Using Intentions to Predict Fertility

Johannes Norling*

July 21, 2020

Abstract

On average, childless women observed by the Panel Study of Income Dynamics report that they intend to have more children than they actual have. A collection of intentions that record only whether respondents intend to have another child can more accurately predict the number of children they have. Errors in the formation of intentions are not required to explain this finding. Rather, if intentions record a survey respondent's most likely predicted number of children, then the average of these intentions does not necessarily equal average actual fertility, even if intentions are formed using rational expectations.

JEL codes: C83; D84; J13

Keywords: Intentions; Fertility; Rational expectations; Survey methods

* Contact information: Department of Economics, Mount Holyoke College, 50 College Street, South Hadley, MA 01075; jnorling@mtholyoke.edu.

1 Introduction

Childbearing involves uncertainty. Imperfect control over conception, marital formation and dissolution, changes in financial resources, and other mutable factors influence whether and when people have children (Heckman and Willis 1976). Because these factors can prevent people from perfectly predicting their future fertility, the number of children people say they intend to have frequently differs from the number of children they end up having (Bachrach and Morgan 2013). In 1985, the Panel Study of Income Dynamics (PSID) asked several thousand people in the United States to state whether they intend to have a child in the future, and, if so, how many children they intend to have (Panel Study of Income Dynamics 2017). The PSID then tracked childbearing over the subsequent decades. Fifteen percent of women had a child after stating that they did not intend to do so, or did not have a child after stating that they intended to do so. Twenty-five percent had fewer or more children than they had stated was their intention. Many studies similarly document differences between stated intentions and actual fertility using longitudinal surveys administered in the United States and other countries.[1]

[1] Stated fertility intentions and actual fertility have been compared using longitudinal data from several surveys, including the British Household Panel Survey (Berrington 2004), the Malaysian Family Life Survey (DaVanzo et al. 2003), the Malaysian Population and Family Survey (Tan and Tey 1994), the Survey of Fertility Intentions in France (Toulemon and Testa 2005), and the following surveys in the United States: the Detroit Area Study (Freedman et al. 1980, Freedman and Thornton 1982), the National Fertility Study (Westoff and Ryder 1977), the National Longitudinal Survey of Labor Market Experiences of Young Women (Stolzenberg and Waite

In this paper, I compare the accuracy of two ways of using fertility intentions to predict the number of children that people will have. I use a sample of women aged 18–44 who had never given birth before 1985 and who subsequently had 0.99 children on average. In 1985, these women reported that they intend to have 1.44 children on average in the future, nearly half a child higher on average than their actual future fertility. Alternatively, for the same group of childless women in 1985, I calculate the share who reported that they intend to have at least one child in the future. Then, among women with one child in 1985, I calculate the share who reported that they intend to have at least one more child. I continue with these calculations at higher parities, which together yield a predicted distribution over total future number of children. The average of this distribution, 1.00, almost exactly matches the actual average of 0.99. Responses to the question "do you intend to have a baby at some time?" collected from a broad group of women better predict fertility for childless women than do these women's own responses to the question "how many do you intend to have?" If the goal is to predict average fertility, then the best way to ask people how many children they intend to have may not be to ask them "how many children do you intend to have?"

I consider the role of rational expectations in explaining this surprising finding. Again, because of uncertain future factors, people may not be able to perfectly predict fertility. A person with rational expectations knows the expected distribution of these factors, even though

---

1977), the 1979 National Longitudinal Survey of Youth (Quesnel-Vallée and Morgan 2003, Hayford 2009, Miller et al. 2010, Morgan and Rackin 2010), the National Survey of Family Growth (Williams et al. 1999), the National Survey of Families and Households (Schoen et al. 1999), and the Princeton Fertility Survey (Westoff et al. 1963).

she doesn't know what her draw will be. Manski (1990) provides a framework for evaluating whether intentions are consistent with an assumption that they are formed using rational expectations. I show that intention to have a child is generally consistent with rational expectations, but intended number of children differs from subsequent fertility to a degree that suggests errors in the formation of intentions. However, I also show that the surprising pattern observed among PSID respondents does not require errors in the formation of intentions. Rather, the structure of intentions themselves can provide a particularly inaccurate prediction. If respondents' state their most likely number of children as their intended number of children, then actual fertility can differ from intended fertility in aggregate, even without any errors in the formation of intentions.

2 Panel Study of Income Dynamics

The PSID is a longitudinal survey of several thousand U.S. families, conducted annually between 1968 and 1996 and biannually starting in 1997. In 1985, the PSID recorded fertility intentions of women who were age 44 or younger and of men whose wives were age 44 or younger. The PSID asked two questions. First, "Looking to the future, do you intend to have a baby at some time?" The PSID permitted three responses: "Yes," "No," and "Maybe; Don't know." Among respondents who stated that they intend to have a child in the future, the PSID

further asked "How many do you intend to have?" The PSID then recorded actual childbearing over the following decades.[2]

I use a sample 3,251women aged 18–44 in 1985 who were present in the household and responded for themselves; who reported their current number of children; who were not pregnant; who reported the year of birth of every child they have had; who never adopted a child; who never had a twin or other multiple birth; and who answered the initial intentions question, "do you intend to have a baby at some time?" Column 1 of Table 1 provides basic demographic characteristics of this sample. These women were 32 years old on average when observed in 1985. Eighty-six percent were white, 66 percent were married, and 85 percent had completed high school (these calculations, and all others in the paper, are made using sampling weights). Eight percent of these women responded with uncertainty when asked whether they intend to

[2] In 1972, the PSID asked household heads (who were generally men) to report the total number of children they expected to have. This question was repeated to their wives in 1976. Forty-nine percent of men reported a greater number of expected children in 1972 than their number of intended children in 1985. The value remained constant for 20 percent of men, and rose for 32 percent of men. Thirty-four percent of women reported a greater number of expected children in 1976 than their number of intended children in 1985, 36 reported the same value over time, and 30 percent reported an increase. Although the 1972 and 1976 questions recorded expectations while the 1985 question recorded intentions, preventing an exact comparison over time, this evidence confirms the findings of Lee (1980), Hayford (2009), and others studies that document average decreases in expected or intended family size as people age.

have a child in the future. I exclude these women from all estimates that require a response of yes or no to this intentions question.

I focus much of the analysis in this paper on a subsample of 394 women who had zero children by 1985; whose birth history continued to be recorded until at least age 40;[3] and who either responded "No" when asked whether they intend to have a child in the future, or who responded "Yes" and then also provided an intended number of children. Columns 2 and 3 of Table 1 compare these women to the 202 women with zero children in 1985 who were not followed until age 40. The PSID continued to follow women who were, on average, older, more likely to be white, less likely to be married, and more likely to have completed high school. Except for the share that was white, these differences are all statistically significant at the five-percent level.

The 394 women who were aged 18–44 and childless in 1985 (and who continued to be observed by the PSID until at least age 40) later had 0.99 children on average. Because completed fertility was not recorded for the 202 childless women who were not followed until age 40, I cannot directly determine whether the sample restriction biases completed fertility higher or lower. However, I estimate the direction of this bias by first regressing completed number of children on the demographic characteristics listed in Table 1, using the 394 women who were followed. Estimated coefficients from this regression, presented in Table 2, indicate

[3] Age 40 is often used to identify women who are generally finished having children (Schoen *et al*. 1999; Hayford 2009; Modrek and Ghobadi 2011; Beaujouan and Solaz 2013; Cornolli and Bernardi 2015). The PSID collects birth histories from few women after age 44. Among women last observed at age 44, just two percent gave birth after age 39.

that completed fertility is higher for women who were younger, white, married, or had graduated high school. Using the relationship between completed fertility and each demographic characteristic suggested by these estimated coefficients, I simulate fertility for the 202 women who were not followed. The average value of this simulation, 1.38 children, substantially exceeds the 0.99 children on average that were born to the women who were followed. This comparison suggests that attrition from the survey leads to an underestimate of future fertility.

3 Two ways of predicting fertility

I begin with the sample of 394 women who were aged 18–44 and childless in 1985, and who continued to be observed until at least age 40. These women went on to have 0.99 children on average: 47 percent remained childless, 17 percent had one child, 28 percent had two children, eight percent had three children, and one percent had four children. Column 1 of Table 3 provides these shares, with standard errors in parentheses. In the rest of this section, I compare this actual fertility to two ways of predicting fertility for these same women using intentions reported in 1985.

3.1 Categorical intentions

The PSID asked these 394 women "do you intend to have a baby?" Thirty-six percent reported that they did not, a smaller share than the 47 percent who actually remained childless. The remaining 64 percent were then asked "how many [children] do you intend to have?" Column 2 of Table 3 provides the distribution of responses to this question. I label these

intentions "categorical" because responses are integers between one and five children. Eight percent intended to have a single child, again a smaller share than actually achieved this family size. Conversely, 36 percent intended to have two children and 15 percent intended to have three children, larger shares than actually achieved these family sizes. The overall average number of intended children, 1.44, therefore substantially exceeds the average number of children actually born to these women. This finding – that actual fertility falls short of intended fertility – is consistent with the evidence in section 2 of an average decline in fertility intentions over time. As women age, they are more likely to revise their intended number of children downward than upward. It therefore follows that average intended number of children exceeds actual fertility.

