# Measuring the Effect of the Timing of First Birth*

Jane Leber Herr†

November 2011

**Abstract**

The existing literature on the effect of the timing of first birth on women's wages generally concludes that there is a benefit to fertility delay, but one that is overstated by the raw correlation. In this paper I reconsider this question, but begin by diverging from the literature to redefine "timing" in terms of a woman's entry into the labor force, rather than her age. When one considers the mechanisms by which fertility timing may affect a woman's wage path, it is clear that each turns on her experience level at first birth, not her age. This transformation also reveals the important distinction between women who have their first child after they enter the labor market versus before. Applying this measure to women from the 1979 cohort of the National Longitudinal Survey of Youth, I show the following. First, when measured correctly, there is little bias captured in the correlation between first-birth timing and wages. Second, the existing literature's focus on age has produced confused inference that has obscured the magnitude, and for some the *sign*, of the link between fertility timing and wages. In particular, estimates based on age at first birth understate the return to delay for women who remain childless at labor market entry. But more importantly, these positive estimates obscure the *negative* return to delay - to having a first birth after labor market entry, rather than before - for the majority of all but college graduates.

*I would like to thank David Card, Ronald Lee, Guido Imbens, Jesse Shapiro, Elizabeth Weber Handwerker, Melanie Guldi, Emily Oster, Colleen Manchester, Sarah Hayford, Lucie Schmidt, Jenna Johnson-Hanks, Ashley Langer, Kevin Stange, Donna Ginther, Martha Bailey, Rachel Sullivan Robinson and participants of the U.C. Berkeley labor seminar series and the University of Chicago's Booth applied microeconomic series for helpful comments. I would also like to thank Jay Zagorsky and Steve McClaskie from the Center of Human Resource Research at the Ohio State University for providing me the sampling weights appropriate for my sample. Lastly, I would like to thank the Center for Human Potential and Public Policy at the University of Chicago's Harris School, and its director, Ariel Kalil, for support and guidance. Financial support for this work was provided by the National Institute for Child Health and Human Development (Interdisciplinary Training Grant No. T32-HD007275).

†The Harris School of Public Policy Studies, University of Chicago, 1155 East 60th Street, Chicago, IL 60637, jlherr@uchicago.edu.

# 1 Introduction

It has long been established that women who delay entry into motherhood have systematically higher long-run wages (Bloom 1986). Identifying the causal effect of first-birth timing is complicated, however, by women's capacity to (imperfectly) control their fertility. Although the existing literature has consistently found positive bias captured in the correlation between first-birth timing and wages, it has also concluded that there remains a benefit to fertility delay. Chandler *et al.* (1994) and Miller (2011) find that a one-year delay of first birth is associated with 2 to 3 percent higher wages, measured among women in their 30s or 40s. Other research has found that a first birth beyond age 28 or 30 closes the "motherhood wage gap" between women with and without children (Taniguchi 1999; Amuedo-Dorantes and Kimmel 2005).

A key limitation of this research, however, is its focus on *age* at first birth. As I show in this paper, to gauge the effect of fertility timing on wages, the appropriate measure is instead a woman's "career timing" of first birth, the point in her labor market career in which children are first present. Furthermore, I show that the existing literature's focus on age has produced confused inference that obscures the magnitude, and for some the *sign*, of the link between timing and wages. I also show that when timing is measured correctly, there is little bias captured in the positive correlation between first-birth timing and long-run wages.

To set the stage for assessing the effect of first-birth timing, I begin by considering the mechanisms by which timing may affect a woman's wage path. This thought experiment reveals a first key insight: each potential mechanism turns on a woman's *experience* level at first birth, not her age. I therefore begin my analysis by translating age at first birth into a woman's "relative timing" – the difference between the year of her first birth and the year she entered the labor force.

This transformation then reveals a second key insight: the relative timing can only

reflect a woman's experience level at first birth for those who have their first child *after* they enter the labor force. For women who instead have a first birth before working, the mechanism by which fertility delay may affect subsequent wages is more ambiguous.

Formalizing these thoughts into a model of the effect of first-birth timing on a woman's wage path, I then show the confused inference that can arise by instead measuring timing in terms of age. First, I show that the coefficient estimate on age is likely to *under*-state the return to delay for those women who remain childless at labor market entry. And second, I show that this coefficient will obscure the relationship between timing and wages for women who instead enter the labor force after motherhood, and may in fact indicate the wrong sign.

I next return to the question of the potential bias captured in the raw correlation between first-birth timing and wages. The existing literature has used various econometric methods to address this identification problem, such as fixed effects or instrumental variables (IV) approaches. Rather than relying on such identification strategies, in this paper I instead use the wage model described above to formalize the economic conditions under which bias may be captured in the estimated return to delay. Focusing on women who have their first child after labor market entry, I show that each source of bias discussed in the literature makes clear predictions for how wages diverge between early and late mothers.

As a first test of causality, I then compare these predictions to the pattern of wages for women from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). Counter to expectations, I find that the raw wage pattern for women who remain childless at labor market entry is strikingly similar to the wage path consistent with an unbiased reflection of the causal effect of fertility delay.

I follow this with a simple OLS framework that considers how the measured relationship between first-birth timing and wages varies as one controls for an increasing number of factors that are correlated with either (*i*) wage growth, (*ii*) taste for early motherhood, or (*iii*) the individual return to fertility delay. In this exercise I now consider the effect of these controls

on the estimated slope between timing and wages for both populations of mothers – those with a first birth before labor market entry, and those with a first birth afterwards. For the latter, I again find little bias captured in the correlation between timing and wages. For the former, I find that the observed positive correlation is driven by unobserved heterogeneity.

Overall, these results show that there is a clear benefit to continued fertility delay for women who enter the labor force before motherhood. For those with at least a high school diploma when they start working, a year of fertility delay leads to 1.5 to 3 percent higher wages 20 years later. By contrast, I find that there is no *linear* relationship between fertility timing and wages for women who have their first child before they start working. I therefore conclude that the relevant measure of the timing of a woman's first birth is her "career timing," the point in her career when children are first present.

Yet counter to expectations, for all but college graduates, I also find clear evidence suggesting that the majority of women who had their first child *after* labor market entry, would have had higher long-run wages if they instead had their first birth earlier, *before* entering the labor force. Thus I find an initial *negative* return to fertility delay – at least across the threshold at labor market entry – followed by a positive return to subsequent delay further into one's career.

The remainder of this paper is structured as follows. Section 2 discusses the rationale for measuring fertility timing in terms of the point of labor market entry, and the implications of instead measuring in terms of age. Section 3 then discusses the key sources of potential bias that may be captured in the correlation between first-birth timing and wages, and how each will influence the path of wages over time. In Section 4 I then introduce my sample and variable definitions, followed by a discussion of my identification strategy in Section 5. Section 6 reports my results, and Section 7 concludes.

# 2 Measurement of the Timing of First Birth

In this section I begin by discussing the rationale for measuring the timing of first birth in terms of the year that a woman enters the labor market. I then discuss the specification error that can arise if timing is instead measured in terms of age.

## 2.1 Potential Mechanisms for the Effect of Fertility Delay

To gauge the appropriate measure of fertility timing, consider the mechanisms by which timing may affect a woman's wages. In the wage profile of male workers, wages rise quickly over a worker's first decade in the labor market, grow at an increasingly more moderate pace in the subsequent decade, and slowly flatten out from that point forward (Murphy and Welch, 1990). Taking this as the potential wage path for a woman entering the labor force at age 20, we see that the steepest period of wage growth falls during her key childbearing years, and wages do not fully plateau until after her childbearing years have passed.

Since wage growth slows throughout this period, one can imagine that a woman faces a payoff to delaying her first birth – putting off for one more year the discontinuous shock to time demands that arise at motherhood. More formally, if the timing of a woman's first birth affects her wages, it must do so through its influence on the underlying economic factors that generate the shape of the standard wage path.[1]

For instance, if a key factor is an especially high rate of return on early-career work experience (Light and Ureta 1995), the economic return to fertility delay may arise from the payoff to one more year of full-intensity work experience during this initial stage.[2] Fertility

[1] One might also ask whether the *reason* for fertility delay influences the return, if women's subsequent labor market behavior varies systematically with that reason. One can imagine isolating the factor that was the binding constraint the year before we see a woman get pregnant with her first child. For instance, "I would have gotten pregnant one year earlier, but (a) I wasn't married yet; (b) my husband's earnings weren't high enough yet; (c) it took me a long time to get pregnant; (d) I was up for a big promotion; (e) I didn't want a baby yet." (Or, in the reverse, (f) "I would have waited one more year, but I got pregnant on accident.") One can therefore think of the average return to a year of delay reflecting the weighted average of these individual returns. See footnote 50 for more on this question.

[2] For the vast majority of mothers, the arrival of a first child marks a clear drop in her labor force presence, whether or not she stops working. For instance, in the NLSY79, among women who are working full-time before motherhood, at least two thirds cut their hours by at least 10 percent in the following year, and only

timing may also influence wages through its effect on a woman's labor supply if the greater human capital accumulated by the time of a delayed first birth lowers her propensity to leave the labor force at motherhood, or the length of time she takes off. Alternatively, following job match theory (Topel and Ward 1992), if the arrival of a baby discontinuously increases the transaction costs of job search, fertility delay may affect a woman's wage path by allowing her to reach a better-quality job match before the transition into motherhood.

Although the purpose of this paper is not to establish the mechanism by which fertility timing affects wages, this thought experiment highlights the fact that each of these possibilities centers on the timing of a woman's first birth in terms of experience, not age. I therefore begin by calculating a woman's "relative timing," $K_1$ – the difference between the year of her first birth and the year she enters the labor force, $t_1$. Thus $K_1 > 0$ for those who have their first child after they begin working, and $K_1 \leq 0$ for those who enter the labor force after motherhood.

This transformation then reveals a second possibility that is obscured by the focus on age: the effect of fertility delay may be non-monotonic. In particular, $K_1$ only reflects the "experience timing" of a woman's first birth for those who remain childless at $t_1$.[3]

For those who instead enter the labor market with children, the mechanism for the link between first-birth timing and wages is more ambiguous. One possibility is that the key factor driving the effect of fertility timing is the timing of the discontinuous change in the marginal value of time at home that occurs at the birth of a first child. If this is the case, then for women with $K_1 \leq 0$, the timing itself is irrelevant; what matters is only that children are already present at the point when they enter the labor force.

Alternatively, $K_1$ may have a direct effect on the wage path of these "pre-$t_1$" ($pt_1$)

20 percent do not reduce their labor supply (even when calculated only among women who remain working).

[3]Although $K_1$ only technically reflects a woman's *potential* experience at first birth, in practice it reflects actual experience for the vast majority of these women. Considering each calendar year between $t_1$ and the year before first birth, if I calculate the proportion of these years in which a woman worked at least 1000 hours, the average proportion across these women is 0.92.

mothers if there is a link between the age of a woman's oldest child at $t_1$ (which for these women is by definition equal to $-K_1$), and her subsequent labor market behavior. For instance, women who enter the labor force with older children may work longer hours, leading to greater wage growth. If so, among $pt_1$ women, the wage effect of fertility delay will be *negative*.[4]

Furthermore, even without such a direct link between timing and wages among $pt_1$ mothers, one can imagine scenarios in which it is strictly better to have one's first birth before labor market entry, rather than shortly thereafter. This would suggest a negative return to delay across the threshold at $t_1$. Thus my redefinition of the timing of first birth highlights the fact that in some range, the direction of the effect of delay on wages may in fact switch sign.

## 2.2 Implications of Measuring Timing in Terms of Age

To consider the implications of the existing literature's focus on age at first birth, it is useful to translate the ideas introduced above into a formal model of the effect of the timing of first birth on a woman's wage path. First consider the possibility that the key element of fertility timing is the point in a woman's career in which children are first present. Since this makes timing otherwise irrelevant for $pt_1$ women, I define the "career timing" of first birth as $k_1 \equiv max(0, K_1)$.

Now suppose that a mother's wage path is a function of $k_1$ in the following way:

$$\begin{aligned} w_t &= w_0 e^{g_1 t} && \forall\ t \leq k_1, \quad \text{and} \\ w_t &= w_0 e^{g_1 k_1} e^{g_2 (t-k_1)} && \forall\ t > k_1, \quad \text{with} \quad g_1 > g_2, \end{aligned} \tag{1}$$

where wages grow at a rate $g_1$ up till the career timing of first birth, $k_1$, at which point wage

[4]The more important factor may instead be the number of children a woman has at $t_1$, or the age of her *youngest*. Although the link between $K_1$ and the latter is ambiguous, total children will be monotonically decreasing in $K_1$. If women with more children at $t_1$ make labor supply choices that lead to lower wage growth, the slope on $K_1$ may therefore be positive.

growth stalls to a slower rate, $g_2$.[5] Note that, by construction, in Equation (1) the wages of $pt_1$ women always grow at the lower rate $g_2$.

Accordingly, a given woman's wage level at some point $\tau > k_1$ will be the following:

$$l(w_\tau) = l(w_0) + \theta k_1 + g_2 \tau, \tag{2}$$

where the coefficient on $k_1$ is the change in wage growth at that point, $\theta \equiv (g_1 - g_2)$. Thus under these assumptions, for women who remain childless at $t_1$, wages are a linear function of their timing of first birth, whereas among $pt_1$ mothers, wages and timing are unrelated.

