.20
	(19.34)	(19.18)	(19.45)	(22.06)	(21.04)	(22.73)
# of Obs.	254	103	151	254	103	151

Note: Standard deviations are in the parenthesis.

Table B.4: Beliefs about the Timing of Family Outcomes, Cohort 2001, Year 3

Probability at Each Age (%)	Marriage			First Child		
	All	Male	Female	All	Male	Female
Before 23	22.57	18.97	24.82	7.76	6.10	8.80
	(32.26)	(28.92)	(34.00)	(17.13)	(13.48)	(18.99)
24 to 25	21.71	19.90	22.84	15.53	13.88	16.57
	(21.42)	(21.90)	(21.04)	(18.10)	(17.98)	(18.09)
26 to 27	19.30	19.79	18.99	26.34	21.92	29.11
	(18.02)	(17.43)	(18.37)	(22.97)	(21.11)	(23.65)
28 to 29	16.12	18.31	14.74	22.00	23.79	20.87
	(17.81)	(19.15)	(16.78)	(19.67)	(20.31)	(19.16)
After 30	10.99	13.51	9.41	15.55	21.64	11.73
	(17.30)	(21.37)	(13.93)	(21.22)	(26.40)	(16.05)
Never	9.31	9.51	9.19	12.82	12.67	12.92
	(20.71)	(23.13)	(19.03)	(25.18)	(24.34)	(25.70)
# of Obs.	218	84	134	218	84	134

Note: Standard deviations are in the parenthesis.

Table B.5: Beliefs about Conditional Labor Supply at Age 28, Cohort 2001, Year 2

Probability (%)	Unmarried	Own - Married			Spousal - Married		
		No Kids	Age 0-2	Age 3-5	No Kids	Age 0-2	Age 3-5
Full Sample, # of Obs. = 254							
Full-time	84.87	82.48	59.64	66.88	77.36	64.83	69.54
	(17.64)	(18.59)	(31.33)	(29.75)	(22.56)	(30.27)	(27.73)
Part-time	11.96	14.17	25.54	22.52	16.72	22.26	20.29
	(14.12)	(15.00)	(20.51)	(20.16)	(16.06)	(19.32)	(18.38)
Not Working	3.17	3.35	14.81	10.60	5.91	12.91	10.17
	(8.54)	(8.49)	(23.51)	(19.61)	(12.35)	(22.36)	(18.33)
Male, # of Obs. = 103							
Full-time	81.28	82.19	79.27	81.47	65.33	46.90	53.83
	(20.03)	(18.83)	(22.48)	(21.17)	(25.76)	(32.32)	(29.20)
Part-time	13.78	14.13	15.68	14.08	23.03	27.98	26.17
	(15.14)	(15.22)	(17.72)	(16.21)	(17.36)	(20.98)	(18.21)
Not Working	4.94	3.68	5.05	4.45	11.63	25.12	20.00
	(12.24)	(10.86)	(12.36)	(11.75)	(17.08)	(29.20)	(24.53)
Female, # of Obs. = 151							
Full-time	87.32	82.68	46.25	56.92	85.57	77.07	80.25
	(15.32)	(18.42)	(29.41)	(30.64)	(15.41)	(21.44)	(20.70)
Part-time	10.72	14.19	32.27	28.28	12.42	18.36	16.27
	(13.25)	(14.85)	(19.54)	(20.56)	(13.52)	(17.03)	(17.39)
Not Working	1.96	3.12	21.48	14.80	2.01	4.58	3.47
	(4.11)	(6.38)	(26.75)	(22.57)	(4.47)	(9.38)	(6.64)

Note: Standard deviations are in the parenthesis.