3.2 Binary intentions

I construct an alternative predicted distribution for these same childless women using only responses to the intentions question "do you intend to have a baby?" Because I use responses of yes and no, I refer to this distribution as "binary." (Ajzen and Fishbein [1980] refer to this intention as a "choice" intention.) Again, 36 percent of women report that they do not intend to have a child, so binary intentions predict that 36 percent of women will remain childless, just as categorical intentions do. However, at higher parities the distributions differ. To explain how I construct the predicted distribution using binary intentions and why it differs from the prediction using categorical intentions, I will first introduce a simple simulated population and discuss a related procedure, described by Udry et al. (1973), on which the procedure I develop is based.

3.2.1 Binary intentions in a simulated population

Consider a population in which a childbearing career lasts for three years, from age one through age three, and births only occur at ages two and three. There are three types of women: type 1 gives birth when age two and age three, type 2 gives birth when age three, and type 3 never gives birth. The population is stable, so there are the same number of women at each age, and the distribution of women by type is the same at each age. Suppose that there are nine women in the population, one of each type observed at each age. Table 4 displays the cumulative number of children, by age and type, in this population (labelled population A).

Each birth occurs at the start of the year, after which each woman is asked to report her fertility intentions. All women have perfect control over conception and intentions exactly predict future fertility. Of the three women observed at age 1, two report that they intend to have a child in the future (only the type 3 woman intends to remain childless). Of the three women observe at age two, only the type 1 woman intends to have a child in the future. No woman observed at age three intends to have a child in the future. No matter their age, all type 1 women intend to have two total children, all type 2 women intend to have one child, and all type 3 women intend to remain childless.

Consider just the six women observed when they are childless. The single type 1 woman will have two children, the two type 2 women will have one child, and the three type 3 women will remain childless. Therefore, 50 percent of childless women have zero children, 33 percent have one child, and 17 percent have two children. Categorical intentions are the same: the same type 1 woman intends to have two children, the same two type 2 women intend to have one child, and the same three type 3 women intend to remain childless.

Udry et al. (1973) propose a procedure for using average binary intentions reported at each family size to predict final number of children. Again start with the six women in the simulated population observed when childless, three of whom (the type 1 and type 2 women) report that they intend to have a child in the future. Therefore, 50 percent of these childless women are predicted to remain childless, while the remaining 50 percent are predicted to have at least one child. Of the two women observed when they already have exactly one child, only one (the type 1 woman) intends to have another child. So, 25 percent (one-half of 50 percent) are predicted to have just one child. Because no woman with two children intends to have a third child, the remaining 25 percent of women are predicted to have two children.

The distribution predicted using binary intentions accurately predicts that 50 percent of childless women will remain childless, but underpredicts the share that will have one child: 25 percent, compared to the actual 33 percent. The share predicted to have two children, 25 percent, exceeds the actual 17 percent. The mean of this binary prediction, 0.75 children, is therefore larger than the actual mean, 0.67 children. The age composition of people observed with one child explains this shortfall. Of the three women who are observed when childless but will go on to have at least one child, just the single type 1 woman then have a second child. Yet, among the two women observed when they have one child, half intend to have another child. Intention to have a third child is overrepresented among women who actually have one child, compared to women who have zero children but are predicted to have a first child.

Resolving this discrepancy requires considering not just parity (number of children a woman has), but also the ages at which each child was born. I predict fertility using binary intentions by adding consideration of these birth histories to the Udry et al. (1973) procedure. For example, consider Angela, the type 1 woman observed at age one. She reports that she

intends to have a child, but this stated intention does not identify her type, and it is not yet revealed whether the child will be born when Angela is aged two or three. I predict the timing of Angela's first birth using the birth histories of all women aged three who also had zero children as of age one. There are two such women, one of whom first gave birth when age two and one when aged three. Because these women have completed childbearing and the population is stable, the timing of their first births predicts that there is a 50 percent likelihood that Angela will give birth when aged two and a 50 percent likelihood she will instead first give birth when aged three.

Suppose Angela gives birth when aged two. There is only one woman observed at age two who has just had her first birth (the same birth history as Angela up to this point), and this woman intends to have another child. Just one woman who has completed childbearing has a birth history in which the first of two births occurs at age one, and this woman's second child was born when she was age two. So, if Angela gives birth when age one, she is also predicted to give birth when age one.

Suppose instead Angela has her first child when she is age two. There is again only one woman observed at age two who has just had her first birth, and this woman does not intend to have another child. So, if Angela gives birth when age two, she is not predicted to have any more children. Angela therefore is predicted to have a zero percent likelihood of having zero children, a 50 percent likelihood of having one child, and a 50 percent likelihood of having two children.

I repeat this procedure for each of the five other women who are observed when childless. Like Angela, the type 2 woman observed at age one intends to have a child, and is predicted to have an equal chance of having one or two total children. The type 2 woman

observed at age two intends to have a child, which is predicted to be born when she is age three. All type 3 women do not intend to have a child, and are therefore predicted to have zero children. I then aggregate these six distributions, yielding a predicted distribution in which 50 percent of women have zero children, 33 percent have one child, and 17 percent have two children. This distribution exactly predicts actual childbearing.

Again, the Udry et al. (1973) procedure for using binary intentions overpredicts average fertility for childless women in population A. The rest of Table 4 simulates alternative populations. Type 4 women give birth when age two. Population B consists of one woman of each type 1, type 3, and type 4 observed at each age. Population C consists of one woman of each type observed at each age. The Udry et al. (1973) procedure underpredicts average fertility in population B (0.53 children per woman, compared to the actual 0.60), and exactly predicts the average of 0.71 children per woman in population C. In each case, though, the revised procedure that uses binary intentions and birth histories exactly predicts fertility. Aggregate binary intentions alone can underpredict, overpredict, or happen to accurately predict fertility, but additionally considering birth histories allows the procedure to always accurately predict fertility in stable populations in which intentions accurately reflect future behavior. Later in the paper, I consider reasons why predicted and actual fertility may differ in an actual population.

3.2.2 Binary intentions in the PSID

I now apply the procedure developed in section 3.2.1 to predict fertility for the 394 childless women observed by the PSID in 1985. Again, for a woman of a certain age who intends to have a child, the procedure predicts when this birth will occur using the birth histories

of all women who have completed childbearing and who had the same birth history through that age. Because few women give birth after age 40, I construct this sample of women who have completed childbearing using all women aged 40 and older, even though I consider binary intentions reported through age 44. This cutoff at age 40 yields a larger group of women who have completed childbearing than would a higher cutoff, allowing more reliable predictions of the likelihood of giving birth at each age.

Column 3 of Table 3 provides this predicted distribution using binary intentions. Thirty-six percent of women indicate that they do not intend to have a baby and are predicted to have zero children, the same share as predicted using categorical intentions. Binary intentions predict that 35 percent of women will have exactly one child, a greater share than the eight percent of women who state that they intend to have exactly one child. Again, the binary prediction is constructed by starting with each women's response to the binary intentions question, and then based on that response, simulating her future childbearing using the intentions of women at higher parities with the same birth history. The comparison with stated categorical intentions indicates that women with zero children are more likely to intend to have two or more total children than are women who have actually already had their first child. At higher parities, this comparison reverses: 36 percent of childless women intend to have two total children and 15 percent intend to have three total children, while binary intentions predict that just 23 percent and four percent of women will do so.

Figure 1 compares the distribution of actual fertility to the distributions predicted using categorical and binary intentions. Each point estimate is bounded by 83-percent confidence intervals, so that any two pairs of confidence intervals that do not overlap indicate that the difference between the two values is statistically significant at the five-percent level. For

example, 36 percent of childless women in 1985 state that they intend to remain childless, substantially less than the 47 percent who actually do so. Binary intentions also predict that 36 percent of women remain childless. The share of women with one child predicted using categorical intentions is substantially closer to the actual share than is the share predicted using binary intentions. However, at every family size between two and four children, the binary distribution more accurately predicts actual fertility, and the categorical distribution is statistically significantly different from the actual distribution.

This comparison between distributions is confirmed in the final two rows of Table 3. Both predicted distributions have an area of overlap of 0.81 in common with the actual distribution of children born. The mean of the categorical distribution, 1.44, is substantially higher than the actual mean of 0.99 because the categorical distribution underpredicts small families and overpredicts larger families. The binary distribution's overprediction of families with one child balances out its underprediction at all other family sizes, and the mean of the binary distribution, 1.00 children per women, nearly exactly matches the actual mean of 0.99 children. This difference between the means of the categorical and binary distributions is substantial and statistically significant, and suggests that binary intentions better predict average fertility.

4 Explaining differences between predicted and actual fertility

The PSID and many other surveys of intentions do not define for the respondent the meaning of intention. For example, the PSID asks "do you intend to have a baby?" If the respondent interprets this question as "do you want to have a baby?" then she would respond yes

if she desires a child. If she instead interprets the question as "do you think you will have a baby?" then she would respond yes if her predicted likelihood of having a child is above some threshold, say 50 percent. I follow Juster (1966), Ajzen and Fishbein (1980), Manski (1990, 2004), and other survey researchers and economists by assuming this second interpretation. The remaining discussion in sections 4 and 5 relies on this interpretation of intentions as a prediction of future behavior, not as a statement of preferences.