In the second possibility described above, the wages of $pt_1$ mothers are instead related to their fertility timing through the age of their oldest child at labor market entry ($-K_1$). For instance, suppose that the wage growth rate for $pt_1$ women, $g_{pt_1}$, is a function of fertility timing in the following way: $g_{pt_1} = g^0_{pt_1} + g'_{pt_1}(-K_1)$. (In the example described above, in which mothers of older children work longer hours, $g'_{pt_1} > 0$). Equation (3) reflects wage growth to time $\tau$ under these conditions:

$$l(w_\tau) = l(w_0) + \theta k_1 + (1 - pt_1) g_2 \tau + pt_1 (g^0_{pt_1} - g'_{pt_1} K_1)\tau. \tag{3}$$

The third possibility discussed above assumes no link between $K_1$ and wages for $pt_1$ mothers ($g'_{pt_1} = 0$), but suggests that women who have their first child before $t_1$ experience greater long-run wage growth than women who have their first child shortly thereafter (corresponding to $g^0_{pt_1} > g_2$). Namely, while $pt_1$ mothers may experience slower wage growth early on (if $g^0_{pt_1} < g_1$), their wages may catch up to those of some later mothers when those women interrupt their careers to have children. Alternatively, if $g^0_{pt_1} = g_2$ (and $g'_{pt_1} = 0$),

[5] One might instead assume that wage growth stalls only temporarily, as in the model in Wilde *et al.* (2010). Figure 3, however, shows no evidence of a return to the pre-birth growth rate, at least within the 10- to 15-year post-birth window observed. Also note that Equation (1) does not model the possible direct effects on a woman's wages of various labor market behaviors (such as the choice to change jobs at first birth, or the length of time taken off from the labor market), because, to the extent that timing influences these choices, they may reflect the *mechanism* by which fertility timing affects a woman's wage path.

Equation (3) reduces to Equation (2), and it is strictly better to have one's first birth after labor market entry.

Taking Equation (3) as the most flexible representation of the effect of first-birth timing on women's wages, now consider the implications of instead estimating the effect of timing using age at first birth, $a_{b_1}$. Rewriting Equation (3) to focus on $K_1$,

$$l(w_\tau) = l(w_0) + \Big(\theta(1 - pt_1) - g'_{pt_1} pt_1\Big) K_1 + \Big(g_2(1 - pt_1) + g^0_{pt_1} pt_1\Big)\tau,$$

we can further substitute for $K_1$ in terms of age at first birth, since $a_{b_1} = a_{t_1} + K_1$ (where $a_{t_1}$ is age at $t_1$). We therefore see that a least squares regression of $l(w_\tau)$ on $a_{b_1}$ will capture an estimate of $\Big(\theta(1 - pt_1) - g'_{pt_1} pt_1\Big)$.

If the sample used in such a regression includes only women with a first birth *after* labor market entry, then the coefficient on $a_{b_1}$ will recover the return to delay for this population, $\theta$. As the proportion of $pt_1$ mothers rises, however, it is likely that the coefficient will increasingly underestimate $\theta$. The two scenarios described above suggest that $g'_{pt_1}$ is either positive or zero. Although one can envision alternate scenarios in which $g'_{pt_1}$ is negative – that delay also increases wages for $pt_1$ mothers – intuition suggests that any return to delay in the years *before* labor market entry will be weaker than the return to delay beyond $t_1$. In a mixed sample, the coefficient on $a_{b_1}$ will therefore only recover $\theta$ in the unlikely scenario that $g'_{pt_1} = -\theta$, namely, if the slope of wages on $K_1$ is the same on both sides of $t_1$.[6]

Since the majority of women have their first birth after labor market entry, the coefficient on $a_{b_1}$ will likely do an even poorer job of reflecting the return to delay for $pt_1$ mothers, and may even reflect the wrong sign. Assuming that $\theta$ is non-negative, this will clearly hold if $g'_{pt_1} > 0$. Yet this may hold even if $g'_{pt_1} = 0$, if it is also the case that $g^0_{pt_1} > g_2$, as in the scenario in which it is better to have a first birth before $t_1$ rather than shortly thereafter.

[6]Surprisingly, if I pool my sample across education levels, in the specification before I control for any individual characteristics (see description in Section 5), the slope on $K_1$ is in fact almost identical on both sides of $t_1$. This result disappears as soon as I add additional controls.

## 3 Potential Sources of Bias in Estimating the Return to Delay

Now I return to the question of how to identify the effect of first-birth timing, given women's capacity to control their fertility. The existing literature has relied on various identification strategies to tackle this issue. Some research has used a fixed effects approach to address unobserved heterogeneity, as in Geronimus and Korenman (1992), Taniguchi (1999), and Amuedo-Dorantes and Kimmel (2005). Other papers have used an IV approach, for instance relying on health shocks such as miscarriages, as in Hotz *et al.* (2005) and Miller (2011).

In the following section, I instead use the wage model introduced above to formalize the conditions under which the observed positive correlation between timing and wages may reflect these various potential sources of bias. For each, I then plot the corresponding path of wages predicted for 'early' versus 'late' mothers.

To begin, consider how women choose their optimal timing of first birth, $k_1^*$. Suppose women are maximizing the following lifetime utility function, $U = log(y) - c(k_1)$, where earnings, $y$, are a function of $k_1$ via the wage path in Equation (1), and the cost of delay, $c(k_1)$, varies across women with their taste for early motherhood, $\psi$. Note that for the sake of simplicity, I am focusing on the ways in which bias may influence the estimated return to delay for women who have their first child after labor market entry ($\theta$), and am therefore building off of Equation (1), defined on the "career timing" of first birth, $k_1$.

To solve for a woman's marginal benefit of fertility delay, we must first solve for the production function of earnings, $y = f(k_1)$. Making the simplifying assumption that labor supply is constant throughout the lifecycle, the net present value of lifetime earnings is

$$f(k_1) = \int_T w_t h_t e^{-rt} dt = \int_0^{k_1} w_0 e^{g_1 t} e^{-rt} dt + \int_{k_1}^{T} w_0 e^{g_1 k_1} e^{g_2 (t-k_1)} e^{-rt} dt,$$

where $r$ reflects the discount rate and $T$ the length of a woman's career.[7] This, in turn,

[7] This assumes $h_t = 1$ for all $t$. If I instead assume that $h_t$ varies by the presence and age of children, but

solves to

$$f(k_1) = \frac{w_0(g_2 - g_1)}{(g_1 - r)(g_2 - r)} e^{(g_1 - r)k_1} + \frac{w_0}{(g_2 - r)} e^{(g_2 - r)T} e^{(g_1 - g_2)k_1} - \frac{w_0}{(g_1 - r)}. \tag{4}$$

The marginal benefit of fertility delay will then equal $f'(k_1)/f(k_1)$. Taking the linear approximation of Equation (4), $MB(k_{1i}) = \theta_i - m_{1i}k_{1i}$, where $m_{1i} \geq 0$.[8] Thus for each woman, the marginal benefit of delaying her first birth by one more year has an intercept term equal to the change in her wage growth rate at first birth, $\theta_i = g_{1i} - g_{2i}$, and the return is decreasing in $k_1$.

I can similarly approximate the marginal cost of delay, $c'(k_1)$, as a linear function of a woman's taste for early motherhood, $\psi$: $MC(k_1) = \psi_i + m_{2i}k_{1i}$, where I assume $m_{2i} \geq 0$.[9] A woman will therefore choose $k^*_{1i}$, her optimal timing of first birth, at the point where the marginal benefit of one more year of delay is just equal to its marginal cost, solving to:

$$k^*_{1i} = \frac{\theta_i - \psi_i}{m_{1i} + m_{2i}}.$$

---

not by the timing of the first:

$$\begin{array}{lll} h_t = 1 & \text{for } t < k_1 \text{ and } t > k_1 + \Delta, & \text{and} \\ h_t = (1-\delta) & \text{for } t \in \{k_1, k_1 + \Delta\}, & \end{array}$$

the production function of earnings, $f(k_1)$, is now equal to

$$f(k_1) = \int_0^{k_1} w_0 e^{g_1 t} e^{-rt} dt + \int_{k_1}^{k_1+\Delta} (1-\delta) w_0 e^{g_1 k_1} e^{g_2(t-k_1)} e^{-rt} dt + \int_{k_1+\Delta}^{T} w_0 e^{g_1 k_1} e^{g_2(t-k_1)} e^{-rt} dt,$$

$$= \Big[\frac{w_0}{(g_1 - r)} - \frac{w_0}{(g_2 - r)}\Big(\delta(e^{(g_2 - r)\Delta} - 1) + 1\Big)\Big] e^{(g_1 - r)k_1} + \frac{w_0}{(g_2 - r)} e^{(g_2 - r)T} e^{(g_1 - g_2)k_1} - \frac{w_0}{(g_1 - r)}.$$

(As expected, this expression reduces to Equation (4) if $\delta = \Delta = 0$.) Without calculating the full linearization of the corresponding marginal benefit function, I can show that the intercept term is again a function of $\theta_i$:

$$= \theta_i \frac{\Big(e^{(g_{2i} - r)T} - 1\Big) - \delta(g_{1i} - r)\Big(e^{(g_{2i} - r)\Delta} - 1\Big)}{\Big(e^{(g_{2i} - r)T} - 1\Big) - \delta\Big(e^{(g_{2i} - r)\Delta} - 1\Big)}.$$

[8] The slope on $k_{1i}$, $m_{1i} \equiv \theta_i \frac{g_{2i} - r}{e^{(g_{2i} - r)T} - 1}$, is non-negative, regardless of the relative size of $g_{2i}$ and $r$.

[9] Since the marginal cost of delay will be a function of declining fertility, one could rewrite this in terms of age at labor market entry, $a_{t_1}$.

Scenario #1: No Bias

Taking this as the underlying framework, now consider a first possible scenario, in which the rate of wage growth before and after first birth, $g_1$ and $g_2$, are equal for all women. The return to delay is therefore invariant across the population, $\theta_i = \theta$ for all $i$. In this scenario, the marginal benefit of delay is thus constant, and observed variation in timing of first birth will arise only through variation in taste for early motherhood, $\psi_i$.

The first panel of Figure 1 plots the path of wages under this scenario, comparing two women, one who chooses an early first birth, $k_1^E$, and the other who chooses a late first birth, $k_1^L$. Assuming an equivalent starting wage, we see that their wages move together, growing at the rate $g_1$, up till the point when the 'early' mother has her first birth, $t = k_1^E$. Her wage path then begins to grow at the slower rate $g_2$, while the later mother's wage continues to grow at $g_1$ until she has her first birth at $t = k_1^L$.

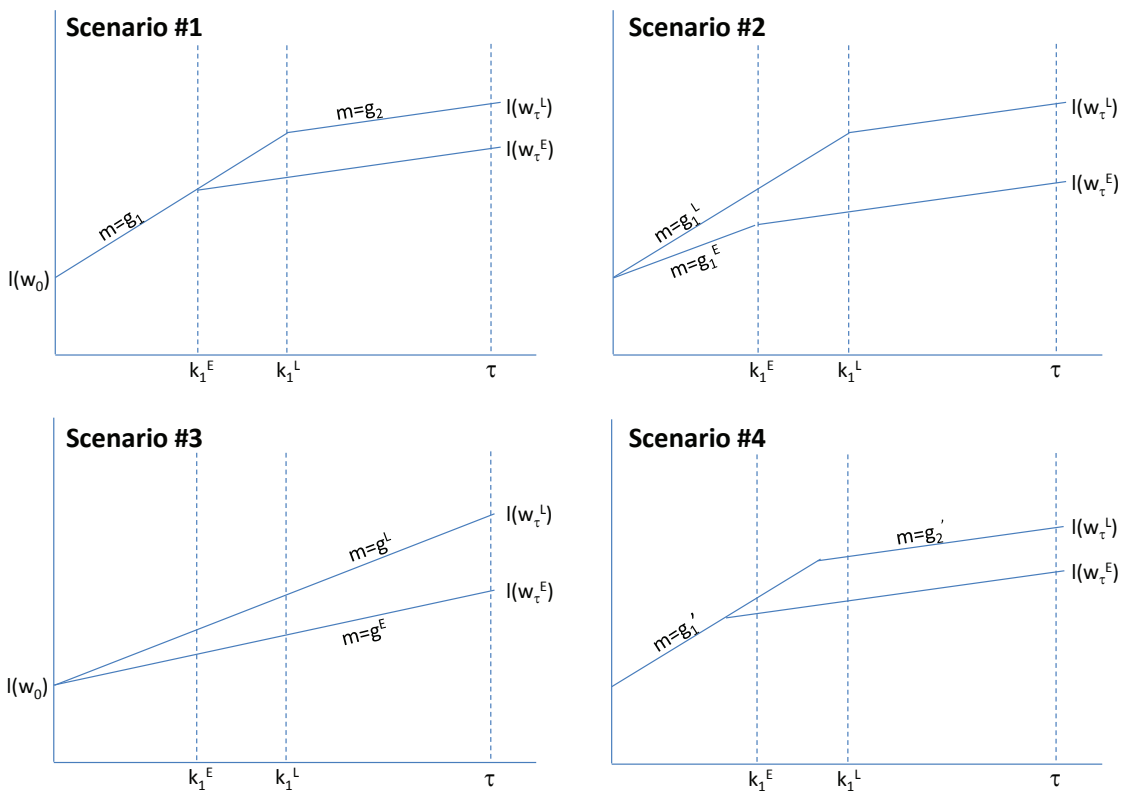


Figure 1: Possible Wage Paths by Timing of First Birth

If we then trace their wage path through some point $\tau$ beyond $k_1^L$, we see that the difference in their final wage level is equal to $\theta$, the return to one more year of fertility delay:

$$l(w_\tau^L) - l(w_\tau^E) = \Big(l(w_0) + \theta k_1^L + \tau g_2\Big) - \Big(l(w_0) + \theta k_1^E + \tau g_2\Big) = \theta(k_1^L - k_1^E) = \theta,$$

where I have made the simplifying assumption that $k_1^L - k_1^E = 1$. Under this scenario, running a least squares regression of $l(w_{\tau i})$ on observed timing of first birth, $k_{1i}$, will therefore provide an unbiased estimate of the average return to delay, $\bar{\theta} = \theta$.