Table B.6: Beliefs about Conditional Labor Supply at Age 28, Cohort 2001, Year 3

Probability (%)	Unmarried	Own - Married			Spousal - Married		
		No Kids	Age 0-2	Age 3-5	No Kids	Age 0-2	Age 3-5
Full Sample, # of Obs. = 218							
Full-time	84.49	82.56	61.19	68.27	79.56	68.67	71.47
	(19.53)	(19.74)	(32.82)	(30.91)	(22.13)	(29.82)	(28.88)
Part-time	12.68	14.47	23.87	19.12	15.87	20.52	18.78
	(16.46)	(17.04)	(20.90)	(17.36)	(17.76)	(20.34)	(18.35)
Not Working	2.83	2.97	14.94	12.62	4.57	10.81	9.75
	(6.08)	(5.81)	(24.46)	(22.13)	(9.13)	(20.40)	(19.07)
Male, # of Obs. = 84							
Full-time	82.87	83.41	80.35	84.69	70.33	49.77	52.88
	(20.54)	(19.07)	(22.54)	(17.30)	(24.34)	(32.18)	(31.21)
Part-time	13.79	13.36	15.79	11.87	22.17	28.32	26.94
	(16.63)	(16.15)	(18.65)	(13.78)	(19.09)	(22.11)	(18.42)
Not Working	3.35	3.23	3.85	3.45	7.50	21.92	20.18
	(6.90)	(5.74)	(7.93)	(6.26)	(12.46)	(27.96)	(26.45)
Female, # of Obs. = 134							
Full-time	85.51	82.03	49.17	57.97	85.34	80.52	83.13
	(18.80)	(20.12)	(32.54)	(33.04)	(18.40)	(20.81)	(19.84)
Part-time	11.99	15.16	28.94	23.66	11.92	15.64	13.66
	(16.32)	(17.54)	(20.65)	(17.82)	(15.63)	(17.46)	(16.34)
Not Working	2.51	2.81	21.89	18.36	2.74	3.84	3.22
	(5.47)	(5.85)	(28.44)	(26.19)	(5.43)	(7.81)	(6.49)

Note: Standard deviations are in the parenthesis.

Table B.7: Beliefs about the Number of Children, Cohort 2001, Year 2

Probability (%)	All	Male	Female
0 Child	12.95	13.73	12.41
	(23.14)	(22.54)	(23.53)
1 Child	21.23	23.62	19.60
	(17.43)	(17.36)	(17.29)
2 Children	33.52	33.18	33.75
	(20.85)	(20.49)	(21.10)
3 Children	22.37	20.64	23.56
	(18.33)	(16.96)	(19.11)
$\geq$ 4 Children	9.93	8.83	10.68
	(17.08)	(16.26)	(17.57)
# of Obs.	254	103	151

Note: Standard deviations are in the parenthesis.

Table B.8: Beliefs about the Number of Children, Cohort 2001, Year 3

Probability (%)	All	Male	Female
0 Child	12.35	12.74	12.11
	(23.63)	(22.13)	(24.51)
1 Child	21.17	22.98	20.04
	(18.60)	(17.76)	(19.01)
2 Children	33.66	33.86	33.54
	(22.53)	(21.40)	(23.21)
3 Children	22.29	21.80	22.60
	(18.63)	(19.38)	(18.14)
$\geq$ 4 Children	10.53	8.63	11.71
	(16.44)	(13.00)	(18.16)
# of Obs.	218	84	134

Note: Standard deviations are in the parenthesis.

C Construction of $F_{A_i}(a_i)$

In this section we discuss how we construct the distribution of A_i which describes student i's beliefs about the age of the youngest child at age 28, a_i. We first note that this distribution can be obtained through a simulation-based approach if information about student i's beliefs about the evolution of her future fertility situation is available. We need to simulate student i's entire fertility history for a large number of times and record the age of the youngest child at age 28 associated with each simulation. The distribution of these recorded ages converges in distribution to the distribution of A_i as the number of simulation increases.

We model student i's beliefs about future fertility outcomes as follows. Let $g_{i,q}$ denote student i's age of having the qth child, and $G_{i,q}$ denote the random variable describing student i's beliefs about $g_{i,q}$. We assume that students believe they will have no more than four children, and that children will not be born after age 40. For ease of notation, we let $G_{i,1} = G_{i,2} = G_{i,3} = G_{i,4} = 40$ if student i has no children in her lifetime.

As discussed earlier, Question D in Appendix A provides direct information on $G_{i,1}$. Specifically, the distribution of $G_{i,1}$ can be exactly determined from student i's responses to Question D under the assumption that the density function of $G_{i,1}$ is 1) flat between age 22 and 23, between 24 and 25, between 26 and 27, and between 28 and 29 and 2) decreases linearly to zero between age 30 and 39. To take advantage of Question E to estimate $G_{i,q}$, for $q \geq 2$, we begin by assuming that student i believes that, net of the 10 months ($\frac{5}{6}$ year) necessary for pregnancy, the age gap between having two consecutive children follows an exponential distribution with mean $\mu_{i,q}$. Formally, we have:

$$G_{i,q+1} - G_{i,q} - \frac{5}{6} \sim \text{Exp}(\mu_{i,q+1}). \qquad \text{(3 revisited)}$$