Consider a person asked to predict a future outcome. Uncertainty about the outcome occurs when at least one factor that influences the outcome is not known to a person at the time the prediction is recorded. A person with rational expectations is subject to this uncertainty but knows the distribution of these factors, and knows how every possible realization of the factors would influence the outcome. For example, some people struggle to conceive a child, but this infecundity is generally unknown at the start of a person's childbearing career. A person with rational expectations would know the probability of infecundity and would take this information into account when predicting future fertility.

This concept of rational expectations has been discussed in the context of expectations about future earnings (Dominitz 1998), intention to work (Manski 1990), and other predictions, whether elicited as an expectation, as an intention, or some other way. In the remainder of this paper, I similarly use the term rational expectations in regard to fertility intentions. Predictions formed using rational expectations do not necessarily average out to the true outcome. For example, if a group of people all accurately predict that they are likely but not guaranteed to have a child, they may all state that they intend to have a child, even though some will end up childless. Even when intentions are formed using rational expectations, there can still be differences between intended and actual fertility in aggregate.

Manski (1990) establishes a framework for evaluating whether stated intentions are consistent with rational expectations. Using a collection of binary (yes/no) intentions, this framework establishes testable bounds on observed behavior under the assumption that intentions are formed using rational expectations. The test cannot prove that survey respondents have rational expectations. Instead, it determines whether observed behavior is consistent with an assumption that intentions were formed using rational expectations. A wide variety of fertility behavior may fit this criterion.

In appendix A, I use Manski's framework to demonstrate that yes/no responses to the first intentions question ("do you intend to have a baby?") asked by the PSID in 1985 are generally consistent with rational expectations. In appendix B, I extend Manski's framework to allow for yes/no/uncertain responses, and I find that stated intentions are again generally consistent with rational expectations. Only among women with three children in 1985 do stated intentions differ from subsequent fertility in a way that suggests errors in the formation of intentions. In appendix C, I further extend Manski's framework to allow for categorical responses. I find that responses to the second intentions question ("how many [children] do you intend to have?") are consistent with rational expectations for women who state that they intend to have zero or two children. For women who intend to have one, three, or four or more children, intended and actual number of children differ enough to suggest errors in the formation of intentions.

These errors could account for the relative inaccuracy of categorical intentions in predicting the number of children that PSID respondents will have. However, in the rest of this section I demonstrate that, even if categorical intentions are formed using rational expectations, binary intentions can still more accurately predict aggregate fertility. I begin with simulated

populations 1 and 2 in Table 5. In population 1, 45 percent of people have a first child. In population 2, 55 percent of people have a first child. In both populations, 80 percent of people with one child have a second child, and nobody has a third child. These parity progression ratios shape the distribution of children across completed families. In population 1, 55 percent of families have zero children, 9 percent have one child, and 36 percent have two children. In population 2, 45 percent of families have zero children, 11 percent have one child, and 44 percent have two children.

In both populations, the most common number of children is zero. Assuming that a person's categorical intention equals the most likely category, these intentions predict that everyone in both populations will have zero children. On the other hand, binary intentions vary across the two populations. Assuming that a person intends to have a child if the probability of doing so is at least one-half, nobody in population 1 with zero children intends to have a child because only 45 percent of people actually do so. Therefore, in a predicted distribution constructed from binary intentions, nobody in population 1 progresses beyond zero children. In population 2, everybody with zero children intends to have a child, as does everybody with one child. Binary intentions predict that everybody in population 2 has two children.

In population 1, the actual distribution of children across completed families has a mean of 0.81, both predicted distributions have a mean of 0, and the overlap between the actual distribution and each predicted distribution is 0.55. The two ways of using intentions to predict fertility perform equally well. In population 2, the mean of the predicted distribution using categorical intentions, 0, is slightly closer to the mean of the actual distribution, 0.99, than is the mean of the predicted distribution using binary intentions, 2. The overlap between the actual distribution and predicted distribution using categorical intentions is 0.45, slightly larger than the

0.44 overlap between the actual distribution and predicted distribution using binary intentions. Categorical intentions predict fertility better than do binary intentions.

In simulated population 1 and in simulated population 2, everyone has the same binary intention, everyone has the same categorical intention, and both ways of using intentions to predict fertility yield distributions in which all probability is on a single number of children. This number of children can differ across the two predicted distributions, but, by assumption, categorical intentions select the most likely number of children. In any homogeneous population, such as populations 1 and 2, categorical intentions predict fertility at least as well as do binary intentions.

This comparison can reverse in a mixed population. In simulated population 3 in Table 5, half of people belong to population 1 and half to population 2. A particular person's membership in population 1 or 2 is known to the person but not the researcher. Fifty percent of people have zero children, 10 percent have one child, and 40 percent have two children. The mean number of children predicted using binary intentions in this mixed population, 1, more closely matches the actual mean, 0.9, than does the mean predicted using categorical intentions, 0. Similarly, the overlap between the actual distribution and distribution predicted using binary intentions is 0.9, which is greater than the 0.5 overlap between the actual distribution and distribution predicted using categorical intentions. As with PSID respondents, binary intentions can better predict the distribution of children across completed families than do categorical intentions in this simulated mixed population in which all intentions are consistent with rational expectations.

5 Discussion

A central purpose of surveyed intentions is to predict behavior (Manski 1990). To predict the number of children that a group of people will have, the most straightforward approach is to ask each person how many children she intends to have. However, responses to this question can provide a particularly poor prediction. Among a sample of childless women aged 18–44 when observed by the PSID in 1985, average intended number of future children is 1.44, yet these women only end up having 0.99 children on average. Alternatively, stated intentions to have or not have another child, collected from parents with various numbers of current children, can be used to calculate a predicted distribution of children across completed families. Using intentions provided by all women aged 18–44 when observed by the PSID in 1985, the average predicted number of children for childless women is 1.00, substantially closer to the actual realized average of 0.99.

The finding that binary intentions better predict average future fertility is not universal. At higher parities, categorical intentions can provide the better prediction. Panel (a) of Figure 2 repeats Figure 1, the distributions of actual fertility, fertility predicted using categorical intentions, and fertility predicted using binary intentions for childless women observed by the PSID in 1985. Again, as given the final two rows of Table 3, each predicted distribution has the same overlap with actual fertility, but binary intentions better predict the actual mean of 0.99 children. These values are repeated in the first two lines of Table 6. Panel (b) of Figure 2 repeats the calculations for women observed by the PSID with one child in 1985. As noted in Table 6, the categorical distribution better predicts actual fertility, as measured by both the overlap and mean of the distributions, although these differences are not statistically significant.

As given in panels (c) and (d) of Figure 2 and the remaining rows of Table 6, both categorical and binary intentions nearly perfectly predict future fertility for women observed at higher parities.

I repeat this comparison using another longitudinal survey that similarly records intentions. In 1992, the British Household Panel Survey (BHPS) asked "do you think you will have any (more) children?" and, if so, "how many (more) children do you think you will have?" (University of Essex Institute for Social and Economic Research 2010). The BHPS then recorded actual fertility over the following years. Panel (e) of Figure 2 presents the actual, categorical, and binary distributions for childless women observed by the BHPS in 1992. The binary distribution better predicts the shares of women who will actually have one child and two children, but the categorical distribution better predicts the share of women who will actually have three children. As given in Table 6, these two predicted distributions have similar overlap with actual fertility, but the binary distribution better predicts average actual number of children born. At higher parities, the two distributions perform about equally well.

Infecundity, inability to find a suitable partner, or other circumstances may stop a person who intends to have a child from actually doing so. Conversely, circumstances like inadequate contraception may lead a person to have an unintended child. For these reasons, the share of people who intend to have a child can differ from the share that actually do so, even if intentions are formed using rational expectations. However, this paper finds that the number of children that women observed by the PSID actually have is substantially lower on average than the number they intended to have. Errors in the formation of these categorical intentions could explain this difference, as could the realization of future events. For example, the experience

raising a first child may affect a parent's desire for a second child. This learning process may explain the relatively poor prediction of categorical intentions among childless women.

A more fundamental problem with using stated intentions to predict fertility lies in how intentions questions are commonly posed. Intentions questions customarily require the respondent to convert her judgment about the probability of future behavior into a binary or categorical response. Even when respondents have rational expectations, conversion from probabilities into intentions can have counterintuitive consequences. For example, more than half of people in simulated population 2 in Table 5 have at least one child, but the most common family size is zero. It is rational for a person in this population at the start of her childbearing career to state that she both intends to have a child in the future and intends to have zero children. Surveys like the PSID do not allow for this combination of intentions, instead restricting intended number of children to be greater than or equal to one for respondents who intend to have a child.