Scenario #2: Endogenous Timing of First Birth

Now suppose that the observed correlation between timing and wages is influenced by endogenous timing of first birth in response to variation in the return to delay. For instance, suppose $\theta$ is an increasing function of ability, $a_i$, as in Mullin and Wang (2002). To translate this idea into the terms of Equation (1), suppose $\theta$ is increasing in ability because $g_{1i} = g_1(a_i)$, with $g_1' > 0$ (and $g_2$ constant).[10] A woman's marginal benefit of delay is now increasing in her ability, and thus, all else equal, higher-ability women choose later first births.[11]

The second panel of Figure 1 plots the wage path under this scenario of endogenous timing, where the two women now vary in ability, $a^E < a^L$. Maintaining the assumption that they start at the same wage level, we see that because $g_1^E < g_1^L$, their wages diverge *before* either have children. If we again let $k_1^L = k_1^E + 1$, comparing their wage level at $\tau$,

$$ln(w_\tau^L) - ln(w_\tau^E) = \theta^L k_1^L - \theta^E k_1^E = \theta^L + k_1^E(\theta^L - \theta^E) = \theta^L + k_1^E(g_1^L - g_1^E),$$

we see that the difference is equal to $\theta^L$ (the return for the higher-ability/later mother), *plus* the difference in their wage level that has already emerged by $t = k_1^E$. Thus under this scenario, a regression of $l(w_{\tau i})$ on $k_{1i}$ will provide a coefficient that not only over-estimates the average return to delay, $\bar{\theta}$, but even over-states the return for higher-ability mothers, $\theta^L$.

[10] Although $g_2$ may also be a function of ability, in order for $\theta_i$ to increase with ability, the slope on $a_i$ for $g_1$ must exceed the slope for $g_2$.

[11] Although $\theta$ appears in both the numerator and denominator of $k_1^*$, for reasonable values of $g_2$, $r$, and $T$, the term in the numerator dominates.

Scenario #3: Unobserved Heterogeneity

Now consider a third scenario, in which the observed correlation between first-birth timing and wages is biased by unobserved heterogeneity (the focus of Geronimus and Korenman [1992], Taniguchi [1999], and others). For simplicity, assume the extreme case in which fertility timing has no influence on wages: $\theta = 0$, namely $g_1 = g_2 \equiv g$. Under these conditions, the marginal benefit of delay is zero, and variation in first-birth timing again only arises from variation in taste for early motherhood, $\psi$.

Yet to satisfy the universal observation that long-run wages and first-birth timing are positively correlated, in this scenario it must be the case that $\psi$ is correlated with factors that directly influence women's wages. For instance, suppose $\psi_i$ and $a_i$ are negatively correlated, suggesting the observed relationship between timing and wages is driven by ability bias. If this holds, 'late' mothers will again be higher-ability women, who in turn have greater wage growth if we assume $g_i = g(a_i)$, $g' > 0$. Alternatively, as in Blackburn *et al.* (1993), women with high $\psi$ may choose to invest less in human capital, thus $g_i = g(\psi_i)$, $g' < 0$.

The third panel of Figure 1 shows the path of wages under Scenario #3. We see that there is now no kink in the wage path at $k_1$ for either mother, and instead wages diverge immediately upon labor market entry because $g^L > g^E$. If we again calculate the level difference in wages at $\tau$,

$$l(w_\tau^L) - l(w_\tau^E) = \tau g^L - \tau g^E = \tau(g^L - g^E),$$

we see that it is unrelated to $k_1$, suggesting that a regression of $l(w_{\tau i})$ on $k_{1i}$ should find no slope. Yet because of the correlation between $\psi$ and ability – and because $\psi$ drives $k_1$ – if we cannot fully control for ability, OLS will suffer from omitted variables bias, providing a positive coefficient on $k_1$.

Scenario #4: Reverse Causality

Lastly, consider a fourth scenario of reverse causality, in which the wage path drives fertility timing, rather than the reverse (Miller [2011] considers the milder case of simultaneity).[12] For instance, suppose that all women enter the labor force with a wage growth rate of $g_1'$, but that at some variable time in their career, each reaches a critical point, $t_c$, where their wage growth stalls to a lower rate $g_2'$. Suppose also that women choose to delay first birth until after $t_c$, thus 'late' mothers are those who hit $t_c$ later in their career.

The last panel of Figure 1 shows the path of wages under this case of reverse causality. As with Scenario #1, we see that the early-career wages of the two mothers are identical. Yet if women cannot perfectly anticipate $t_c$, in Scenario #4 their wages diverge *before* $k_1^E$. If we let $\vartheta \equiv g_1' - g_2'$, the level wage difference at $\tau$ will equal $\vartheta(t_c^L - t_c^E)$, which again is unrelated to $k_1$. But if women take roughly the same amount of time to conceive and give birth, $t_c^L - t_c^E \approx k_1^L - k_1^E$, a regression of $l(w_{\tau i})$ on $k_{1i}$ will produce an OLS estimate consistent with $\vartheta$, but that does not reflect a causal effect of fertility timing on wages.

The purpose of this exposition has been to highlight the ways in which these potential sources of bias will influence both the correlation between first-birth timing and the long-run wage level, and the path of wages over time. After introducing my sample below, I will return to compare these figures to the raw path of wages observed for the women in the NLSY79, as a first gauge of the sources of bias that may be captured in the observed correlation between first-birth timing and long-run wages.

# 4 Data and Variable Definitions

In this analysis I rely on data for the women from the NLSY79, who by 2008 had reached the ages of 44 to 51.[13] In the following section I first discuss how I calculate their relative

[12] See, for instance, Heckman and Walker (1990) and Perry (2005) for evidence on wages influencing timing.

[13] Because the NLSY79 is well known, I leave a more detailed discussion of the data to the appendix. As discussed there, my sample builds from both the cross-sectional and minority supplemental samples of the NLSY79, thus throughout this analysis I report weighted results. The appendix also provides a more detailed discussion of how I define the variables introduced here, and explores whether my sample selection criteria may influence my results.

timing of first birth, $K_1$, based on the year of labor market entry, $t_1$. I also consider whether $t_1$, or a woman's education level at that point, are endogenous to her timing of first birth. I then discuss my dependent variable, $w_{20}$, the wage level approximately 20 years after labor market entry, and the selection issues surrounding the women for whom I lack an observed value. Last, I discuss my sample selection criteria, and report some summary statistics for my final sample.

## 4.1 Defining Relative Timing of First Birth, $K_1$

To determine the year a woman entered the labor market, $t_1$, I first establish her 'graduation year,' roughly the year in which she completed continuous schooling.[14] From that point forward, I search for the first twelve-month period in which she worked at least 1000 hours, defining this as $t_1$.[15] Given $t_1$, and the calendar year of a woman's first birth, $B_1$, I then define $K_1$, the relative timing: $K_1 = B_1 - t_1 + 1$.

Figure 2 shows the distribution of $K_1$ for the population of NLSY79 women observed through age 40.[16] From the dip in this distribution at $K_1 = \{0, 1\}$ – women who have their first birth within a year before or after entering the labor force – it is clear that $t_1$ is endogenous to $B_1$, at least within this range. One can further imagine such endogeneity for many $pt_1$ women, especially for those who have an unplanned birth shortly before their intended labor market entry (shifting their value of $K_1$ from positive to negative). If women with unplanned pregnancies are negatively selected, this will dampen any result suggesting the benefit of a pre-$t_1$ first birth.

[14] In determining her last year of schooling, in some cases I allow gaps in her post-high school education, as long as upon her return she completed more years of schooling at a sufficient pace (see the appendix for more detail). As an example, for a woman who takes a year off after high school, then returns to complete a college degree, her 'graduation year' will be based on the latter (as long as she was not working full-time throughout college). If she then works for a few years before going to graduate school, however, her 'graduation year' will remain the year she completed college.

[15] Using this criterion, of the NLSY79 women observed till age 40, only 106 never enter the labor force. If I limit the analysis to women who work at least 1500 hours in their first year, the results are substantively equivalent. The only clear difference is that for women with some college, the coefficient estimate of $\delta$ (see notation introduced in Section 5) is no longer positive.

[16] Figure 2 uses the cross-section sample of the NLSY79. At one extreme I show the proportion of women who never work (NW), and at the other the proportion who never have kids (NK), the first bar reflecting those who have ever married by their 40s, and the second, those who have not married by this point.

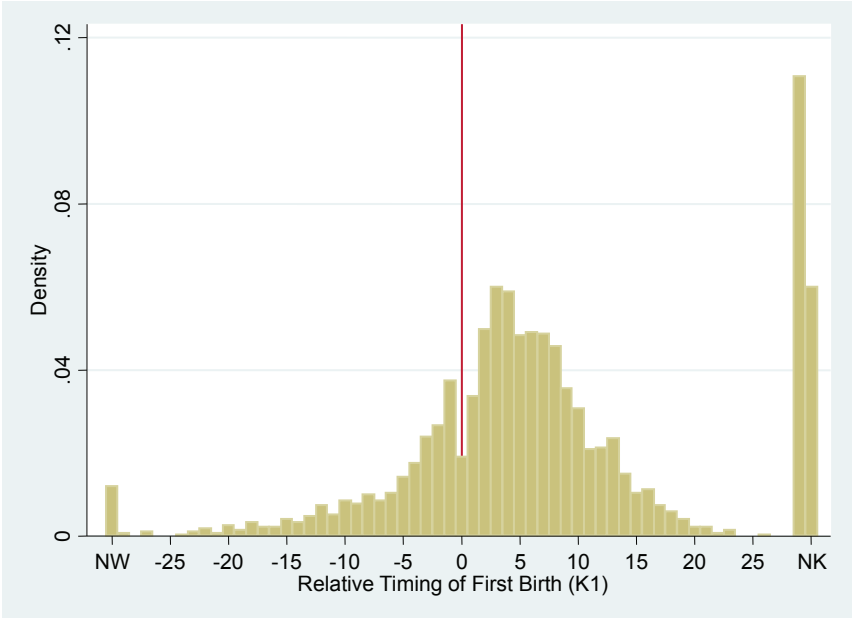


Figure 2: Distribution of Relative Timing of First Birth, $K_1$

Since I group women by their education level at $t_1$, this raises another potential endogeneity issue if education and first-birth timing are jointly determined. For instance, if some women get unexpectedly pregnant while in school, education at $t_1$ may be endogenous to fertility timing for certain $pt_1$ mothers. Education may also be endogenous to *intended* first-birth timing: women who want an early birth may choose less schooling (Blackburn *et al.*, 1993). This suggests that some early mothers will be observed in an education level below their 'potential.' If these women have higher wage growth than the average woman in their observed education group, because they also have earlier first births, this will dampen any positive slope on $K_1$ within these education levels.

In the NLSY79 I find at best weak evidence of either source of endogenous schooling. First, consistent with Rindfuss *et al.* (1980) and Stange (2011), among $pt_1$ women, surprisingly few were still in school when they conceived, and of those, many got more schooling before $t_1$.[17] (These results are also consistent with the finding of at most weak negative effects of a teen birth on education, *e.g.*, Geronimus and Korenman [1992] and Hotz *et al.*

[17]Only 11 percent of high school dropouts, and 20 to 25 percent of women with more education, were in school when they conceived. Among these, for those who were mid-way through high school or college, 40 to 50 percent got more schooling before $t_1$, with the majority completing the given degree by then.

[2005].) I also find little evidence that women who are likely to prefer an earlier first birth intend to get less schooling.[18]

## 4.2 Defining Long-Run Wages, $w_{20}$

My dependent variable is a woman's wage approximately twenty years after labor market entry, $w_{20}$.[19] Specifically, I use any wage observed in a five-year window between a woman's $19^{th}$ and $23^{rd}$ 'career years' (calendar years measured from $t_1$), defining this as "$t_{20}$."

Although using this five-year window limits the bias that may arise from capturing only women with strong labor force attachment, I am still restricted to those women ($i$) who I observe at least to this point, and ($ii$) who work at some point during this five-year stretch.[20] As discussed in Appendix Section D, the first issue has the strongest implications for $pt_1$ mothers, where I systematically lose those who take a long time to enter the labor force after motherhood. Controlling for background characteristics, however, these women show no evidence of systematically different wage patterns, thus their exclusion should have little influence on my results.

The second issue, mothers who do not work during "$t_{20}$," is more pertinent among women with $K_1 > 0$. (Although my sample captures the vast majority of women who remained childless at labor market entry, Appendix Table A-2 shows that this is the most common reason for exclusion.) As I discuss in the appendix, my selection criteria may therefore lead me to somewhat overstate $\theta$ for women with some college at $t_1$, but to understate it for women with a college degree.

[18]To test for a link between taste for early motherhood and intended schooling, I compare highest grade wanted and mother's age at first birth, a strong correlate of own age at motherhood. (Although the NLSY79 collects expected year of first birth, its correlation with expected schooling could reflect the reverse direction of causation.) Among girls observed in high school, I only find a strong positive correlation between these two factors among daughters of high school drop-outs (using father's education to group women into "potential education" levels). I am therefore asking, for instance, whether, among daughters of college graduates, those who may want an earlier first birth also want less schooling. Among daughters of high school or college graduates, I instead find no significant relationship (whether or not I control for mother's education).

[19]See footnote 42 for a discussion of my results if I instead use wages observed approximately 15 years after $t_1$.

[20]The discussion that follows on the implications of missing data incorporates the two sample restrictions discussed in Section 4.3, although the same results broadly hold before their application.