The value of $\mu_{i,q+1}$ can be computed from student i's beliefs about the number of children he/she will have and information on $G_{i,1}$. Note that, if $\mu_{i,q+1}$, $q = 1, 2, 3$ (and the distribution of $G_{i,1}$) are known, we can compute the probability that student i has Q children given that $Q \geq 1$ in her lifetime using a simulation-based approach similar to the one described above. We denote this model-implied probability $\tilde{P}^K_{i,Q}(\mu_{i,2}, \mu_{i,3}, \mu_{i,4})$, $Q = 1, 2, 3, 4$ and denote the directly elicited probability of having Q children given that $Q \geq 1$ in her lifetime $\hat{P}^K_{i,Q}$. For each student, we numerically search for the set of parameters $\{\mu_{i,2}, \mu_{i,3}, \mu_{i,4}\}$ that minimizes a weighted sum of the discrepancies between observed and model implied probabilities. We weight each category by its associating probability. Formally, we have:

$$\{\widehat{\mu_{i,2}}, \widehat{\mu_{i,3}}, \widehat{\mu_{i,4}}\} = argmin \sum_{Q \in \{1,2,3,4\}} \hat{P}^K_{i,Q}(\tilde{P}^K_{i,Q}(\mu_{i,2}, \mu_{i,3}, \mu_{i,4}) - \hat{P}^K_{i,Q})^2. \tag{C.1}$$

Once parameters $\{\mu_{i,2}, \mu_{i,3}, \mu_{i,4}\}$ are estimated, we can approximate the distribution of A_i by simulation using the method described in the first paragraph of this appendix.

D Classical Measurement Error

In this appendix, we quantify the magnitude of measurement error contained in responses to survey questions under the assumption that it is classical using a method similar to that developed in Gong, Stinebrickner, and Stinebrickner (2019). Table 1 showed that there is substantial cross-sectional variation (measured by the standard deviations in the parenthesis) in the reported probabilities of working at age 28. To provide some quantitative evidence about the contribution of measurement error and true heterogeneity to this variation, we take advantage of the fact that the BPS makes it possible to obtain student i's perceived probability of having work status j at age 28, P^j_i, in two distinct ways. We refer to the perceived probability elicited directly using survey Question A.2 in Appendix A as $\tilde{P}^j_i$. We refer to the perceived probability computed using the alternative method detailed in Section 3.2.1 as $\hat{P}^j_i$. The intuition underlying our method for estimating the magnitude of measurement error is that the two probabilities will be identical if the responses to the survey questions used to compute these values are not affected by measurement error. However, when the two values are different, measurement error is present and its importance can be quantified if one specifies the manner in which measurement error influences responses to the survey questions.

Formally, we write the directly elicited probability as:

$$\tilde{P}^j_i = P^j_i + \varsigma^j_i, j \in \{F, P, N\}, \tag{D.1}$$

where ς^j_i is the classical measurement error attached to the true value P^j_i. We allow ς^j_i to be correlated across j. Since the sum of the probabilities $\tilde{P}^j_i$ over j and the sum of P^j_i over j are each equal to one, the sum of ς^j_i over j is equal to zero.

D.1 Magnitude of the Measurement Error

We are interested in characterizing the variance in the true value, P_i^j, across students because this variance represents a measure of how much heterogeneity exists in actual beliefs. Taking the variance of both sides of Equation (D.1) we see that dispersion in the reported value, $\tilde{P}_i^j$, across students originates from both variation in the true value across students and randomness caused by measurement error, ς_i^j:

$$var(\tilde{P}_i^j) = var(P_i^j) + var(\varsigma_i^j), j \in \{F, P, N\}. \tag{D.2}$$

A simple rearrangement of Equation (D.2) reveals that the object of interest, $var(P_i^j)$, can be obtained by subtracting the variance of the measurement error term, $var(\varsigma_i^j)$, from the directly-observable variance of the reported probabilities, $var(\tilde{P}_i^j)$. Thus, the remainder of this section focuses on estimating the variance of ς_i^j.

Section 3.2.1 discusses how the computed probability $\hat{P}_i^j$ can be obtained from responses to questions eliciting beliefs about labor supply at age 28 conditional on having various family outcomes (Question D), as well as questions eliciting beliefs about family outcomes (Questions B, C and E). Similar to the assumption made in Equation (D.1), we assume that measurement error influences the responses to the former types of questions in a classical manner, that is,

$$\tilde{P}_i^{j,k} = P_i^{j,k} + \varsigma_i^{j,k}, \qquad k \in \{N, 02, 35\}, \tag{D.3}$$

where $\tilde{P}_i^{j,k}$ is the reported value of $P_i^{j,k}$, the actual perceived conditional probability of working given family outcome $k \in \{N, 02, 35\}$, and $\varsigma_i^{j,k}$, $k \in \{N, 02, 35\}$, are the corresponding classical measurement errors.