Asking a respondent to report the probability with which she will engage in a behavior more precisely measures her prediction of the future than does a binary or categorical intention (Juster 1966, Manski 1990). Increasing numbers of surveys elicit these probabilities (Manski 2004). For example, the 1997 National Longitudinal Survey of Youth asked respondents aged 16–22 in 2001 to report their probabilities of having zero, one, two, and three or more children in the future (Bureau of Labor Statistics 2015). Among 456 women who reported a complete predicted probability distribution, summing these distributions yields an aggregate predicted distribution: 30 percent of women are predicted to have zero additional children, 24 percent to have one child, 33 percent to have two children, and 13 percent to have three or more children. By 2013, 34 percent of these women had given birth to zero additional children, 27 percent had

one child, 25 percent had two children, and 13 percent had three or more children. The overlap of 0.92 between the two distributions demonstrates the accuracy of using reported probabilities to predict actual fertility. However, an additional 399 women, or 47 percent of the sample, reported probabilities that do not sum to one. Many survey respondents struggle to provide complete probability distributions.

The main finding in this paper – that binary intentions can predict average fertility better than categorical intentions – suggests an alternative way for surveys to elicit a probability distribution: because children are accumulated sequentially, a predicted probability distribution can be formed from a sequence of binary probabilities. For example, a survey would ask a respondent to report the probability that she will have a child. Regardless of her response, the survey would then ask the respondent to report the probability of having a second child, in the hypothetical case that she did in fact have a first child. Additional questions would elicit hypothetical probabilities upon reaching larger family sizes. This sequence of reported probabilities could then be used to calculate the person's predicted probability distribution. For example, the probability of having only one child equals the probability of having a first child multiplied by one minus the probability of having a second child conditional on having the first child. Such a sequence of conditional probability questions may provide a more complete prediction if respondents struggle to calculate their entire probability distribution all at once.

The comparison between binary and categorical fertility intentions in this paper requires several decades' worth of longitudinal survey data. Determining the most accurate way to ask probability questions about lifetime fertility will similarly require a longitudinal survey to ask probability questions multiple ways. However, a person's number of children is a quantity that is reached sequentially but is reversible: particularly in areas with poor health, child mortality

may distort prediction of total number of children. Sequential quantities that are accumulated more rapidly and are nonreversible, such as months spent trying to conceive a child or weeks spent searching for a job, may allow for a quicker, more accurate test of relationship between predicted probabilities and observed behavior.

Appendix A: A best-case analysis of binary (yes/no) intentions

In this section, I summarize Manski's (1990) framework for using surveyed intentions to estimate bounds on behavior, under the assumption that people have rational expectations. This framework applies to any question in which a survey respondent is asked to state whether she intends, expects, or anticipates engaging in a behavior. These directives may suggest different degrees of certainty a person needs to believe in order to provide a response of "yes." Because I use survey data that record fertility intentions, I refer to intentions throughout this document.

Suppose a person is asked whether she intends to have a child in the future. Information, s, available to her at time of survey and information, z, that is realized later determine whether she has a child, $y(s,z) = 1$, or does not have a child, $y(s,z) = 0$. $P(y|s)$ is the objective probability distribution of y conditional on s, $P_z|s$ is the objective probability distribution of z conditional on s, and $P(y = 1|s) = P_z[y(s,z) = 1|s]$. The respondent does not know what the realization of z will be. However, under the assumption of rational expectations, she knows the distribution of z conditional on s, $P_z|s$, and she knows the function, $y(s,\cdot)$, that maps between every possible realization of z and her future childbearing. Whether she states that she intends to have a child, $i = 1$, or does not intend to have a child, $i = 0$, depends on the relative losses she associates with the two possible prediction errors ($i = 0$, $y = 1$; and $i = 1$, $y = 0$). If the probability of having a child is above some threshold, π, she states that she intends to have a child:

$$i = 0 \Rightarrow P(y = 1|s) \leq \pi,$$
$$i = 1 \Rightarrow P(y = 1|s) \geq \pi. \qquad (1.1)$$

A researcher may wish to predict whether the respondent will have a child. The researcher has access to some subset, x, of the full information, s, available to the respondent, and the researcher knows the respondent's stated intention, i. $P_s|xi$ is the probability distribution of s conditional on x and i. The probability that the respondent has a child, given her observable x and i, is $P(y = 1|x, i) = \int P(y = 1|s)dP_s|xi$. Therefore, assuming that π is the same for everyone and known to the researcher,

$$P(y = 1|x, i = 0) \leq \pi \leq P(y = 1|x, i = 1). \quad (1.2)$$

In order for inequality 1.2 to be defined, the probability of having a child must be higher among people who intend to have a child than among people who do not intend to have a child.

The probability of having a child is defined as:

$$P(y = 1|x) \equiv P(y = 1|x, i = 0)P(i = 0|x) + P(y = 1|x, i = 1)P(i = 1|x). \quad (1.3)$$

Inequality 1.2 and equation 1.3 can be combined to provide bounds on $P(y = 1|x)$. Because probabilities are non-negative, $P(y = 1|x, i = 0) \geq 0$. From inequality 1.2, $\pi \leq P(y = 1|x, i = 1)$. Therefore, $P(y = 1|x) \geq P(y = 1|x, i = 1)P(i = 1|x) \geq \pi P(i = 1|x)$. Similarly, because probabilities cannot exceed one, $P(y = 1|x, i = 1) \leq 1$. From inequality 1.2, $P(y = 1|x, i = 0) \leq \pi$. Therefore, $P(y = 1|x) \leq \pi P(i = 0|x) + P(i = 1|x)$. Together, these two inequalities yield bounds on $P(y = 1|x)$:

$$\pi P(i = 1|x) \leq P(y = 1|x) \leq \pi P(i = 0|x) + P(i = 1|x). \quad (1.4)$$

Stated intentions are consistent with rational expectations if the probability of having a child lies within this range.

Bounds 1.2 and 1.4 can be tested using survey data, under the assumption of random sampling. Testing the bounds further requires that future events, z, are not subject to aggregate shocks. For example, assume that intention to have a child is based entirely on ability to have a child, and women face a 10 percent chance of infecundity. If this risk of infecundity is independent across women, then $P(y = 1|s) = P_z[y(s, z) = 1|s] = 0.9$. However, if infecundity instead affects or does not affect all women, then $P(y = 1|s)$ equals 0.9 but $P_z[y(s, z) = 1|s] \in \{0,1\}$. The best-case analysis in this paper therefore evaluates whether stated intentions are consistent with rational expectations, assuming that each respondent states her best prediction, π is the same for everyone and known to the researcher, and there are no aggregate shocks.

I test bounds 1.2 and 1.4 using PSID data. In 1985, the Panel Study of Income Dynamics (PSID) asked, "Looking to the future, do you intend to have a baby at some time?" (Panel Study of Income Dynamics 2017). The PSID permitted three responses: "Yes," "No," and "Maybe; Don't Know." I omit all respondents who chose the third option (in appendix B, I extend Manski's framework to permit this uncertain option). The PSID then recorded all births over the subsequent decades.

Table A.1 presents estimates for women aged 18–44 when surveyed in 1985 who continued to be observed until at least age 40. I group women by the total number of children to whom they had given birth when surveyed in 1985, x. As given in column 1, among the 401

women with zero children[4], 66 percent stated that they intend to have a child in the future, and 34 percent stated that they do not intend to have a child. The share of women who subsequently had a child is 76 percent among women who stated that they intend to do so, and 16 percent among women who stated that they did not intend to do so. These shares provide bounds on the probability threshold, π, above which a woman states that she intends to have a child. That these bounds (and many of the bounds in the rest of appendices A, B, and C) are wide indicates that, given observed intentions, an assumption that these intentions were formed using rational assumptions is not especially restrictive. Intentions formed using rational expectations can be consistent with a variety of actual fertility outcomes. Finally, the overall share of women who had a child after 1985, 56 percent, falls within the estimated bounds of 33 percent and 83 percent, calculated assuming a symmetric loss function of $\pi = 0.5$.

The remaining columns of Table A.1 similarly estimate bounds 1.2 and 1.4 among women with larger families in 1985. For women with zero, one, or two children, bounds 1.2 include $\pi = 0.5$. Assuming $\pi = 0.5$, bounds 1.4 all include the observed share of women who had a child after 1985. The assumption of a symmetric loss function is not required to reach this conclusion. Alternative assumed thresholds, such as $\pi = 0.25$, also satisfy bounds 1.2 and 1.4. Across 1,000 samples drawn with replacement, these findings generally hold: bounds 1.2 are defined for women in at least 95 percent of samples, and bounds 1.4 are satisfied at all parities in all samples. Even though some women have a child after stating they do not intend to do so, and

[4] These 401 women include the 394 women from column 2 of Table 1, as well as seven more who state that they intend to have a child in the future, but decline to report how many children they intend to have.

vice versa, these estimates indicate that, under the best-case hypothesis in which the threshold is the same for all women and known to the researcher, stated intentions are generally consistent with rational expectations.