## 4.3 Final Sample Selection

My final sample starts with the set of NLSY79 mothers for whom I can observe $w_{20}$. I then apply two additional sample restrictions: $(i)$ I exclude women with a first birth beyond their $17^{th}$ career year, and $(ii)$ I exclude women who are over 35 when they have their first child.[21] I apply the first restriction, which excludes only 2 percent of mothers, in order to limit any bias incorporated by variation in the pace at which women return to work after motherhood. I apply the second because I find evidence suggesting that among women from the NLSY79 cohort, those with a first birth beyond age 35 are on average negatively selected in terms of some characteristic important in both the marriage and labor markets.[22]

In particular, as shown in Appendix Figure A-1, I find a sharp drop after age 35 in the proportion of women who were married when they conceived their first child or at the birth of that child, and among those married, the older mothers had relatively lower-earning spouses. As I discuss in Appendix Section C, this pattern holds at all education levels, even as the proportion of older mothers rises. Although I only find a corresponding pattern for wages among college graduates, they are likewise the largest group. As I show in Appendix Figure A-2, women who had a first birth after 35 – who have clearly lower wages at $t_{20}$ – already suffered discontinuously lower wage growth long before motherhood.

These restrictions provide a sample of 2,567 mothers, of which 816 (25 percent weighted) have their first child before $t_1$. For this sample, Table 1 reports mean values of $w_{20}$ and starting wages, as well as several background characteristics, grouping women by their first-birth timing and their education level at $t_1$. Throughout this analysis I use four education categories: less than a high school diploma or a GED (Cameron and Heckman, 1993), exactly a high school diploma, some college education, or at least a college degree. Appendix Table A-1 provides a larger set of summary statistics.

[21] Of all mothers observed through age 40 who ever work, 2.2 percent have their first child beyond their 17th year (1.1 percent of high school dropouts, and 2.3 to 2.5 percent in all other education groups). Among all mothers observed through age 40, 4.7 percent have a first child after age 35 (2.0 percent of high school dropouts, 3.2 percent of high school graduates, 5.7 percent of women with some college, and 8.7 percent of college graduates).

[22] This may no longer hold for more recent cohorts, in which first births after age 35 are more common.

Table 1: Summary Statistics by Timing of First Birth

| | **All** | **Fertility Timing ($K_1$)** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $\leq^{-}7$ | $^{-}7$-$^{-}4$ | $^{-}3$-0 | 1-3 | 4-6 | 7-9 | 10+ |
| **Less Than High School Diploma:** | | | | | | | | |
| $\ln(w_{20})$ | 2.25 | 2.13 | 2.26 | 2.32 | 2.15 | 2.36 | 2.30 | 2.22 |
| Starting wage (2000$) | 6.75 | 6.93 | 6.68 | 6.56 | 7.13 | 7.02 | 6.11 | 6.16 |
| $w_{20}$ (2000$) | 10.69 | 9.02 | 10.32 | 11.65 | 9.43 | 12.16 | 10.76 | 10.51 |
| $\ln(w_{20})$-$\ln(w_0)$ | 0.41 | 0.27 | 0.42 | 0.50 | 0.24 | 0.49 | 0.56 | 0.37 |
| Total education at $t_1$ | 10.5 | 10.4 | 10.7 | 10.5 | 10.3 | 10.6 | 10.0 | 10.7 |
| AFQT score (percentile) | 29.9 | 22.8 | 26.2 | 31.3 | 30.6 | 32.7 | 28.6 | 39.1 |
| Highest grade expected | 12.0 | 11.6 | 12.1 | 11.9 | 12.3 | 12.2 | 11.9 | 12.5 |
| **Sample Size:** | **506** | **82** | **64** | **144** | **86** | **67** | **35** | **28** |
| % of all (weighted) | | 13.3 | 10.9 | 27.3 | 18.1 | 15.8 | 8.2 | 6.4 |
| **High School Diploma:** | | | | | | | | |
| $\ln(w_{20})$ | 2.41 | 2.23 | 2.17 | 2.39 | 2.37 | 2.39 | 2.46 | 2.56 |
| Starting wage (2000$) | 7.33 | 6.15 | 6.08 | 7.04 | 7.40 | 7.78 | 7.41 | 7.51 |
| $w_{20}$ (2000$) | 12.38 | 10.28 | 9.49 | 12.32 | 11.78 | 12.01 | 12.97 | 14.24 |
| $\ln(w_{20})$-$\ln(w_0)$ | 0.46 | 0.52 | 0.43 | 0.50 | 0.39 | 0.39 | 0.52 | 0.56 |
| Total education at $t_1$ | 12.0 | 12.0 | 12.0 | 12.0 | 12.0 | 12.0 | 12.0 | 12.0 |
| AFQT score (percentile) | 43.9 | 37.1 | 32.4 | 37.2 | 41.9 | 46.4 | 47.0 | 51.5 |
| Highest grade expected | 13.2 | 13.0 | 12.9 | 13.2 | 13.3 | 13.0 | 13.1 | 13.5 |
| **Sample Size:** | **1,207** | **35** | **74** | **256** | **285** | **218** | **178** | **161** |
| % of all (weighted) | | 2.5 | 4.6 | 16.4 | 22.8 | 20.4 | 16.8 | 16.6 |
| **Some College Education:** | | | | | | | | |
| $\ln(w_{20})$ | 2.59 | 2.67 | 2.61 | 2.70 | 2.54 | 2.49 | 2.63 | 2.66 |
| Starting wage (2000$) | 8.71 | 9.58 | 9.22 | 8.34 | 9.07 | 8.66 | 9.26 | 7.74 |
| $w_{20}$ (2000$) | 15.25 | 15.85 | 14.30 | 17.28 | 14.20 | 13.69 | 16.35 | 16.44 |
| $\ln(w_{20})$-$\ln(w_0)$ | 0.50 | 0.52 | 0.49 | 0.63 | 0.41 | 0.38 | 0.52 | 0.66 |
| Total education at $t_1$ | 13.8 | 13.6 | 13.6 | 13.8 | 13.9 | 13.8 | 13.9 | 13.6 |
| AFQT score (percentile) | 55.2 | 25.9 | 39.1 | 46.8 | 55.9 | 59.9 | 60.2 | 56.5 |
| Highest grade expected | 15.1 | 14.9 | 14.8 | 15.3 | 15.1 | 14.8 | 15.4 | 15.0 |
| **Sample Size:** | **524** | **20** | **38** | **78** | **110** | **108** | **83** | **87** |
| % of all (weighted) | | 2.2 | 5.0 | 11.1 | 22.1 | 23.7 | 18.0 | 17.8 |
| **College Graduate:** | | | | | | | | |
| $\ln(w_{20})$ | 2.86 | 2.76 | 2.53 | 2.53 | 2.78 | 2.77 | 2.95 | 3.07 |
| Starting wage (2000$) | 11.26 | 13.53 | 11.45 | 9.52 | 11.63 | 11.19 | 11.91 | 10.22 |
| $w_{20}$ (2000$) | 20.52 | 21.54 | 14.43 | 14.33 | 18.77 | 18.26 | 22.11 | 25.25 |
| $\ln(w_{20})$-$\ln(w_0)$ | 0.51 | 0.26 | 0.17 | 0.36 | 0.41 | 0.41 | 0.53 | 0.83 |
| Total education at $t_1$ | 16.2 | 17.0 | 16.0 | 16.2 | 16.3 | 16.1 | 16.2 | 16.2 |
| AFQT score (percentile) | 77.2 | 61.1 | 57.1 | 67.4 | 75.2 | 79.3 | 77.4 | 79.0 |
| Highest grade expected | 16.3 | 15.7 | 16.0 | 16.4 | 16.2 | 16.1 | 16.5 | 16.4 |
| **Sample Size:** | **330** | **5** | **6** | **14** | **61** | **98** | **91** | **55** |
| % of all (weighted) | | 1.1 | 0.9 | 4.1 | 16.7 | 30.8 | 29.2 | 17.3 |

**NOTES:** All values reflect weighted means. See the appendix for more details on these weights and how I build these variables. AFQT scores reflect age-adjusted percentiles (measured in 1979), and highest grade expected is reported at approximately age 18. Monetary values are translated into year-2000 dollars using the Consumer Price Index for all urban consumers.

Looking first at women who remain childless at $t_1$ ($K_1 > 0$), for all groups other than high school dropouts, Table 1 echoes the existing literature's finding that long-run wages are increasing in timing of first birth. Excepting high school graduates, we also see no clear link between $K_1$ and either AFQT scores or educational expectations, providing a first suggestion that the correlation between timing and wages is not driven by ability bias.

For $pt_1$ women, Table 1 likewise shows a positive link between wages and later first births among high school dropouts, although not among women with more education. We also see, however, that AFQT scores are likewise increasing with timing, suggesting that the link between timing and wages among $pt_1$ mothers may only reflect this slope on ability.

Lastly, if we compare the outcomes for women with their first birth just before $t_1$ versus just after, we also see initial evidence suggesting that it may be relatively better to have one's first child before entering the labor market. For instance, among high school dropouts and women with some college, we see that women with their first birth within 3 years before $t_1$ have higher long-run wages than women with their first birth in the 3 years following. Furthermore, since $pt_1$ women have lower starting wages, the difference is more striking – and now evident for all but college graduates – when measured in terms of wage growth over this 20-year span.[23]

# 5 Identification Strategy

As a first gauge of the relationship between first-birth timing and wages, I plot the path of wages for the women from the NLSY79, grouping them by their career timing of first birth, $k_1$. To assess whether the positive correlation between timing and wages may be contaminated by bias, I compare this to the hypothetical wage paths in Figure 1, each associated with a potential source of bias discussed in Section 3. Because two of these scenarios also have implications for the relationship between timing and ability, I also plot average AFQT scores by timing of first birth.

[23] I find that the lower starting wages of these $pt_1$ mothers is largely explained by individual characteristics observable at $t_1$ (*e.g.*, AFQT scores and family background).

Building on the model from Section 2.2, I then estimate the following wage equation:

$$l(w_{20}) = l(w_0) + \theta k_1 + \gamma k_1^{pt_1} + \delta pt_1 + X\beta, \tag{5}$$

where $k_1^{pt_1} \equiv K_1 pt_1$ and $\gamma \equiv -g'_{pt_1}$. Notice that I have also rearranged Equation (3) to group the multipliers of $pt_1$ that are unrelated to $K_1$, $\delta \equiv (g^0_{pt_1} - g_2)t_{20}$. Thus in Equation (5), $\theta$ reflects the return to fertility delay for women who enter the labor market without children, $\gamma$ reflects the same for $pt_1$ mothers, and $\delta$ reflects the level difference in long-run wages for women with their first birth just before versus just after labor market entry.

As a second test for bias, I use the following simple ordinary least squares approach to consider how the estimated relationship between timing and wages shifts as I control for an increasing number of individual characteristics. My attention throughout this exercise is on changes in the coefficient estimates of $\gamma$ and $\theta$, as evidence to whether the initial 'raw' coefficients are capturing bias as well as, or instead of, a causal effect.

I begin by estimating Equation (5) including very few additional controls beyond those that speak to the distribution of wages from which a woman's values of $w_0$ and $w_{20}$ were drawn.[24] I treat these OLS-estimated coefficients for $\gamma$ and $\theta$ as my 'raw' measures of the partial correlation between the timing of first birth and long-run wages.

As a first check on the robustness of these coefficients, I then control for a set of observable characteristics, $X_{a_i}$, that reflect ability, career motivation, and other underlying characteristics that may directly influence a woman's wage path.[25] These include:

[24] These include characteristics of where she lived at $t_1$ and $t_{20}$, and the calendar year of the latter. In all specifications I also control for the following: (*i*) years education at $t_1$; (*ii*) total children by $t_{20}$ (so that the coefficients on timing do not capture the indirect effect on wage growth through timing's influence on total fertility), and (*iii*) dummies for women with $K_1 = \{0, 1\}$ (given the pattern in Figure 2 suggesting that these women are relatively select). See the notes to Table 2 for additional controls included in these regressions, and Appendix Section B for more detail on how I define them.

[25] In both this step and the next, I replace missing values with their means because I do not yet want to capture any unobserved heterogeneity reflected in the incidence of missing data. When I later control for family background, $X_{b_i}$, I substitute for these means with missing value indicators.

1. her age-adjusted AFQT score (measured in 1979), as a proxy for ability;
2. her educational and career expectations, measured as expected total schooling (reported at approximately age 18), and expected labor force participation and occupation at age 35 (reported at approximately $t_1$);
3. the length of time between her 'graduation year' and $t_1$ (entered quadratically), to allow for human capital depreciation, or as a reflection of motivation or other characteristics of her underlying labor market 'quality'; and
4. a measure of her self-esteem and her score on the Rotter locus of control scale (measured in 1980 and 1979, respectively), where the latter reflects the degree to which she believes she has control over the events of her life.

If the wages of 'early' and 'late' mothers diverge due to unobserved heterogeneity (as in Scenario #3 described above), the elements of $X_{a_i}$ may be the key factors driving the separation of wages over time. Furthermore, $X_{a_i}$ may influence the optimal timing of a woman's first birth, $k^*_{1i}$, if they enter into her marginal benefit of delay, leading to endogenous sorting as in Scenario #2. Thus if I find that the OLS-estimates of $\gamma$ or $\theta$ fall with the inclusion of $X_{a_i}$, this suggests that the 'raw' estimates were biased upwards by either endogeneity or omitted variables bias.

I next control for a set of characteristics, $X_{\psi_i}$, that either indirectly reflect taste for early motherhood, $\psi_i$, or otherwise influence $k^*_{1i}$ through the marginal cost of delay. These include how many children a woman expects to have, her mother's age at first birth (if $\psi_i$ is inherited), her views on gender–role norms, and her age at first intercourse. If, as in Scenario #3, the positive coefficient on timing arises from a correlation between $\psi_i$ and factors that influence wages, including $X_{\psi_i}$ should drive the OLS-estimates of $\gamma$ and $\theta$ further towards zero by absorbing more of this unobserved heterogeneity.