Taking into account that the reports of the actual perceived conditional probabilities $P_i^{j,k}$ may be noisy and are given by $\tilde{P}_i^{j,k}$, the probability $\hat{P}_i^j$ can be computed using Equation (2):

$$\begin{aligned}
\hat{P}_i^j &= \sum_{k \in \{N,02,35\}} \pi_i^{A,k} \tilde{P}_i^{j,k} \\
&= \sum_{k \in \{N,02,35\}} \pi_i^{A,k} P_i^{j,k} + \sum_{k \in \{N,02,35\}} \pi_i^{A,k} \varsigma_i^{j,k} \\
&= P_i^j + \sum_{k \in \{N,02,35\}} \pi_i^{A,k} \varsigma_i^{j,k}, j \in \{F, P, N\}.
\end{aligned} \tag{D.4}$$

Here, we assume that no error is introduced during the computation of $\pi_i^{A,k}$. When using a similar approach, Gong, Stinebrickner and Stinebrickner (2019) show that relaxing this assumption and specific other assumptions that were needed in Section 3.2.1 to arrive at Equation (2) will lead to a smaller estimate for the magnitude of measurement error.

The intuition underlying identification is that the difference between $\tilde{P}_i^j$ and $\hat{P}_i^j$ is informative about the amount of measurement error. Taking this difference,

$$\tilde{P}_i^j - \hat{P}_i^j = \varsigma_i^j - \sum_{k\in\{N,02,35\}} \pi_i^{A,k}\varsigma_i^{j,k}, j \in \{F,P,N\}. \tag{D.5}$$

Using equation (D.5) to estimate $var(\varsigma_i^j)$ requires assumptions about the joint distribution of ς_i^j, $\varsigma_i^{j,N}$, $\varsigma_i^{j,02}$ and $\varsigma_i^{j,35}$. The prior assumption that ς_i^j and $\varsigma_i^{j,k}$, $k \in \{N, 02, 35\}$, represent classical measurement error implies that they have mean zero and are independent of other factors. In addition, we assume that the four measurement error terms are independent and identically distributed.

Under these assumptions, taking the variance of both sides of Equation (D.5), we have:

$$\begin{aligned}
var(\tilde{P}_i^j - \hat{P}_i^j) &= var(\varsigma_i^j - \sum_{k\in\{N,02,35\}} \pi_i^{A,k}\varsigma_i^{j,k}) \\
&= var(\varsigma_i^j) + \sum_{k\in\{N,02,35\}} var(\pi_i^{A,k}\varsigma_i^{j,k}) && \text{(independence of MEs)} \\
&= var(\varsigma_i^j) + \sum_{k\in\{N,02,35\}} E((\pi_i^{A,k})^2)E((\varsigma_i^{j,k})^2) - (E(\pi_i^{A,k})E(\varsigma_i^{j,k}))^2 \\
& && (\pi_i^{A,k} \perp\!\!\!\perp \varsigma_i^{j,k}) \\
&= var(\varsigma_i^j) + \sum_{k\in\{N,02,35\}} E((\pi_i^{A,k})^2) var(\varsigma_i^{j,k}) && (E(\varsigma_i^j) = 0 \text{ and } E(\varsigma_i^{j,k}) = 0) \\
&= var(\varsigma_i^j)[1 + \sum_{k\in\{N,02,35\}} E((\pi_i^{A,k})^2)]. && (var(\varsigma_i^j) = var(\varsigma_i^{j,k}))
\end{aligned}$$

Therefore,

$$var(\varsigma_i^j) = \frac{var(\tilde{P}_i^j - \hat{P}_i^j)}{1 + \sum_k E((\pi_i^{A,k})^2)}. \tag{D.6}$$

Note that the sample analogs of $var(\tilde{P}_i^j - \hat{P}_i^j)$ and $E((\pi_i^{A,k})^2)$ can be computed in a straightforward manner from the data.[36] Hence, $var(\varsigma_i^j)$ (std(ς_i^j)), and, therefore, $var(P_i^j)$, can be estimated.