Appendix B: A best-case analysis of yes/no/uncertain intentions

In this section, I extend Manski's framework to estimate best-case probability bounds on behavior using intentions questions that permit three responses: yes, no, and uncertain. In the binary case, there is some threshold, π, on the probability of having a child, below which the respondent states an intention of no and above which the respondent states an intention of yes. Allowing for an uncertain response requires two thresholds on the probability of having a child: below the lower threshold, α, the respondent states an intention of no, $i = 0$; above the upper threshold, β, the respondent states intention of yes, $i = 1$; and between the two thresholds the respondent states an intention of uncertainty, $i = 9$:

$$\begin{aligned} i = 0 \Rightarrow \quad & P(y = 1|s) \leq \alpha, \\ i = 9 \Rightarrow \alpha \leq\ & P(y = 1|s) \leq \beta, \\ i = 1 \Rightarrow \quad & P(y = 1|s) \geq \beta. \end{aligned} \tag{2.1}$$

Therefore, the probability of having a child should be highest among people who state that they intend to have child, lowest among people who state that they do not intend to have a child, and between the two for people who express uncertainty:

$$P(y = 1|x, i = 0) \leq \alpha \leq P(y = 1|x, i = 9) \leq \beta \leq P(y = 1|x, i = 1). \tag{2.2}$$

Again, the condition under which these bounds on α and β are defined is not especially strict, and just requires that the probability of having a child increases with stated intention to do so.

By definition,

$$P(y=1|x) \equiv P(y=1|x,i=0)P(i=0|x) + P(y=1|x,i=9)P(i=9|x) + P(y=1|x,i=1)P(i=1|x). \quad (2.3)$$

Inequality 2.2 and equation 2.3 can be combined to provide bounds on $P(y=1|x)$. Because probabilities are non-negative, $P(y=1|x,i=0) \geq 0$. By inequality 2.2, $P(y=1|x,i=0) \geq \alpha$ and $P(y=1|x,i=9) \geq \beta$. Therefore, $P(y=1|x) \geq \alpha P(i=9|x) + \beta P(i=1|x)$ Similarly, because probabilities cannot exceed one, $P(y=1|x,i=1) \leq 1$. By inequality 2.2, $P(y=1|x,i=0) \leq \alpha$ and $P(y=1|x,i=9) \leq \beta$. Therefore, $P(y=1|x) \leq \alpha P(i=0|x) + \beta P(i=9|x) + P(i=1|x)$. Together, these two inequalities yield bounds on $P(y=1|x)$:

$$\alpha P(i=9|x) + \beta P(i=1|x) \leq P(y=1|x) \leq \alpha P(i=0|x) + \beta P(i=9|x) + P(i=1|x). \quad (2.4)$$

I test bounds 2.2 and 2.4 using PSID data. Table B.1 presents estimates of these bounds using the same sample and intention question as in appendix A, but now including respondents whose stated intention is "Maybe; Don't Know" ($i=9$). As given in column 1, among the 465 women with zero children in 1985, 29 percent stated that they do not intend to have a child in the future, 14 percent expressed uncertainty, and the rest stated that they intend to have a child. The share of women who subsequently had a child is 16 percent among women who stated that they

did not intend to do so, 34 percent among women who expressed uncertainty, and 76 percent among women who stated that they intend to have a child. These shares provide bounds on the probability thresholds α and β. Finally, the overall share of women who had a child after 1985, 53 percent, falls within the estimated bounds of 34 percent and 75 percent, calculated assuming that $\alpha = 0.3$ and $\beta = 0.5$.

The remaining columns of Table B.1 similarly estimate bounds 2.2 and 2.4 among women with larger families in 1985. All bounds on α are defined, as are bounds on β in all cases except for women with three children.[5] Assuming that $\alpha = 0.3$ and $\beta = 0.5$, bounds 2.4 all contain the observed shares of women who had a child after 1985. These assumed thresholds imply that a woman provides an uncertain fertility intention when her predicted probability of having a child is between 30 and 50 percent. Unlike the assumption of $\pi = 0.5$ in appendix A, this loss function is not symmetric and suggests a greater willingness to fail to have a child after stating an intention of yes, compared to having a child after stating an intention of no. However, as in appendix A, alternative assumed thresholds would yield the same conclusion.

These estimates suggest that, even though stated intentions sometimes disagree with subsequent fertility, intentions are generally consistent with rational expectations. Across 1,000 samples drawn with replacement, these findings generally hold. Bounds 2.2 on α are defined in

[5] The share of women with three children who have a child is higher among women who report uncertainty, 56 percent, than among women who report that they intend to have a child, 28 percent. One possible reason for this pattern is that women with several children implicitly interpret the binary intentions question as asking "are you too old to have another child" and are inclined to answer "no."

at least 93 percent of the samples at all parities. Bounds 2.2 on β are defined in at least 73 percent of the samples for all women except those with three children. Bounds 2.4 contain the observed share of women who have a child for at least 99 percent of the samples at all parities.

Appendix C: A best-case analysis of categorical intentions

In this section, I extend Manski's framework to estimate best-case probability bounds on behavior using intentions questions that permit a single choice among multiple unordered options. The three possible intentions in appendix B (no, uncertain, and yes) are ordered in the probability of the outcome. With more thresholds, the approach in appendix B could be modified to permit additional ordered responses (such as no, unlikely, even chance, likely, and yes). When the response options are unordered in probability, it is no longer possible to identify thresholds, and another criterion must be used. I assume that respondents select the modal category. That is, when a person who can have at most m children is asked to state how many children she intends to have, she reports the number that she judges to be most likely:

$$i = q \Rightarrow P(y = q|s) \geq P(y = r|s) \quad \forall q, r \in \{0,1,2, \dots, m\}, \;\; q \neq r. \tag{3.1}$$

I assume that there is a single mode. Therefore, among respondents who state an intention to have q children, the probability of having q children should be greater than the probability of having any other number of children:

$$P(y = q|x, i = q) \geq P(y = r|x, i = q) \quad \forall q, r \in \{0,1,2, \dots, m\}, \;\; q \neq r. \tag{3.2}$$

By definition,

$$P(y = q|x) \equiv \sum_{r=0}^{m} P(y = q|x, i = r)P(i = r|x) \quad \forall q \in \{0,1,2, \dots, m\}. \tag{3.3}$$

Inequality 3.2 and equation 3.3 can be combined to provide bounds on $P(y = q|x)$. Equation 3.3 can be rewritten as $P(y = q|x) \equiv P(y = q|x, i = q)P(i = q|x) + P(y = q|x, i \neq q)P(i \neq q|x)$. Because probabilities are non-negative, $P(y = q|x, i \neq q) \geq 0$. By assumption, if a respondent states that she intends to have q children, then she views this as the most likely outcome. The probability of this outcome, $P(y = q|x, i = q)$, should therefore be at least $1/(m + 1)$. Therefore:

$$P(y = q|x) \geq P(y = q|x, i = q)P(i = q|x) \geq P(i = q|x)/(m + 1).$$

Similarly, because probabilities cannot exceed one, $P(y = q|x, i = q) \leq 1$. By assumption, if a respondent states that she intends to have $r \neq q$ children, then she does not view q as the most likely outcome. The probability of having q children should be at most one-half: $P(y = q|x, i \neq q) \leq 1/2$. Therefore,

$$\begin{aligned} P(y = q|x) &\leq P(i = q|x) + \frac{1}{2}P(i \neq q|x) \\ &= P(i = q|x) + \frac{1 - P(i = q|x)}{2} \\ &= \frac{1 + P(i = q|x)}{2}, \end{aligned}$$

Together, these two inequalities yield bounds on $P(y = q|x)$:

$$\frac{P(i = q|x)}{m+1} \leq P(y = q|x) \leq \frac{1 + P(i = q|x)}{2} \quad \forall q \in \{0,1,2,\dots,m\}. \tag{3.4}$$

I test bounds 3.2 and 3.4 using PSID data. Among respondents who state that they intend to have a child in the future, the PSID further asked "How many do you intend to have?" For respondents who state that they do not intend to have a child in the future, I assign an intended number of children equal to zero. Therefore, as in appendix A, the sample excludes women who are uncertain about whether they intend to have a child in the future. Table C.1 presents estimates for the 394 women who currently have $x = 0$ children. Each column of this table represents a number of intended children, q. As given in the first row, the most commonly intended numbers of children are zero and two. Few women intend to have or end up having more than four children, so I top-code the number of children at four.

The third through seventh rows of Table C.1 present statistics used to evaluate bounds 3.2. Among the 136 women who state that they intend to have zero children, 84 percent in fact do not have a child. This share is greater than the share of women who have any other number of children, satisfying bounds 3.2. Similarly, among women who state that they intend to have two children, the most common outcome is two children. However, among women who intend to have one, three, or four or more children, this intention exceeds the modal outcome by one child, violating bounds 3.2.

Among women who state that they intend to have zero children, 44 percent do not have a child. This share falls within the estimated bounds 3.4 of seven percent and 67 percent.

Similarly, for women with all other intended numbers of children, bounds 3.4 contain the observed share who end up with that number of children. Across 1,000 samples drawn with replacement, these findings generally hold. Bounds 3.2 are satisfied for women who intend to have zero or two children in at least 99 percent of samples. Bounds 3.4 are satisfied at all intended family sizes below 4, and in 79 percent of samples for women who intend to have four or more children.