Benefiting from the richness of the NLSY79, I next control for a broader set of background characteristics, $X_{b_i}$, that may also indirectly build into a woman's optimal timing of first birth by generating variation in either the marginal benefit or marginal cost of delay. These include her family background, such as her race and ethnicity, the religion in which

she was raised, her parents' education, and her mother's labor force participation.[26]

Notice that as I control for an increasing number of factors that influence $k^*_{1i}$, the OLS-estimated coefficients on timing will increasingly be identified off of the difference between a woman's observed and optimal timing of first birth. As a last step, I take this further to control directly for a woman's predicted timing, $k_1^{pred}$.[27] Taking this value as reported at $t_1$ or the year before, for a woman with $K_1 > 0$, this reflects her estimate of $k^*_{1i}$, which will be driven in part by her expected return to delay, $E_{t_1}[\theta_i]$.

Lastly, note that throughout this progression I never control for factors that may reflect part of the *mechanism* of how first-birth timing affects the wage path (Buckles, 2008). For instance, because timing may influence labor supply, I do not control for accumulated work experience, or elements such as tenure or part-time status at $t_{20}$.

Thus to summarize, my focus throughout this process is on how the estimates $\hat{\gamma}$ and $\hat{\theta}$ shift with the inclusion of this series of controls. To the extent that they change, the stage at which this occurs will provide insight into which types of bias are captured in the raw coefficients. By comparison, if the coefficients remain stable, despite the addition of these rich measures of individual characteristics, this suggests that the original estimates largely reflect only a causal effect of the timing of first birth.

[26] At this point I also control for the career year used as "$t_{20}$," which primarily captures the unobserved heterogeneity influencing whether a woman lacked an observed wage for the middle years of the 5-year stretch used for "$t_{20}$." (Since this may itself be endogenous to timing, I rerun these specifications without this control: I find no effect on the estimates of $\theta$, but find that the estimates of $\delta$ rise by 15 to 20 percent for high school dropouts and women with some college.)

[27] In most survey waves, women were asked when they expected to have their next child, which for childless women reflects their predicted timing of first birth. (For women who anticipate an unplanned pregnancy, this may not reflect $k_1^*$.) To the extent that I am now estimating $\hat{\theta}$ using only the unexpected-at-$t_1$ factors that create the difference between predicted and observed timing, one concern is that these unexpected elements may be correlated with a woman's individual return to delay, $\theta_i$. For instance, women may learn new information about $\theta_i$ after entering the labor market, and may adjust their timing accordingly. I find no evidence to this effect. For instance, among college graduates with $K_1 > 3$, wage growth to $t_3$ is unrelated to the difference between $k_1^{pred}$ and $k_1$. Furthermore, plotting average AFQT scores against this difference, the slopes are surprisingly flat – higher-ability women do not systematically delay for longer.

# 6 Results

In Section 6.1, I first discuss my results on whether the observed correlation between first-birth timing and wages is contaminated by bias, comparing my findings to those in the literature. In Section 6.2, I then discuss the economic implications of my findings. I also test my earlier predictions on the effect of instead measuring timing in terms of age, and end with the question of whether the return varies over time or across women.

## 6.1 Testing for Bias

As initial evidence for the possible causality of first-birth timing on women's long-run wages, Figure 3 plots the experience profile of wages for the women in the NLSY79, grouped by their career timing of first birth, $k_1$.[28] The main pattern evident is that women's wages move together before motherhood, and diverge sharply thereafter.

For instance, consider women with $k_1 = 3 - 10$. Up till their third year in the labor market – before any have children – their wages move together (although women with $k_1 = \{3, 4\}$ are at a slightly lower level). At that point, the first group reaches motherhood, and their growth rate stalls. By comparison, the remaining two groups continue to move together, diverging only after the next group begins to have children.

Thus overall, we see that the raw wage path in Figure 3 looks strikingly similar to the wage pattern for Scenario #1, where the difference in long-run wages between early and late mothers reflects a causal effect of the timing of first birth. In particular, we see that wages move together until each group has their first child, at which point the wage path for mothers

[28] This figure includes all mothers in the NLSY79 who are observed at least 20 years beyond $t_1$. (The pattern is effectively identical, but noisier, if I instead use the cross-section, or only my analysis sample.) If I plot this figure in more narrowly-defined education groups (*e.g.*, only high school graduates, only women with at least some college, or all but college graduates), in each I find the same clear pattern of a kink in the wage path after motherhood. Although each point in Figure 3 captures data only for women with observed wages in the given year, changes in composition do not drive this pattern. (For instance, if I instead plot average AFQT scores by timing of first birth, including in each year only those women with an observed wage, there is no discontinuous drop at first birth.) For the sake of comparison, the wage path of childless women starts just below that for women with $k_1 = 5 - 10$, and continues to fall behind over the next five to ten years, catching up only after each group of mothers have their first child.

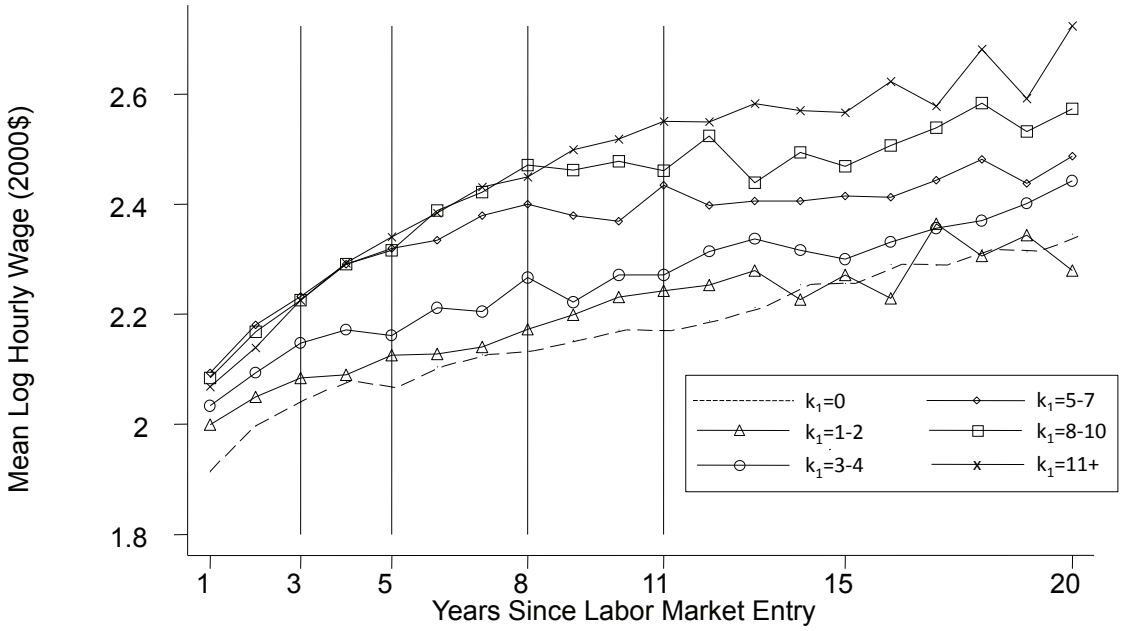


Figure 3: Wage Path by Timing of First Birth

kinks.[29] We do not see that wages diverge directly upon labor market entry, as anticipated by both Scenario #2 (endogenous timing) and Scenario #3 (unobserved heterogeneity), nor do we see that wages diverge *before* first birth, as in Scenario #4 (reverse causality).[30]

Furthermore, the figure looks surprisingly similar to the piece-wise linear wage path depicted in Equation (1). We see that the rate of wage growth before first birth, $g_1$, is very similar across timing groups, and tracing out an average wage path after first birth, $g_2$ looks surprisingly similar as well. This in turn fits the description of Scenario #1, in which the return to delay is invariant because $g_1$ and $g_2$ are equal across the population, and variation in observed timing of first birth arises only from variation in taste for early motherhood.

[29]By contrast, when lining up wages instead by age, Miller (2011) finds that wages diverge before first birth, although Wilde *et al.* (2010) do not.

[30]Regression results likewise show no evidence of reverse causality. For instance, if I regress a woman's wage level at $t_5$ on her *subsequent* first-birth timing (among women with $k_1 > 5$), I find no 'effect' of $k_1$, even though in every education group roughly 45 percent of these women have their first birth within the next two years.

In Figure 4 I next plot average AFQT scores by relative timing of first birth, $K_1$, separating women by their education level at $t_1$.[31] As described in Scenarios #2 and #3, positive ability bias may drive the link between timing and wages if high-ability women systematically choose later first births.

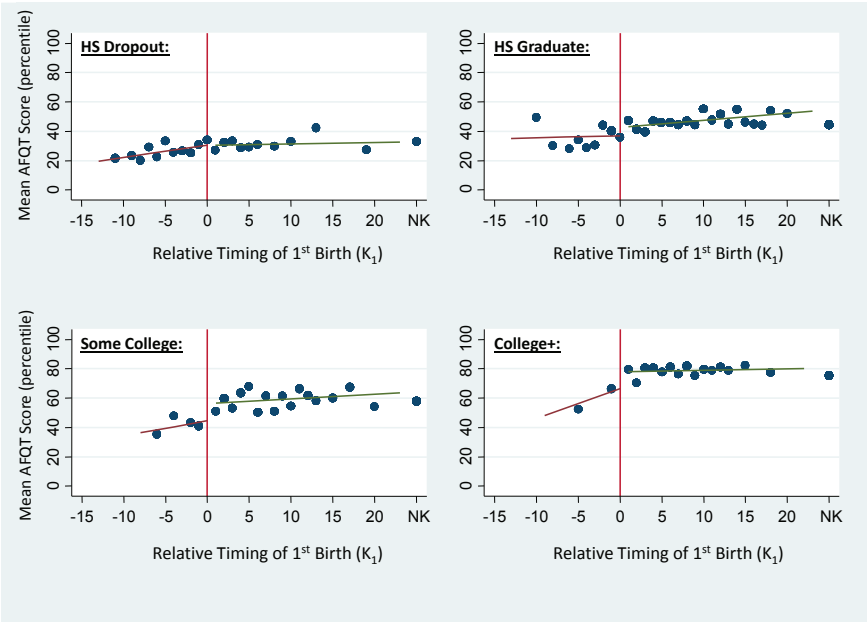


Figure 4: Mean AFQT Scores by Timing of First Birth

As evident in Figure 4, however, among women with their first birth after $t_1$, timing and AFQT scores are almost completely unrelated (the slope is only statistically positive for high school graduates). By contrast, as we also saw in Table 1, $pt_1$ women who enter the labor market more promptly after first birth have generally higher AFQT scores.[32] Thus if I find a positive correlation between timing and wages among $pt_1$ mothers, Figure 4 suggests that this may only reflect unobserved heterogeneity bias. But for women with a first birth after $t_1$, Figure 4 largely undermines the possibility that ability bias is driving the link between timing and wages.

[31] This figure reflects data for women from the NLSY79 cross-section sample who are observed at least 20 years beyond $t_1$ (including women with no kids, 'NK'). The figures are very similar using my analysis sample.

[32] The slopes are significantly positive for high school dropouts and women with some college.

As the second test for bias, Table 2 reports the results of the series of regressions described in Section 5. Grouping women by their education level at $t_1$, in each panel I report the OLS-estimates of $\theta$, $\gamma$, and $\delta$ (the coefficients on $k_1$, $k_1^{pt_1}$, and $pt_1$, respectively), as well as the coefficients on several of the key covariates added at the various stages of the identification process.

The first column of Table 2 reports the 'raw' coefficient estimates of $\theta$ and $\gamma$ – the partial correlations between fertility timing and $l(w_{20})$ before controlling for individual characteristics. For $\theta$, we see no evidence of a link between wages and timing among high school dropouts, but a clear positive correlation at higher education levels. For $\gamma$ we see the reverse: a positive slope on $K_1$ for high school dropouts, and no clear link between timing and wages for other education groups.[33]

In the second column we see the influence on these estimates of controlling for $X_{a_i}$ (*e.g.*, AFQT scores and highest grade expected), factors that directly influence a woman's wage, and may likewise enter into her marginal benefit of delay. For $\theta$, we see relatively little change. The coefficient for high school graduates drops by approximately 25 percent, driven by the correlation between AFQT scores and $K_1$ evident in Figure 4, but for all other education groups, $\hat{\theta}$ is almost completely unchanged.

[33]Throughout this discussion I ignore the coefficients estimated on the small number of $pt_1$ college graduates. I will also largely disregard the coefficient estimates of $\gamma$ for women with some college, which hover near statistical significance throughout these specifications, because they are largely driven by two outliers with very early first births.