Table D.1 reports estimates for the standard deviation of the reported probability, std$(\tilde{P}_i^j)$, the standard deviation of measurement error, std(ς_i^j), and the standard deviation of the actual perceived probability, std(P_i^j). We allow the distribution of ς_i^j to vary by gender. Comparing the second column to the third column reveals that responses to survey questions indeed contain a non-negligible amount of measurement error; the magnitude of measurement error is comparable to the magnitude of heterogeneity for

[36]Computation of the sample analog of $var(\tilde{P}_i^j - \hat{P}_i^j)$ involves finding the difference between $\tilde{P}_i^j$ and $\hat{P}_i^j$ for each individual and then computing the variance of this difference across all individuals in the sample. Computation of the sample analog of $E((\pi_i^{A,k})^2)$ involves computing $\pi_i^{A,k}$ for each individual and then taking the sample average of $(\pi_i^{A,k})^2$.

females and is roughly 50% of the magnitude of heterogeneity for males. Comparing Panel A to Panel B, we find that, while the magnitude of heterogeneity in reported beliefs about labor supply for females is substantially larger than that for males, female students' responses to survey questions contain more measurement error as well.[37] As a result, the third column shows that the magnitude of heterogeneity in actual beliefs about labor supply is somewhat similar between males and females.

Table D.1: Heterogeneity and Measurement Error in Beliefs

Unit: %	std($\tilde{P}_i^j$)	std(ς_i^j)	std(P_i^j)
Panel A: Male, # of Obs. = 153			
Full-time	20.85	9.18	18.72
	(1.62)	(1.09)	(1.80)
Part-time	16.24	7.88	14.20
	(1.12)	(1.02)	(1.26)
Not Working	8.89	4.72	7.54
	(1.71)	(1.24)	(1.80)
Panel B: Female, # of Obs. = 265			
Full-time	25.65	17.23	19.00
	(1.00)	(1.03)	(1.36)
Part-time	18.74	14.66	11.67
	(0.86)	(0.91)	(1.17)
Not Working	16.53	12.60	10.70
	(1.92)	(1.04)	(2.75)

Note: Bootstrapped standard errors are in parentheses.

D.2 Correcting the Attenuation Bias

The presence of measurement error in elicited beliefs leads to the standard attenuation bias when beliefs are used as independent variables in a regression framework. In this appendix, we describe who to correct for this bias given that the variance of the measurement error is known.

Let vector $\boldsymbol{z}_i$ denote the independent variables that are accurately measured and x_i denote the independent variable that is measured with classical measurement error η_i. We allow the variance of η_i to depend on observable $\boldsymbol{g}_i$ and denote this variance $\sigma_{ME}^2(\boldsymbol{g}_i)$. Let $\tilde{x}_i = x_i + \eta_i$ denote the measured value of x_i. Then, the dependent variable y_i is given by:

[37] Conceptually, when reporting the probability of having work status j, students need to take their beliefs about all factors that influence labor supply into consideration. Misperceptions about labor supply arise because of misperceptions about these factors. Then, the observed gender difference in measurement error would be consistent with a traditional view that women have more factors that influence whether they work, while men tend to think they will most likely work "no matter what."

$$\begin{aligned} y_i &= \boldsymbol{z}_i' \boldsymbol{\alpha} + \beta x_i + \epsilon \\ &= \boldsymbol{z}_i' \boldsymbol{\alpha} + \beta \tilde{x}_i + (\epsilon - \beta \eta_i). \end{aligned} \tag{D.7}$$

By construction, $\tilde{x}$ and $\epsilon - \beta\eta_i$ are correlated. Hence, the OLS estimator is biased. To correct for this bias, we notice that:

$$E\left[(y_i - (\boldsymbol{z}_i'\boldsymbol{\alpha} + \beta\tilde{x}_i)) \begin{pmatrix} \boldsymbol{z}_i \\ \tilde{x}_i \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \beta\sigma^2_{ME}(\boldsymbol{g}_i) \end{pmatrix} \right] = E\left[(\epsilon - \beta\eta_i) \begin{pmatrix} \boldsymbol{z}_i \\ \tilde{x}_i \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \beta\sigma^2_{ME}(\boldsymbol{g}_i) \end{pmatrix} \right] = \mathbf{0}. \tag{D.8}$$