Therefore, among women who state that they intend to have zero or two children, these intentions are consistent with the assumption of rational expectations. Among women who report intentions for another family size, bounds 3.2 are violated. This violation may indicate error in the formation of intentions, but could also be a consequence of some women using an alternative criterion rather than selecting the modal category. For example, a woman may feel social pressure to avoid stating intention to have a particular number of children, leading her to select a number other than her predicted mode.

Manski's framework for testing whether intentions are consistent with rational expectations could be further extended to consider alternative ways of eliciting intended number of children, such as allowing a respondent to report a minimum and maximum intended number of children. By forcing these two values to be equal, the PSID prevents a respondent from reporting a range of number of children that she is equally likely to have. The framework could also be applied to other fertility predictions, such as expected time to next birth. For example, annually between 1969 and 1972, the PSID asked respondents to report whether they expect to have a child in the future. Respondents who expect to have a child were then asked to indicate whether they expect the child to be born within the next year.

Disclosure of potential conflicts of interest

This work was supported by the NICHD (T32 HD007339) as part of the University of Michigan Population Studies Center training program. The author gratefully acknowledges the use of the services and facilities of the Population Studies Center (funded by NICHD Center Grant R24 HD041028). The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and the National Science Foundation under award number 1157698. The opinions and conclusions expressed herein are solely the author's and do not represent the opinions or policy of these funders or any agency of the federal government.

The author declares that he has no conflict of interest.

Works cited

Ajzen, Icek, and Martin Fishbein. 1980. *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, N.J.: Prentice-Hall.

Bachrach, Christine A., and S. Philip Morgan. 2013. "A Cognitive–Social Model of Fertility Intentions." *Population and Development Review*, 39(3): 459–485.

Beaujouan, Eva, and Anne Solaz. 2013. "Racing Against the Clock? Childbearing and Sterility among Men and Women in Second Unions in France." *European Journal of Population*, 29(1): 39–67.

Berrington, Ann. 2004. "Perpetual Postponers? Women's, Men's and Couple's Fertility Intentions and Subsequent Fertility Behaviour." *Population Trends*, 117: 9–19.

Bureau of Labor Statistics. 2015. *National Longitudinal Survey of Youth 1997 cohort, 1997-2013 (rounds 1-16)*. Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

Cornolli, Chiara Ludovica, and Fabrizio Bernardi. 2015. "The Causal Effect of the Great Recession on Childlessness of White American Women." *IZA Journal of Labor Economics*, 4(1): 1–24.

DaVanzo, Julie, Christine E. Peterson, and Nathan R. Jones. 2003. "How Well Do Desired Fertility Measures for Wives and Husbands Predict Subsequent Fertility? Evidence from Malaysia." RAND Labor and Population Program, Working Paper Series, 03-16.

Dominitz, Jeff. 1980. "Earnings Expectations, Revisions, and Realizations." *Review of Economics and Statistics*, 80 (3): 374–388.

Freedman, Deborah S., and Arland D. Thornton. 1982. "Income and Fertility: The Elusive Relationship." *Demography*, 19(1): 65–78.

Freedman, Ronald, Deborah S. Freedman, and Arland D. Thornton. 1980. "Changes in Fertility Expectations and Preferences between 1962 and 1977: Their Relation to Final Parity." *Demography*, 17(4): 365–378.

Hayford, Sarah R. 2009. The Evolution of Fertility Expectations over the Life Course." *Demography*, 46(4): 765–783.

Heckman, James J., and Robert J. Willis. 1976. "Estimation of a Stochastic Model of Reproduction: An Econometric Approach." In Nestor E. Terleckyj, editor, *Household Production and Consumption*. New York: National Bureau of Economic Research.

Juster, F. Thomas. 1966. "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design." *Journal of the American Statistical Association*, 61(315): 658–696.

Lee, Ronald D. 1980. "Aiming at a Moving Target: Period Fertility and Changing Reproductive Goals." *Population Studies*, 34(2): 205–226.

Manski, Charles F. 1990. "The Use of Intentions Data to Predict Behavior: A Best-Case Analysis." *Journal of the American Statistical Association*, 85(412): 934–940.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica*, 72(5): 1329–1376.

Miller, Warren B., Joseph Lee Rodgers, and David J. Pasta. 2010. "Fertility Motivations of Youth Predict Later Fertility Outcomes: A Prospective Analysis of National Longitudinal Survey of Youth Data." *Biodemography and Social Biology*, 56(1): 1–23.

Modrek, Sepideh, and Negar Ghobadi. 2011. "The Expansion of Health Houses and Fertility Outcomes in Rural Iran." Studies in Family Planning, 42(3): 137–146.

Morgan, S. Philip, and Heather Rackin. 2010. "The Correspondence between Fertility Intentions and Behavior in the United States." *Population and Development Review*, 36(1): 91–118.

Panel Study of Income Dynamics. 2017. Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. Available at <https://psidonline.isr.umich.edu>, accessed July 4, 2017.

Quesnel-Vallée, Amélie, and S. Philip Morgan. 2003. "Missing the Target? Correspondence of Fertility Intentions and Behavior in the U.S." *Population Research and Policy Review*, 22(5/6): 497–525.

Schoen, Robert, Nan Marie Astone, Young J. Kim, Constance A. Nathanson, and Jason M. Fields. 1999. "Do Fertility Intentions Affect Fertility Behavior?" *Journal of Marriage and Family*, 61(3): 790–799.

Stolzenberg, Ross M., and Linda J. Waite. 1977. "Age, Fertility Expectations and Plans for Employment." *American Sociological Review*, 42(5): 769–783.

Tan, Poo Chang, and Nai Peng Tey. 1994. "Do Fertility Intentions Predict Subsequent Behavior? Evidence from Peninsular Malaysia." *Studies in Family Planning*, 25(4): 222–231.

Toulemon, Laurent, and Maria Rita Testa. 2005. "Fertility Intentions and Actual Fertility: A Complex Relationship." *Population and Societies*, 415: 1–4.

Udry, J. Richard, Karl E. Bauman, and Charles L. Chase. 1973. "Population Growth Rates in Perfect Contraceptive Populations." *Population Studies*, 27(2): 365–372.

University of Essex Institute for Social and Economic Research. 2010. *British Household Panel Survey: Waves 1-18, 1991-2009*. Seventh Edition. UK Data Service.

Westoff, Charles F., and Norman B. Ryder. 1977. "The Predictive Validity of Reproductive Intentions." *Demography*, 14(4): 431–453.

Westoff, Charles F., Robert G. Potter, and Philip C. Sagi. 1963. *The Third Child: A Study in the Prediction of Fertility*. Princeton: Princeton University Press.

Williams, Lindy, Joyce Abma, and Linda J. Piccinino. 1999. "The Correspondence between Intention to Avoid Childbearing and Subsequent Fertility: A Prospective Analysis." *Family Planning Perspectives*, 31(5): 222–227.

Table 1: Sample characteristics in 1985

	All	Zero children by 1985	
		Followed until age 40	Not followed until age 40
Number of women aged 18–44	3,251	394	202
Average age, in years	31.9 (0.2)	28.8 (0.4)	25.3 (0.4)
Percent white	85.6 (0.7)	92.2 (1.4)	89.5 (2.3)
Percent married	66.2 (1.1)	40.0 (2.7)	55.0 (4.5)
Percent completed high school	85.1 (0.8)	97.2 (0.9)	91.5 (2.1)

Notes: The sample consists of women who were present in the household and responded for themselves; who reported their current number of children; who were not pregnant; who reported the year of birth of every child they have had; who never adopted a child; who never had a twin or other multiple birth; and who answered the initial intentions question, "do you intend to have a baby at some time?" The sample in columns 2 and 3 is further restricted to women who either respond "No" when asked whether they intend to have a child in the future, or who respond "Yes" and then also provide an intended number of children. See section 2. Standard errors are provided in parentheses. *Data source*: Panel Study of Income Dynamics.

Table 2: Regression of future fertility on demographic characteristics

	Estimated coefficient
Average age, in years	-0.097*** (0.0069)
Percent white	0.032 (0.13)
Percent married	0.48*** (0.097)
Percent completed high school	0.37 (0.35)
Constant	3.2*** (0.41)

Notes: The dependent variable is number of children after 1985. The sample consists of 394 women who were childless when observed in 1985. R^2 is 0.35. Standard errors are provided in parentheses. Statistical significance at the 10 percent, five percent, and one percent levels is denoted by *, **, and ***. See section 2. *Data source*: Panel Study of Income Dynamics.

Table 3: Actual and predicted number of children in the PSID

	Actual future fertility	Predicted using categorical intentions	Predicted using binary intentions
Share with 0 children	0.47 (0.03)	0.36 (0.03)	0.36 (0.03)
Share with 1 child	0.17 (0.02)	0.08 (0.02)	0.35 (0.05)
Share with 2 children	0.28 (0.03)	0.36 (0.03)	0.23 (0.04)
Share with 3 children	0.08 (0.01)	0.15 (0.02)	0.04 (0.02)
Share with 4 children	0.01 (0.00)	0.03 (0.01)	0.01 (0.01)
Share with 5 children	0.00 (0.00)	0.01 (0.00)	0.00 (0.00)
Share with 6+ children	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Overlap with actual		0.81 (0.03)	0.81 (0.04)
Mean	0.99 (0.06)	1.44 (0.07)	1.00 (0.07)

Notes: Column 1 provides the distribution of subsequent births to 394 women who were childless in 1985. Column 2 provides the distribution of intended number of births, stated by these same women in 1985. Column 3 provides a distribution predicted for these women using binary intentions collected in 1985. Bootstrap standard errors are provided in parentheses, calculated using 1,000 samples drawn with replacement from the original sample. See section 3.2.2. *Data source*: Panel Study of Income Dynamics.