Table 2: Estimates of the Effect of First-Birth Timing on Wages

| | 'Raw' | $+ X_{a_i}$ | $+ X_{\psi_i}$ | $+ X_{b_i}$ | $+ k_1^{pred}$ |
|---|---|---|---|---|---|
| Less than High School Diploma: | | | | | |
| $k_1$ ($\theta$) | 0.007 | 0.006 | 0.005 | 0.001 | 0.000 |
| | (0.011) | (0.012) | (0.012) | (0.011) | (0.011) |
| $k_1^{pt_1}$ ($\gamma$) | 0.026* | -0.006 | -0.004 | -0.007 | -0.004 |
| | (0.014) | (0.022) | (0.022) | (0.022) | (0.023) |
| $pt_1$ ($\delta$) | 0.250** | 0.230** | 0.226** | 0.134 | 0.391*** |
| | (0.103) | (0.112) | (0.113) | (0.115) | (0.148) |
| Starting wage ($w_0$) | 0.118* | | | | |
| | (0.069) | | | | |
| AFQT score (x $10^1$, in 1979) | | 0.022* | | | |
| | | (0.011) | | | |
| Expected education (at 18) | | 0.022 | | | |
| | | (0.013) | | | |
| Self-esteem (in 1980) | | 0.076** | | | |
| | | (0.033) | | | |
| Total kids expected (at $t_1$) | | | 0.016 | | |
| | | | (0.023) | | |
| $k_1^{pred}$ (at $t_1$) | | | | | 0.037** |
| | | | | | (0.017) |
| $R^2$ | **0.22** | **0.28** | **0.29** | **0.39** | **0.40** |
| High School Diploma: | | | | | |
| $k_1$ ($\theta$) | 0.019*** | 0.014*** | 0.015*** | 0.014*** | 0.014*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| $k_1^{pt_1}$ ($\gamma$) | 0.010 | -0.002 | 0.002 | 0.000 | 0.001 |
| | (0.018) | (0.017) | (0.017) | (0.017) | (0.017) |
| $pt_1$ ($\delta$) | 0.117* | 0.172*** | 0.161*** | 0.154** | 0.132 |
| | (0.066) | (0.066) | (0.067) | (0.066) | (0.081) |
| Starting wage ($w_0$) | 0.214*** | | | | |
| | (0.057) | | | | |
| AFQT score (x $10^1$, in 1979) | | 0.032*** | | | |
| | | (0.006) | | | |
| Expected education (at 18) | | 0.048*** | | | |
| | | (0.009) | | | |
| Self-esteem (in 1980) | | 0.029* | | | |
| | | (0.016) | | | |
| Total kids expected (at $t_1$) | | | 0.024* | | |
| | | | (0.014) | | |
| $k_1^{pred}$ (at $t_1$) | | | | | -0.003 |
| | | | | | (0.009) |
| $R^2$ | **0.17** | **0.25** | **0.25** | **0.31** | **0.31** |

**NOTES:** This table lists the coefficients (and robust standard errors) from a weighted regression of Equation (5), controlling for an increasing number of individual characteristics (see footnote 24 for a listing of the controls included in all specifications). 'Raw' estimates exclude individual characteristics; '$+X_{a_i}$' adds controls for AFQT and Rotter scores, self-esteem, and expected total education and (*continued on next page*)

Table 2 – Continued

| | 'Raw' | $+ X_{a_i}$ | $+ X_{\psi_i}$ | $+ X_{b_i}$ | $+ k_1^{pred}$ |
|---|---|---|---|---|---|
| Some College: | | | | | |
| $k_1$ ($\theta$) | 0.018** | 0.019** | 0.018** | 0.018** | 0.017** |
| | (0.009) | (0.009) | (0.009) | (0.008) | (0.008) |
| $k_1^{pt_1}$ ($\gamma$) | 0.030 | 0.037 | 0.044* | 0.036 | 0.035 |
| | (0.019) | (0.024) | (0.024) | (0.023) | (0.023) |
| $pt_1$ ($\delta$) | 0.319*** | 0.341*** | 0.325** | 0.259** | 0.289* |
| | (0.122) | (0.132) | (0.135) | (0.128) | (0.149) |
| Starting wage ($w_0$) | 0.457*** | | | | |
| | (0.074) | | | | |
| AFQT score (x $10^1$, in 1979) | | 0.059*** | | | |
| | | (0.012) | | | |
| Expected education (at 18) | | 0.021 | | | |
| | | (0.017) | | | |
| Self-esteem (in 1980) | | 0.058* | | | |
| | | (0.032) | | | |
| Total kids expected (at $t_1$) | | | 0.038* | | |
| | | | (0.023) | | |
| $k_1^{pred}$ (at $t_1$) | | | | | 0.014 |
| | | | | | (0.013) |
| $R^2$ | **0.21** | **0.30** | **0.31** | **0.48** | **0.49** |
| College Degree or More: | | | | | |
| $k_1$ ($\theta$) | 0.034** | 0.033** | 0.033** | 0.030** | 0.029** |
| | (0.013) | (0.013) | (0.013) | (0.014) | (0.014) |
| $k_1^{pt_1}$ ($\gamma$) | 0.010 | -0.029 | -0.028 | 0.004 | 0.009 |
| | (0.069) | (0.065) | (0.064) | (0.066) | (0.067) |
| $pt_1$ ($\delta$) | 0.125 | 0.050 | 0.054 | 0.104 | 0.101 |
| | (0.227) | (0.229) | (0.230) | (0.251) | (0.263) |
| Starting wage ($w_0$) | 0.390*** | | | | |
| | (0.086) | | | | |
| AFQT score (x $10^1$, in 1979) | | 0.014 | | | |
| | | (0.018) | | | |
| Expected education (at 18) | | -0.020 | | | |
| | | (0.035) | | | |
| Self-esteem (in 1980) | | -0.013 | | | |
| | | (0.043) | | | |
| Total kids expected (at $t_1$) | | | -0.050 | | |
| | | | (0.046) | | |
| $k_1^{pred}$ (at $t_1$) | | | | | 0.014 |
| | | | | | (0.023) |
| $R^2$ | **0.26** | **0.29** | **0.30** | **0.44** | **0.45** |

work status at age 35; '$+X_{\psi_i}$' adds controls for total kids expected, gender-role attitudes, mother's age at first birth, and own age at first sexual intercourse; '$+X_{b_i}$' adds controls for family background (*e.g.*, race, religion, parents' education, total siblings, and characteristics of the household at age 14); '$+k_1^{pred}$' controls for a woman's prediction of her own timing of first birth, made at approximately $t_1$. Significance levels are indicated by * (at the 10% level), ** (at 5%), and *** (at 1%).

For $\gamma$, we instead see that controlling for $X_{a_i}$ eliminates the positive correlation between timing and wages observed among high school dropouts. This is driven in part by AFQT scores, as anticipated by Figure 4. Yet for $pt_1$ mothers, the more important element of $X_{a_i}$ generating the raw correlation between timing and wages is the length of the gap between the point when a woman left school and when she entered the labor force. For both high school dropouts and graduates, there is a negative relationship between years off and long-run wages (which remains after controlling for all individual characteristics), suggesting that at these education levels, $pt_1$ mothers who start working promptly are "higher wage" types.[34]

In the next column I control for $X_{\psi_i}$, factors that may be correlated with taste for early motherhood.[35] If, as in Scenario #3, the link between timing and wages arises from a correlation between $\psi$ and factors that influence wage growth, controlling for $X_{\psi_i}$ should drive the estimates $\hat{\theta}$ and $\hat{\gamma}$ towards zero. Instead we see that their inclusion has little effect, in terms of either the coefficient estimates or the fit.[36] When I next control for background characteristics, $X_{b_i}$, the $R^2$ rises throughout, but the estimates $\hat{\theta}$ and $\hat{\gamma}$ still remain unchanged.[37]

Lastly I control for $k_1^{pred}$, a woman's prediction of $k_1$ based on her information set at $t_1$ (the term is undefined for $pt_1$ women). If, as in Scenario #2, variation in $\theta_i$ introduces positive bias because women facing a higher return to delay choose later first births, then controlling for $k_1^{pred}$ should absorb this, driving the estimates $\hat{\theta}$ towards zero. Instead we see that the coefficients are very stable, even though $k_1^{pred}$ is a good predictor of observed timing.[38] Furthermore, since $k_1^{pred}$ is unrelated to $w_{20}$ – the same holds even if I include

[34]This negative relationship is largely linear, rather than the quadratic pattern we would expect if this reflected the depreciation of human capital over time. Controlling for years off also drives the change in $\hat{\gamma}$ and $\hat{\delta}$ for high school graduates.

[35]All of the elements of $X_{\psi_i}$ are significantly correlated with timing for at least one education group, although among non-$pt_1$ mothers, the link is more common with *predicted* than observed timing. (The incidence of a significant link with observed timing is highest for college graduates, where I also find the highest correlation between predicted and observed timing.)

[36]If I instead add $X_{\psi_i}$ before $X_{a_i}$, the coefficients are again completely unaffected by their inclusion.

[37]For both high school dropouts and women with some college, the drop in $\hat{\delta}$ is driven by controlling for a woman's family characteristics at age 14 and the religion in which she was raised.

[38]Across these four education groups, $k_1$ increases by 0.3, 0.5, 0.6, and 0.7 years for each one-year increase in $k_1^{pred}$, respectively. (Including $k_1^{pred}$ has no effect on the coefficients $\hat{\theta}$ even if I add this before any other

$k_1^{pred}$ before all other controls – the *type* of women who expect to have kids later do not have higher wages.

Overall, in line with the evidence in Figure 3, the results in Table 2 show little indication of bias captured in the correlation between first-birth timing and long-run wages for women who remain childless at $t_1$.[39] Among these mothers, I only find evidence of bias for the high school graduates, where the initial coefficient is somewhat inflated by positive ability bias.[40] In contrast, among $pt_1$ mothers I find clear evidence of positive unobserved heterogeneity bias among high school dropouts, but not among women with more education.

By comparison, when measuring first-birth timing instead in terms of age, the existing literature has consistently found that the raw correlation between timing and wages is biased upwards. For instance, relative to OLS estimates, using fixed effects, Amuedo-Dorantes and Kimmel (2005) find a smaller estimated return to delaying first birth beyond age 30, and using IV, Miller (2011) finds a smaller return to delay among women in their 20s and 30s. Geronimus and Korenman (1992) and Hotz *et al.* (2005) likewise find a smaller negative impact of a teen birth using fixed effects and IV approaches, respectively, but this may mirror the result I find for $pt_1$ women, since teen mothers will largely be high school dropouts.

The comparison of my findings to these results for women with a first birth beyond their teens, suggests that the positive bias reported in the existing literature is largely a byproduct of the mismeasurement of the timing of first birth. I therefore conclude that when timing is correctly measured in terms of a woman's entry into the labor force, the raw correlations between timing and wages largely reflect the causal effect of the timing of a woman's first birth.

individual-specific controls.)

[39] If I run the analysis without the survey weights, the pattern of how the estimates $\hat{\theta}$ change with the inclusion of these controls is very similar; for $\hat{\gamma}$ I instead find that the initial positive coefficient for high school dropouts is insignificantly different from zero.

[40] The results when adding $k_1^{pred}$ suggest that this ability bias is more likely the result of unobserved heterogeneity than endogeneity. (It also seems improbable that variation in $\theta_i$ would lead to endogenous sorting among high school graduates, but not among more educated women.)

## 6.2 The Returns to Delayed First Birth

Given this conclusion, consider now the economic implications of the results in Table 2. First, we see no effect of first-birth timing for women who have their first child before they start working: $\hat{\gamma} = 0$, thus $g'_{pt_1} = 0$. (For $pt_1$ mothers, the relevant coefficients are those reported in the penultimate column of Table 2.[41]) I therefore conclude that the appropriate measure for gauging the effect of first-birth timing on women's wages is the 'career timing' of first birth, $k_1$ – the point in a woman's career in which children are first present.

Next we see that among women who remain childless at labor market entry, for all but high school dropouts, there is a clear positive return to additional fertility delay.[42] For women with a high school diploma or some college at $t_1$, one year of delay leads to 1.4 to 1.7 percent higher wages 20 years after labor market entry, corresponding to roughly 3 percent of total wage growth over this 20-year stretch. By comparison, the return to a one-year delay of first birth for college graduates is 2.9 percent, or a much larger 6 percent of total wage growth by this point.[43]

Yet the positive estimates for $\delta$ suggest that in a range around $t_1$, it may be strictly better to have one's first birth before entering the labor force. (Given my conclusion that $g'_{pt_1} = 0$, when I re-estimate Equation (5) excluding $k_1^{pt_1}$, $\hat{\delta}$ is almost identically 0.15 for all education levels below college graduates.[44])

[41] Since $k_1^{pred}$ is undefined for $pt_1$ women, when I include it in the regressions, to the extent that the intercept of the slope between $k_1^{pred}$ and $l(w_{20})$ is non-zero, this will be loaded onto the coefficient for $pt_1$.

[42] Another possibility is that the effect of children on a woman's wage path is the same, regardless of the timing of her first, but that the cost evolves over time. In that case, when I observe the sample at $t_{20}$, I am capturing these women at different stages of this process. To consider this possibility, I compare the estimates of $\theta$ when measured at $t_{15}$ and $t_{20}$ (for women with $K_1 \leq 12$, approximately 85 percent of each education group). Although I find that $\hat{\theta}$ falls by 20 percent between $t_{15}$ and $t_{20}$ for women with some college, $\hat{\theta}$ is completely unchanged for both high school and college graduates.

[43] As noted in Section 4.2, my sample selection criteria may lead me to overstate the return for women with some college, $\hat{\theta}_{SC}$, but to understate it for college graduates. For the former, however, I also find evidence that measurement error in $K_1$, generated by error in $t_1$, may be simultaneously driving $\hat{\theta}_{SC}$ towards zero. I find no evidence of measurement error at other education levels.

[44] For high school dropouts and women with some college, this coefficient is statistically significant at the 15-percent level (for high school graduates the coefficient is unchanged from that reported in Table 2). Per the comment in footnote 26, when I exclude $t_{20}$ from the specification, $\hat{\delta}$ rises to 0.175 for both groups, and is now statistically significant at the 10-percent level.

More precisely, given its definition, $\delta \equiv (g^0_{pt_1} - g_2)t_{20}$, $\hat{\delta} > 0$ suggests that $g^0_{pt_1} > g_2$. Namely, $pt_1$ mothers experience a higher average wage growth rate than the post-motherhood growth rate observed for non-$pt_1$ mothers. Furthermore, one can define the "cross-over" point, $\tau^x$, as the point beyond which a woman must delay her first birth, in order to experience greater wage growth by $t_{20}$ than if she had her first child before $t_1$: $g_1\tau^x + g_2(t_{20} - \tau^x) = g^0_{pt_1}t_{20}$. Rearranging and substituting terms, $\tau^x = \delta/\theta$.