Equation system (D.8) has the same number of equations and parameters which are equal to the number of observables. Hence, it can be estimated using the Method of Moments, i.e., the estimator of $\begin{pmatrix} \boldsymbol{\alpha} \\ \beta \end{pmatrix}$ is the solution to the sample analog of the moment conditions defined by Equation D.8. It is easy to show that this estimator has an easy-to-implement matrix-form expression. Letting θ denote $\begin{pmatrix} \boldsymbol{\alpha} \\ \beta \end{pmatrix}$ and $\boldsymbol{q}_i$ denote $\begin{pmatrix} \boldsymbol{z}_i \\ \tilde{x}_i \end{pmatrix}$, we have:

$$\hat{\theta} = \left[Q'Q - \begin{pmatrix} \mathbf{0} & 0 \\ 0 & \sum_i \sigma^2_{ME}(\boldsymbol{g}_i) \end{pmatrix} \right]^{-1} Q'Y, \tag{D.9}$$

where and Y and Q are the matrices of y_i and $\boldsymbol{q}_i$, respectively.

In the context of this paper, $\tilde{x}_i$ is the reported perceived probability of having certain outcome (e.g., being married, having a child, working at all and working full-time at age 28), and $\boldsymbol{g}_i$ is the gender of the student. $\sigma^2_{ME}(\boldsymbol{g}_i)$ can be estimated using the method detailed in online Appendix D.1. This information allows us to compute $\hat{\theta}$ using Equation (D.9).

Table D.2: Regression of Actual Outcomes at Age 28 on Perceived Probabilities of Outcomes

	Married	Young Child	Working		Full-time	
Constant	0.035	-0.018	0.634	0.462	0.631	0.580
	(0.075)	(0.049)	(0.112)	(0.197)	(0.070)	(0.099)
Probability	0.66	0.559	0.242	0.428	0.156	0.227
	(0.097)	(0.077)	(0.120)	(0.213)	(0.093)	(0.135)
Correlation	0.3554	0.3784	0.1117	0.1485	0.0932	0.1125
ME Correction	✗	✗	✗	✓	✗	✓
# of Obs	328	328	327	327	325	325

Note: Standard errors are in parentheses. Probability is the perceived (during school) percent chance (0 to 100) of the outcome in a particular column occurring at age 28.

We now correct for the attenuation bias present in the regressions performed in Section 4.2 and revisit the results. The second columns of the last two panels of Table D.2 indicate that, after correcting for measurement error, a one percentage point increase in the perceived probability of working at all and working full-time, respectively, is associated with a 0.43 and 0.23 percentage point increase in the actual probability of working at all and working full-time at age 28, respectively. The correlations increase to .15 and .11, respectively. Thus, our results suggest that, while, on average, perceptions about labor supply are reasonably accurate (shown in Section 4.1), the relationship between individual perceptions and actual outcomes is weaker than what is seen for the family outcomes.

E Table 6 for Respondents Who Appeared in Both In-school and Post-college Samples

Table E.1: Average Perceived and Actual Labor Supply Probabilities

Probability (%)	# of Obs.	Outcomes		Beliefs	
		Working	Full-time	Working	Full-time
Panel A: Full Sample					
Unmarried	154	87.66	75.89	97.01	83.65
		(2.65)	(3.27)	(0.46)	(1.50)
Married without Children	85	92.94	81.18	98.18	86.48
		(2.78)	(4.15)	(0.61)	(1.97)
Married with Young Children	78	74.36	62.62	80.85	63.31
		(4.94)	(4.75)	(2.66)	(3.46)
Panel B: Male					
Unmarried	55	87.27	74.55	95.47	78.42
		(4.49)	(5.57)	(1.09)	(2.99)
Married without Children	21	95.24	95.24	98.12	88.80
		(5.48)	(7.46)	(1.32)	(3.79)
Married with Young Children	33	93.94	90.91	98.12	88.80
		(4.65)	(4.43)	(1.74)	(4.47)
Panel C: Female					
Unmarried	99	87.88	76.64	97.87	86.57
		(3.28)	(4.04)	(0.38)	(1.63)
Married without Children	55	94.55	83.64	98.27	87.02
		(3.06)	(4.91)	(0.62)	(2.25)
Married with Young Children	57	66.67	50.45	74.48	49.54
		(6.24)	(5.68)	(3.54)	(3.94)

Note: The second and the third columns report the fraction of students who were working at all and working full-time at age 28 given each of the family outcomes. Similarly, the fourth and fifth columns report the average perceived conditional probability of working at all and working full-time at age 28 given a certain family outcome for students who actually had that family outcome at age 28. The sample sizes are reported in the first column. Standard errors are in the parenthesis.