Table 4: Simulated populations

	Cumulative number of children when observed, by age			
	Type 1	Type 2	Type 3	Type 4
Age 1	0	0	0	0
Age 2	1	0	0	1
Age 3	2	1	0	1

	Among women who were childless when observed, share with each completed number of children			
	Actual future fertility	Predicted using categorical intentions	Predicted using binary intentions, Udry et al. (1973) procedure	Predicted using binary intentions, incorporating birth histories
Population A: One woman of type 1, type 2, and type 3 observed at each age				
0 children	0.50	0.50	0.50	0.50
1 child	0.33	0.33	0.25	0.33
2 children	0.17	0.17	0.25	0.17
Mean	0.67	0.67	0.75	0.67
Population B: One woman of type 1, type 3, and type 4 observed at each age				
0 children	0.60	0.60	0.60	0.60
1 child	0.20	0.20	0.27	0.20
2 children	0.20	0.20	0.13	0.20
Mean	0.60	0.60	0.53	0.60
Population C: One woman of each type observed at each age				
0 children	0.43	0.43	0.43	0.43
1 child	0.43	0.43	0.43	0.43
2 children	0.14	0.14	0.14	0.14
Mean	0.71	0.71	0.71	0.71

Notes: The population consists of nine women, one of each type observed at each age. See section 3.2.1.

Table 5: Additional simulated populations

	Population 1			Population 2			Population 3: ½ Pop 1, ½ Pop 2		
Number of children	0	1	2	0	1	2	0	1	2
Actual fertility									
Conditional on reaching this number of children, share of people that have another child	0.45	0.8	0	0.55	0.8	0	0.5	0.8	0
Distribution of completed number of children	0.55	0.09	0.36	0.45	0.11	0.44	0.5	0.1	0.4
Categorical intentions									
Share of people who state they intend to have this number of children	1	0	0	1	0	0	1	0	0
Predicted distribution of completed number of children	1	0	0	1	0	0	1	0	0
Binary intentions									
Conditional on reaching this number of children, share of people who state they intend to have another child	0	1	0	1	1	0	0.5	1	0
Predicted distribution of completed number of children	1	0	0	0	0	1	0.5	0	0.5
Overlap between actual distribution and									
Predicted distribution using categorical intentions		0.55			0.45			0.5	
Predicted distribution using binary intentions		0.55			0.44			0.9	
Mean of									
Actual distribution		0.81			0.99			0.9	
Predicted distribution using categorical intentions		0			0			0	
Predicted distribution using binary intentions		0			2			1	

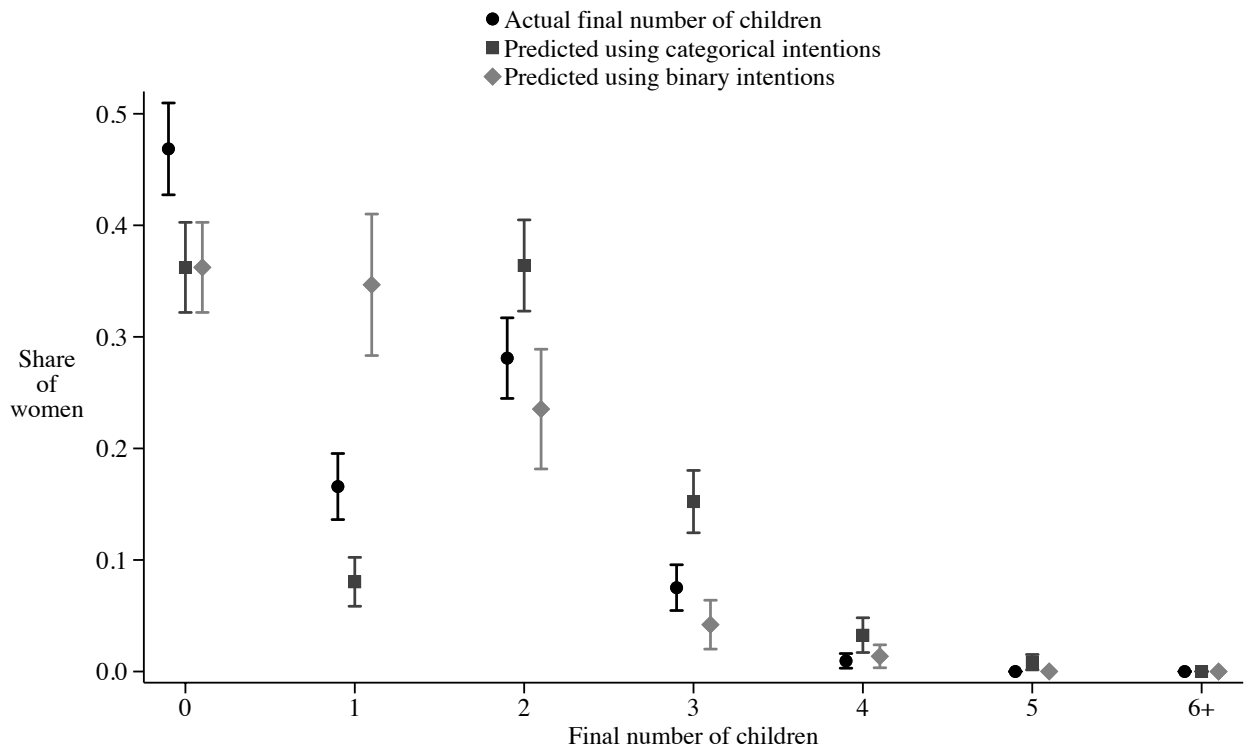
Notes: The actual distribution of children across completed families in the third row is calculated from the parity progression ratios in the second row. Intentions are formed using rational expectations. For categorical intentions, a person indicates that she intends to have a certain number of children if she judges that number to be the most likely. For binary intentions, a person indicates that she intends to have a child if the probability of doing so is more than 50 percent. Half of population 3 consists of parents from population 1, the other half from population 2. See section 4.

Table 6: Actual and predicted number of children in the PSID and BHPS, by parity

	Panel Study of Income Dynamics			British Household Panel Survey		
	Actual final number of children	Predicted using categorical intentions	Predicted using binary intentions	Actual final number of children	Predicted using categorical intentions	Predicted using binary intentions
Currently have 0 children						
Overlap with actual		0.81 (0.03)	0.81 (0.04)		0.79 (0.02)	0.83 (0.03)
Mean	0.99 (0.06)	1.44 (0.07)	1.00 (0.07)	0.81 (0.06)	1.25 (0.07)	1.14 (0.08)
Currently have 1 child						
Overlap with actual		0.95 (0.02)	0.90 (0.03)		0.96 (0.02)	0.95 (0.02)
Mean	1.62 (0.05)	1.55 (0.05)	1.49 (0.04)	1.40 (0.04)	1.47 (0.05)	1.45 (0.04)
Currently have 2 children						
Overlap with actual		0.98 (0.01)	0.99 (0.01)		0.98 (0.01)	0.98 (0.01)
Mean	2.12 (0.02)	2.11 (0.01)	2.10 (0.01)	2.13 (0.02)	2.11 (0.02)	2.10 (0.01)
Currently have 3+ children						
Overlap with actual		0.99 (0.01)	0.99 (0.01)		0.98 (0.01)	0.99 (0.01)
Mean	3.57 (0.03)	3.55 (0.03)	3.54 (0.03)	3.45 (0.02)	3.45 (0.02)	3.44 (0.02)