Based on my coefficient estimates, the break-even points for high school graduates and women with some college are strikingly far along: $\hat{\tau}^x_{HS} = 11.0$ and $\hat{\tau}^x_{SC} = 8.4$. Since among non-$pt_1$ mothers, 82 percent of the former, and 67 percent of the latter, have their first child before these cross-over points, my results imply that for both groups, the vast majority of women would have higher long-run wages if they had their first child before entering the labor force rather than after.[45] Furthermore, since $\hat{\theta} = 0$ for high school dropouts, the same conclusion holds for all but college graduates.[46]

It is unclear why women with a first birth before $t_1$ fare so much better than those with later children. This result is not driven by the subset of mothers who complete their childbearing before entering the labor force.[47] Thus it may be that the effect of children on women's wages largely arises from the career disruption of the shift into motherhood, rather than the presence of children themselves. Alternatively, the early arrival of a child may induce women to "buckle down" at an earlier stage, consistent with the labor supply results in Hotz *et al.* (2005).

It is uncertain, however, whether an earlier first birth will provide better outcomes on other dimensions, for either the mother or the child, especially if this would require a

[45] These results do not suggest that $g^0_{pt_1} > g_1$, which will hold if $\frac{\delta}{t_{20}} > \theta$. For instance, $\hat{\delta}_{HS}/t_{20} = 0.007 < \hat{\theta}_{HS}$, with $\bar{t}_{20} = 21.5$.

[46] Given the positive point estimate for $\delta$ among college graduates, there may be such a cross-over point even among the most educated women. By these estimates, however, the cross-over point is fairly early ($\tau^x_C = 3.6$), thus the majority of college graduates still benefit from a first birth beyond $t_1$.

[47] Among $pt_1$ mothers, 40 percent of high school dropouts, and 30 percent of high school graduates and women with some college, have completed their fertility by $t_1$. Controlling for this lowers $\hat{\delta}_{HS}$ by 15 percent, but has no effect for the other two education groups.

pre-marital birth.[48] I leave to future work this question of the mechanism of the benefit of a pre-$t_1$ first birth, as well as the mechanism for the positive return to delay among women who instead enter the labor force before having children.

Now consider the impact on these results if I instead regress $l(w_{20})$ on age at first birth, consistent with the existing literature. Using this alternate specification, I find a 'return' of 0.9 percent for women with a high school diploma or some college, and 2.1 percent for college graduates. (For high school dropouts, I again find no link between timing and wages.) Thus, consistent with my prediction in Section 2.2, when fertility timing is incorrectly measured in terms of age, the estimated return to one year of delay understates $\theta$ by as much as 30 to 50 percent.[49] The difference is smallest for college graduates, because of the lower proportion of $pt_1$ mothers. More importantly, however, these estimates based on age completely obscure the benefit of an *earlier* first birth.

As a last point, I consider whether $\theta$ varies over time or across women. In Section 2.1 I suggested that the return to fertility delay may arise from the postponement of the interruption of motherhood further into one's career, when the underlying wage path has started to plateau. This suggests that the return should fall with time, as the gradient of the wage profile becomes less steep. Furthermore, since the wage path is steeper among the more educated (Murphy and Welch 1992), the return to delay may increase with schooling, or

[48] Comparing women with their first birth within 3 years to either side of $t_1$, the average age at birth for the pre-$t_1$ mothers is 17.9 for high school dropouts, 18.8 for high school graduates, and 20.3 for women with some college, compared to 20.1, 21.0, and 23.0 for the non-$pt_1$ mothers, respectively. The proportion married at the conception of their first child is starkly lower among the $pt_1$ mothers (17 to 26 percent versus 45 to 60 percent), although the difference in the proportion married at the birth of that child is smaller (50 percent versus 60 to 80 percent). Yet between 55 and 75 percent of the non-$pt_1$ mothers were married by the year they started working or the year following, suggesting many could have had a pre-$t_1$ marital first birth.

[49] Note that the estimates of the return to delay in Chandler *et al.* (1994) and Miller (2011), 2.1 and 3.0 percent, respectively, are much higher than $\hat{\theta}$ for my pooled sample (1.5 percent). (Only these papers report the return to a single year of delay.) This is driven exclusively by their sample selection criteria. Building a sample consistent with Chandler *et al.*, I approximately match their result measuring in terms of age (1.9 percent), but measured in terms of 'career timing,' $\hat{\theta}$ is instead 3.2 percent, almost 70 percent higher. I can likewise reproduce the result for Miller (which uses the NLSY79), and for this sample, find that $\hat{\theta}$ is 3.3 percent when measured in terms of career timing, only slightly higher than her coefficient on age. Consistent with my discussion in Section 2.2, this is because only a very small proportion of Miller's sample has a first birth before $t_1$.

with any other factor that leads to a steeper wage profile.[50]

I first test for a declining return to delay by including the quadratic of $k_1$ in Equation (5). I find no evidence of a decreasing return, and a plot of the coefficients on dummies for each value of $k_1$ shows a linear pattern, rather than evidence that the level difference in $l(w_{20})$ tails off as $k_1$ increases. When we consider that the vast majority of women have their first birth during their first decade in the labor force, a range in which the standard wage profile remains steep, this result may seem less surprising.

Testing for whether $\theta$ varies with education or ability, I find mild evidence of variation across the full population, but no evidence of additional variation within education levels. In particular, consistent with my results in Table 2, in the pooled sample I find that $\theta$ is increasing in years of schooling at $t_1$; the return rises by 0.3 percentage points for each year of education (significant at the 10-percent level). By comparison, and potentially surprisingly, I do not find that the return to delay is increasing in AFQT scores. Furthermore, among women with at least a high school diploma, within each education level, I find no evidence of additional variation in the return by either education or ability.

This conclusion in turn helps answer a puzzle posed by my results: since $\theta$ is relatively large, why do we see no evidence that women respond by shifting their timing of first birth? The answer for why I find no such endogenous sorting is that there is surprisingly little variation in the return to delay across women.

[50] In footnote 1 I also raised the possibility that the return to delay varies with the *reason* for that delay. To the extent that I can test this, I find only limited evidence of variation in $\theta$. In particular, I can (imperfectly) isolate three of the possible reasons listed there: (a) "I wasn't married yet," (c) "it took me longer to get pregnant than I expected," and (f) "I got pregnant unexpectedly." For (a), I compare whether the return to delay is systematically different for women who conceive immediately after marriage, to approximate the population with pent-up demand for a (marital) first birth. Likewise, for (c), I use as the comparison women who were actively trying to get pregnant at least one year before they conceived, and for (f), I use the sample of women who report that they did not want to get pregnant when they conceived their first child. For both (c) and (f), I find no evidence that the returns are different (and where sample sizes allow sufficient precision, the estimated returns are very similar). Only for (a) do I find any evidence that the return varies. Among high school graduates, the return is significantly smaller (and may be zero) for women who may be delaying only because they remain unmarried.

# 7 Conclusion

The first conclusion of this analysis is that the existing literature has mis-estimated the return to fertility delay by its focus on a woman's age at first birth. When one considers the mechanisms by which fertility timing may influence a woman's wage path, it is clear that the critical factor is her experience level at first birth, not her age. Furthermore, this leads to the insight that the effect may vary for women with a first birth before versus after they enter the labor force. In the analysis above, I show that the relevant measure is the career timing of a woman's first birth, the point in her career when children are first present.

One result that emerges from this conclusion is that when timing is measured correctly, there is little bias captured in the raw correlation between the timing of a woman's first birth and her long-run wages. This result is clearly evident in Figure 3, where women's wages move together before motherhood, and stall directly afterwards.

The more striking results that emerge from this re-definition are the updated estimates of the magnitude, and sign, of the return to fertility delay. First, because there is no link between $K_1$ and wages among $pt_1$ women, the existing literature's mis-specification of timing has led to an underestimate of the benefit of fertility delay for women who remain childless when they enter the labor force. More importantly, the literature's universally-reported positive estimates of the return to delay have concealed the benefit of an *earlier* first birth, before labor market entry.

# References


[1] Amuedo-Dorantes, Catalina and Jean Kimmel, "The Motherhood Wage Gap for Women in the United States: The Importance of College and Fertility Delay," *Review of Economics of the Household*, 3 (2005), 17-48.

[2] Blackburn, McKinley L., David E. Bloom, and David Neumark, "Fertility Timing, Wages, and Human Capital," *Journal of Population Economics*, 6 (1993), 1-30.

[3] Bloom, David E., "Fertility Timing, Labor Supply Disruptions, and the Wage Profiles of American Women," In *Proceedings of the Social Statistics Section*, (1986), 49-63.

[4] Buckles, Kasey, "Understanding the Returns to Delayed Childbearing for Working Women," *American Economic Review, Papers and Proceedings*, 98 (2008), 403-07.

[5] Cameron, Stepehn V., and James J. Heckman, "The Non-Equivalence of High School Equivalents," *Journal of Labor Economics*, 11 (1993), 1-47.

[6] Chandler, Timothy D., Yoshinori Kamo, and James D. Werbel, "Do Delays in Marriage and Childbirth Affect Earnings?" *Social Science Quarterly*, 75 (1994), 838-853.

[7] Geronimus, Arline T., and Sanders Korenman, "The Socioeconomic Consequences of Teen Childbearing Reconsidered," *Quarterly Journal of Economics*, 107 (1992), 1187-1214.

[8] Hamermesh, Daniel S., and Jeff E. Biddle, "Beauty and the Labor Market," *The American Economic Review*, 84 (1994), 1174-1194.

[9] Heckman, James J. and James R. Walker, "The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data," *Econometrica*, 58 (1990), 1411-1441.

[10] Hotz, Joseph V., Susan Williams McElroy, and Seth G. Sanders, "Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment," *The Journal of Human Resources*, 40 (2005), 683-715.

[11] Light, Audrey, and Manuelita Ureta, "Early-Career Work Experience and Gender Wage Differentials," *Journal of Labor Economics*, 13 (1995), 121-154.

[12] Miller, Amalia R., "The Effects of Motherhood Timing on Career Path," *Journal of Population Economics*, 24 (2011), 1071-1100.

[13] Murphy, Kevin M., and Finis Welch, "Empirical Age-Earnings Profiles," *Journal of Labor Economics*, 8 (1990), 202-229.

[14] Murphy, Kevin M., and Finis Welch, "The Structure of Wages," *The Quarterly Journal of Economics*, 107 (1992), 285-326.

[15] Mullin, Charles H., and Ping Wang, "The Timing of Childbearing Among Heterogenous Women in Dynamic General Equilibrium," NBER Working Paper No. 9231, 2002.

[16] Perry, Cynthia D., "How do Female Earnings Affect Fertility Decisions?" Working Paper, 2005.

[17] Rindfuss, Ronald R., Larry Bumpass, and Craig St. John, "Education and Fertility: Implications for the Roles Women Occupy," *American Sociological Review*, 45 (1980), 431-447.

[18] Schmidt, Lucie, "Murphy Brown Revisited: Human Capital, Search, and Nonmarital Childbearing among Educated Women," Working Paper, 2007.

[19] Stange, Kevin, "A Longitudinal Analysis of the Relationship Between Fertility Timing and Schooling," *Demography*, 48 (2011), 931-56.

[20] Taniguchi, Hiromi, "The Timing of Childbearing and Women's Wages," *Journal of Marriage and the Family*, 61 (1999), 1008-1019.

[21] Topel, Robert H. and Michael P. Ward, "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107 (1992), 439-479.

[22] Wilde, Elizabeth Ty, Lily Batchelder, and David T. Ellwood, "The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Differing Human Capital," NBER Working Paper No. 16582, 2010.

# Appendix

## A Sample

This analysis uses the sample of women from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). The survey began in 1979 with 6,283 14- to 22-year old women, who by 2008 had reached the ages of 44 to 51. The largest groups lost across time are the military and supplemental white/poor samples, which were dropped from the NLSY79 in 1984 and 1991, respectively. Because the earliest year used for "$t_{20}$" in my analysis is 1994, this means that these two samples are largely excluded. Exclusive of these two groups, we observe 89 percent of the remaining sample through at least age 40, or approximately the end of the childbearing years. Of these, 83 percent had at least one child by their last year observed, when the median woman was 47.

My sample builds from the NLSY79 cross-sectional sample, plus the supplemental minority samples. Because of the inclusion of these oversamples, throughout the analysis I apply sampling weights that were built specifically for my final sample by the Center for Human Resource Research at the Ohio State University.[1]

## B Variable Definitions

To define $t_1$, the year of labor market entry, I first establish a woman's 'graduation year,' the year in which she completed continuous schooling. (Under certain circumstances I allow gaps in a woman's education, as long as upon her return she completed more years of schooling at a sufficient pace.[2]) Starting from June of that year, I then search for the first twelve-month

[1] When I calculate statistics for a population broader than my sample, I use the 1979 sampling weights. In my final sample, by number, women from the supplemental sample make up 41 percent of high school dropouts, 35 percent of high school graduates, 45 percent of women with some college, and 21 percent of college graduates. Applying the appropriate weights, these women instead represent 17, 15, 21, and 7 percent of each education group, respectively.

[2] I allow a 1-year gap if, after returning to school, her reported education level increased. (Many respondents report school attendance without their completed education level increasing thereafter.) I also allow up to a 2-year gap if upon return she stayed in school for at least 2 years, completed at least 2 years of schooling in that timeframe, and completed it at a rate of at least 0.65 school years per calendar year (*e.g.*, finishing 2 school years in 3 years but not 4). Lastly, I allow up to a 4-year gap if during that time she never worked more than 1000 hours per year, and when she returned to school she completed at least 1 year of

period in which she worked at least 1000 hours.[3] I define that year — the second calendar year in that 12-month stretch — as $t_1$, her first year in the labor market.