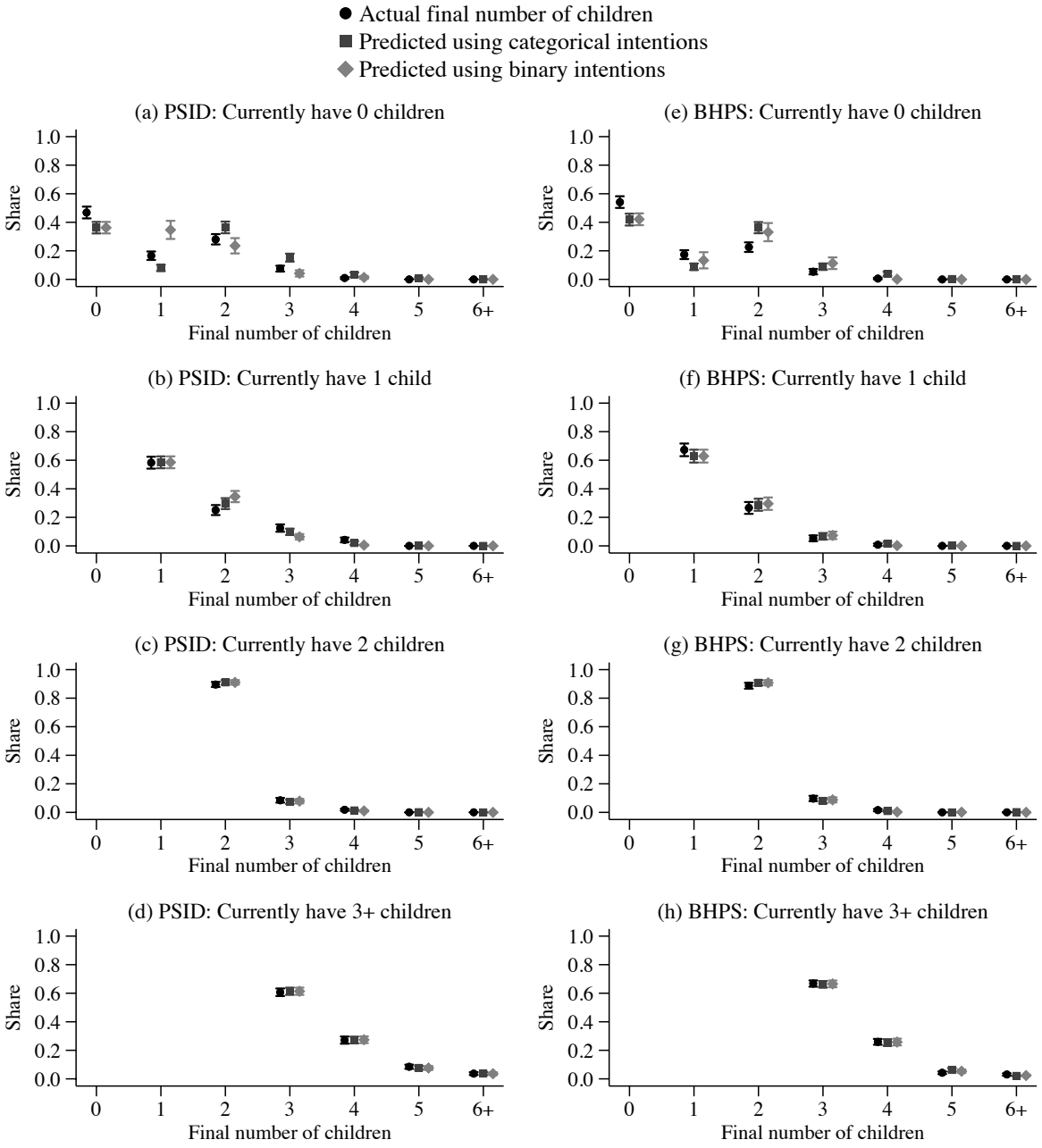
Notes: See notes to Table 3. PSID intentions recorded in 1985, BHPS intentions recorded in 1992. Bootstrap standard errors are provided in parentheses, calculated using 1,000 samples drawn with replacement from the original sample. See section 5. *Data sources*: Panel Study of Income Dynamics; British Household Panel Survey.

Figure 1: Actual and predicted final number of children, among currently childless women



Notes: Distribution from the first six rows of Table 3. Bootstrapped 83-percent confidence intervals calculated using 1,000 samples drawn with replacement from the original sample. See section 3.2.2. *Data source*: Panel Study of Income Dynamics.

Figure 2: Actual and predicted number of children, by current number of children



Notes: Distributions generated according to methods described in section 3. Bootstrapped 83-percent confidence intervals calculated using 1,000 samples drawn with replacement from the original sample. See section 5. *Data sources*: Panel Study of Income Dynamics; British Household Panel Survey.

Table A.1: Consistency between yes/no fertility intentions and actual fertility

	Number of children when question is asked (x)				
	0	1	2	3	4+
Number of women	401	456	755	426	247
$P(i = 0\|x)$	0.34	0.55	0.89	0.94	0.97
$P(i = 1\|x)$	0.66	0.45	0.11	0.06	0.03
$P(y = 1\|x, i = 0)$	0.16	0.17	0.07	0.05	0.03
$P(y = 1\|x, i = 1)$	0.76	0.79	0.62	0.28	0.43
$P(y = 1\|x)$	0.56	0.45	0.14	0.07	0.04
Bounds 1.2 on π	[0.16, 0.76]	[0.17, 0.79]	[0.07, 0.62]	[0.05, 0.28]	[0.03, 0.43]
Bounds 1.4 on $P(y = 1\|x)$	[0.33, 0.83]	[0.23, 0.73]	[0.06, 0.56]	[0.03, 0.53]	[0.01, 0.51]
Bounds 1.2 on π are defined	Yes	Yes	Yes	Yes	Yes
Bounds 1.4 contain observed $P(y = 1\|x)$	Yes	Yes	Yes	Yes	Yes
Across 1,000 samples, percentage in which					
Bounds 1.2 on π are defined	100	100	100	99	95
Bounds 1.4 contain observed $P(y = 1\|x)$	100	100	100	100	100

Notes: Bounds 1.2 on the threshold probability π above which a respondent indicates that she intends to have a child are estimated according to inequality 1.2. Bounds 1.4 are estimated according to inequality 1.4, assuming $\pi = 0.5$. 1,000 samples are drawn with replacement from the original sample. See appendix A. *Data source*: Panel Study of Income Dynamics.

Table B.1: Consistency between yes/no/uncertain fertility intentions and actual fertility

	Number of children when question is asked (x)				
	0	1	2	3	4+
Number of women	465	508	803	444	257
$P(i = 0\|x)$	0.29	0.49	0.83	0.90	0.93
$P(i = 9\|x)$	0.14	0.10	0.06	0.04	0.04
$P(i = 1\|x)$	0.57	0.41	0.11	0.06	0.03
$P(y = 1\|x, i = 0)$	0.16	0.17	0.07	0.05	0.03
$P(y = 1\|x, i = 9)$	0.34	0.44	0.48	0.56	0.30
$P(y = 1\|x, i = 1)$	0.76	0.79	0.62	0.28	0.43
$P(y = 1\|x)$	0.53	0.45	0.16	0.09	0.05
Bounds 2.2 on α	[0.16, 0.34]	[0.17, 0.44]	[0.07, 0.48]	[0.05, 0.56]	[0.03, 0.30]
Bounds 2.2 on β	[0.34, 0.76]	[0.44, 0.79]	[0.48, 0.62]	N/A	[0.30, 0.43]
Bounds 2.4 on $P(y = 1\|x)$	[0.33, 0.73]	[0.23, 0.60]	[0.07, 0.39]	[0.04, 0.35]	[0.03, 0.33]
Bounds 2.2 on α are defined	Yes	Yes	Yes	Yes	Yes
Bounds 2.2 on β are defined	Yes	Yes	Yes	No	Yes
Bounds 2.4 contain observed $P(y = 1\|x)$	Yes	Yes	Yes	Yes	Yes
Across 1,000 samples, percentage in which					
Bounds 2.2 on α are defined	100	100	100	100	93
Bounds 2.2 on β are defined	100	100	93	3	73
Bounds 2.4 contain observed $P(y = 1\|x)$	100	100	100	100	99

Notes: Bounds 2.2 on the threshold probabilities α (below which a respondent states that she does not intend to have a child) and β (above which a respondent states that she intends to have a child) are estimated according to inequality 2.2. Bounds 2.4 are estimated according to inequality 2.4, assuming $\alpha = 0.3$ and $\beta = 0.5$. 1,000 samples are drawn with replacement from the original sample. See appendix B. *Data source*: Panel Study of Income Dynamics.

Table C.1: Consistency between categorical fertility intentions and actual fertility

	Number of children intend to have (q)				
	0	1	2	3	4+
Number of women	136	39	146	58	15
$P(i = q\|x)$	0.35	0.10	0.37	0.15	0.04
$P(y = 0\|x, i = q)$	0.84	0.41	0.26	0.09	0.13
$P(y = 1\|x, i = q)$	0.09	0.33	0.20	0.29	0.13
$P(y = 2\|x, i = q)$	0.06	0.23	0.42	0.45	0.27
$P(y = 3\|x, i = q)$	0.01	0.03	0.12	0.12	0.47
$P(y = 4 + \|x, i = q)$	0.00	0.00	0.01	0.05	0.00
$P(y = q\|x)$	0.44	0.19	0.27	0.09	0.01
Bounds 3.4 on $P(y = q\|x)$	[0.07, 0.67]	[0.02, 0.55]	[0.07, 0.69]	[0.03, 0.57]	[0.01, 0.52]
Bounds 3.2 are satisfied	Yes	No	Yes	No	No
Bounds 3.4 contain observed $P(y = q\|x)$	Yes	Yes	Yes	Yes	Yes
Across 1,000 samples, percentage in which					
Bounds 3.2 are satisfied	100	30	99	0	0
Bounds 3.4 contain observed $P(y = q\|x)$	100	100	100	100	79

Notes: Bounds 3.2 are estimated according to inequality 3.2. Bounds 3.4 are estimated according to inequality 3.4. 1,000 samples are drawn with replacement from the original sample. See appendix C. *Data source*: Panel Study of Income Dynamics.