I define a woman's 20th-year wage, $w_{20}$, using any wage for years $t_{19}$ through $t_{24}$, centered on $t_{21}$ or $t_{22}$.[4,5] Given this approach, I observe $w_{20}$ for 77 percent of the NLSY79 women observed through age 40. I define a woman's starting wage as her wage in her first or second year in the labor force; of women who ever work, I observe starting wages for 90 percent of all women observed through 40.[6]

For other variable definitions:

- The NLSY79 collected information on fertility expectations – when a woman expected to have her next child, and how many (more) children she expected – in 1979, 1982 to 1986, and biannually from that point forward. For women with $K_1 > 0$, I build $k_1^{pred}$ using data for $t_1$ or the year before. For all women, I measure total children expected at approximately $t_1$.
- The gender-attitudes measure is the sum of the responses to a set of questions concerning family and gender attitudes. For instance, a woman was asked in 1979 how strongly she agreed with the following statement: "A wife who carries out her full family responsibilities doesn't have time for outside employment," or "A working wife feels more useful than one who doesn't hold a job." I align the seven questions so that a higher score reflects a more conservative view, and sum these responses.
- For father's occupation and women's expected occupation at age 35, I group detailed occupations into six broad categories: traditionally male professional jobs, traditionally female professional jobs, sales, clerical, services, and blue-collar jobs. Because of sample

full-time, or 2 years of part-time, schooling.

[3]When determining the last year of school versus the first year of work, in some instances of part-time (post-high school) education, I count her as already in the labor force if she was simultaneously working at a rate of 1000 hours or more per year.

[4]For all wage data in a given year $t$ I use the wage for the current job, or if not currently working, for the job that corresponds with a time closest to the middle of the calendar year (based on the date the job ended). In some cases of implausible values I use an alternate wage (corresponding to different jobs worked), or correct values that appear to be typos. In some cases where neither gives me a reasonable alternative, I treat it as a missing value. I also exclude all hourly wages below \$2 and above \$200, both in year 2000 dollars, with the upper bound lower for earlier survey years. (Wages are converted to year-2000 dollars using the Consumer Price Index for all urban consumers.) Because for each woman the 20th year falls beyond 1993, when the survey became biennial, I use only data reported for the survey years in building $w_{20}$. For those who entered the labor market in an odd year, I use the first available wage for their 22nd, 24th, and 20th career year (in that order), and for those who entered in an even year, I use their 21st, 23rd, or 19th.

[5]Throughout this appendix I refer to career years by number, for instance, $t_{17}$ for the 17th career year. For the purposes of this appendix, $t_{20}$ will therefore reflect the 20th career year, and "$t_{20}$" will reflect the five-year window used for defining $w_{20}$.

[6]Of those lacking a starting wage, 71 percent entered the labor force before 1978; the initial survey did not capture retrospective wage data.

size issues, for father's occupation I also combine various of these categories, depending on the woman's education level.

- For women with a high school diploma or less, I separately distinguish Hispanics from African-Americans. For women with more education, I group minorities (because the results for the two are very similar, or because of small sample sizes).

# C Sample Restrictions

$K_1 > 17$:

Given that my dependent variable is a woman's wage level at "$t_{20}$," I limit the sample to women with $K_1 \leq 17$, since among the latest mothers, I do not want to capture a sample that is influenced by the rate at which women return to work after $K_1$, which may be correlated with wages. Among the cross-section sample who I observe for at least 10 years after first birth, 86 percent return to work within 36 months, and 91 percent within five years. With the bulk of my sample using wages from $t_{21}$ or $t_{22}$, by limiting myself to women with $K_1 \leq 17$, most women will be back to work by this point.[7]

Age at First Birth > 35:

Among women from the NLSY79 cohort, there is evidence suggesting that as women approached the end of their 30s and the oncoming fertility decline, some chose to have a first child under lesser circumstances. To show this, Figure A-1 plots, by age at first birth, the proportion of women who were married at the conception of their first child, married by the birth of that child, and among the latter who were married, their husband's earnings that year.[8]

In the first panel, after age 35 we see a discontinuous drop in the proportion of women married at the conception of their first child. Comparing women with their first birth between the ages of 30 and 35, to those with a first birth after 35, 78 versus 61 percent are married (significantly different at the 1-percent level).

[7] In addition, in building $w_{20}$ I use only those wages that reflect a point greater then 36 months after a woman's first child was born.

[8] Age is measured in months, rounded to the closest integer. Figure A-1 includes all mothers observed through age 40; the figures are substantively identical if limited to the cross-section sample. (Women at later ages are grouped into cells of approximately the same size as those among women in their mid 30s.)

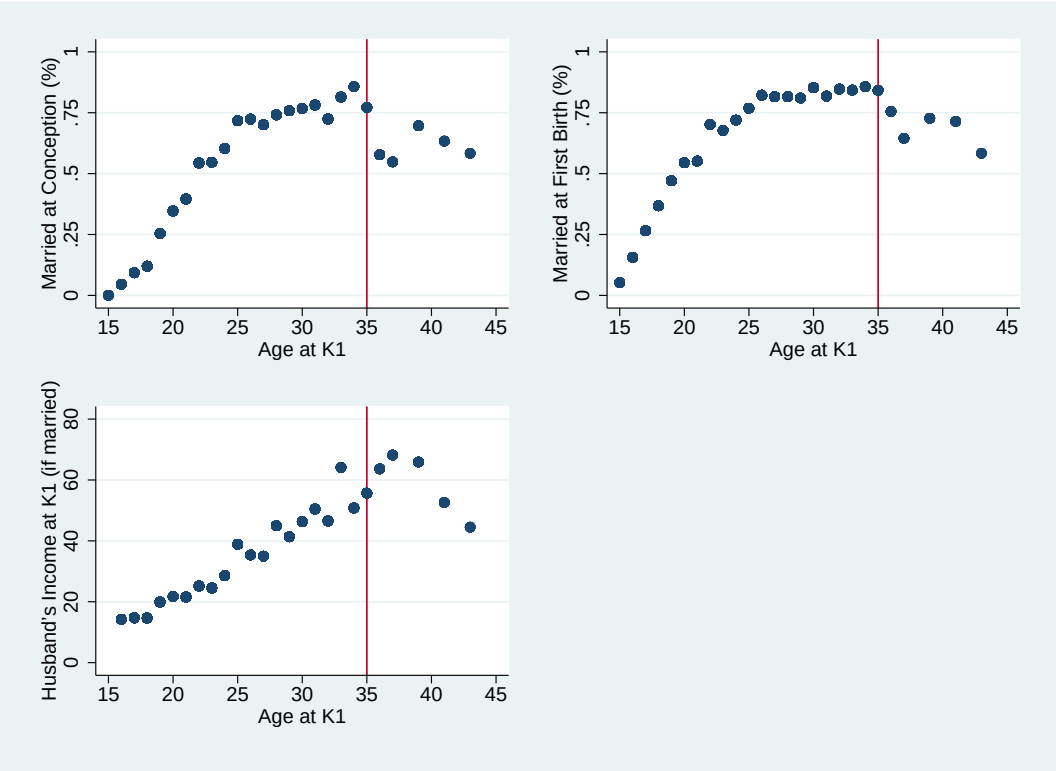


Figure A-1: Marriage Outcomes by Age at First Birth

Although the proportion of women who are married by the *birth* of that child is always higher, the same level drop appears after age 35. Furthermore, these differences are apparent within each education level. For instance, although a larger 11 percent of college graduates have their first birth after age 35 (compared to 5 percent overall), college graduates over 35 are 8 percentage points less likely to be married at conception than those in their early 30s (significant at the 15-percent level), and 11 percentage points less likely to be married by the birth (significant at the 1-percent level).

For women married at first birth, the last panel of Figure A-1 plots husband's average earnings that year. As expected, we see that husband's earnings rise with her age (and thus, on average, with his). In the late 30s, however, we see that the positive slope reverses.[9] Thus, even among women who were married, it appears that later mothers may have 'settled' for lower-earning spouses in order to have a child. All of these results are consistent with the predictions of the "spousal search" model in Schmidt (2007).

[9]If I regress husband's earnings on the quadratic of a woman's age at first birth, the coefficient on the quadratic term is almost significant at the 10-percent level, despite the small number of later mothers.

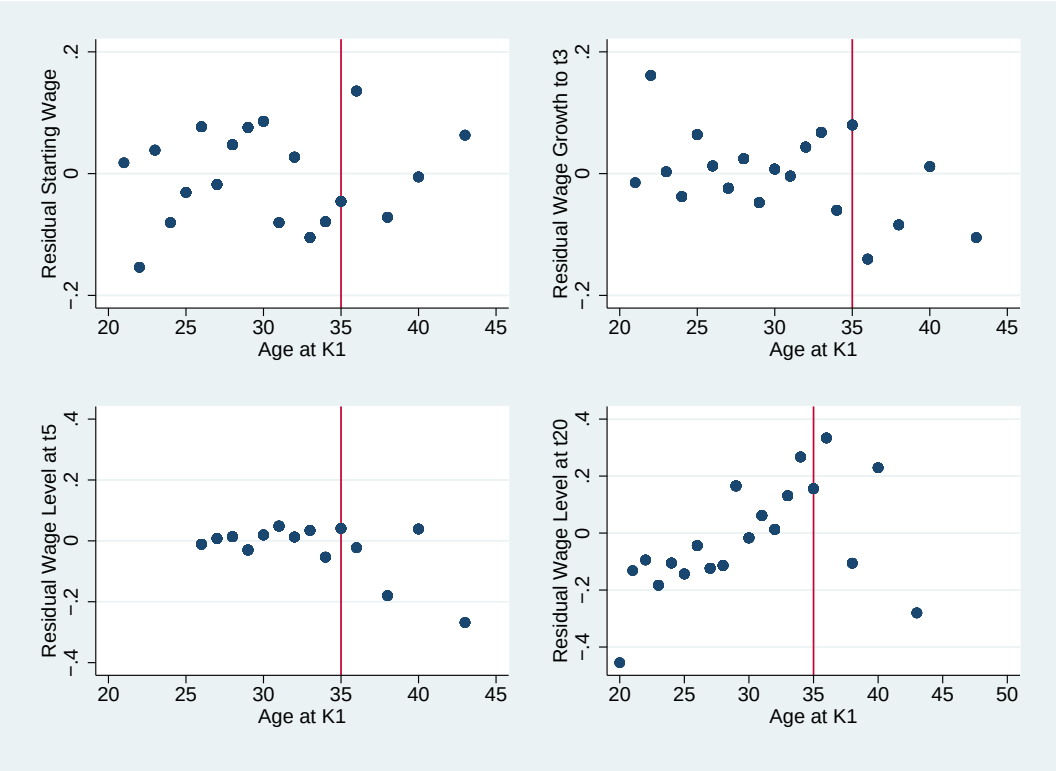


Figure A-2: Wage Outcomes by Age at First Birth

The evidence in Figure A-1 suggests that the older mothers from the NLSY79 cohort may possess some characteristic that was penalized in the marriage market. If that same characteristic influenced their labor market outcomes – consider the literature on the relationship between earnings and beauty (*e.g.*, Hamermesh and Biddle 1994) – later mothers may likewise have had lower wages because of this underlying characteristic.

The data suggest this is true, but only among the college graduates (the largest group of older mothers). To show this, Figure A-2 plots college graduates' mean residual starting wage ($w_0$), wage growth to $t_3$, wage level at approximately $t_5$ (limited to women with their first birth at least one year later), and wage level at "$t_{20}$" (after all had had a first child).[10]

As evident in the first panel of Figure A-2, women who went on to have a first birth after age 35 showed no systematic difference in starting wages. Yet the next two panels show that their wage growth lagged behind over time. In comparison to women with a first birth

[10]Using all college graduates observed to age 40 with a first child by $t_{17}$, these values reflect the residual of a regression of the given dependent variable on race, education at $t_1$, AFQT scores interacted by race, highest grade expected at age 18, and location at $t_1/t_5/$"$t_{20}$."

in their early 30s, they had significantly lower wage growth by $t_3$, and likewise significantly lower wages at $t_5$.[11]

The last panel of Figure A-2 plots the residual wage level at "$t_{20}$," after all of these college graduates had reached motherhood. Consistent with my results, we see an initial clear rise by age at first birth (reflecting the return to fertility delay), but this pattern reverses in the late 30s. The late mothers, who exhibited lower wage growth *before* motherhood, likewise have clearly lower wages afterwards. Based on this evidence, I have therefore excluded from my analysis those NLSY79 women with a first birth after age 35. If I were to instead include these women, my regression results in Section 6 would suggest a much lower return to delay for college graduates, driven by the spuriously low wages of these late mothers.

Given these two restrictions, my final sample of 2,567 mothers reflects 74 percent of all mothers observed till at least age 40.[12,13] Appendix Table A-1 reports an expanded set of summary statistics for this sample.

[11]The first comparison is significant at the 5-percent level, the second at the 1-percent level. If I limit myself to the cross section sample, the results are largely the same.

[12]Some women are also dropped because of missing labor force or school attendance data. Since the NLSY79 only captured retrospective labor force participation back to 1975, I also drop the handful of women who completed their schooling, or had a first child, before 1974.

[13]Of mothers observed till age 40, the sample captures 63 percent of high school dropouts, 79 percent of high school graduates, 74 percent of women with some college, and 69 percent of college graduates, with education measured at $t_1$ or when last observed (for those who never work). (I do not limit my sample to those women observed till age 40, although only 1.4 percent do not meet this criterion.)

Table A-1: Other Summary Statistics

| | All | Fertility Timing ($K_1$) | | | | | |
|---|---|---|---|---|---|---|---|
| | | $\leq -4$ | $-3-0$ | $1-3$ | $4-6$ | $7-9$ | 10+ |
| **Less than High School:** | | | | | | | |
| Age at first birth | 20.4 | 17.7 | 17.9 | 20.1 | 22.8 | 25.4 | 29.5 |
| Age at first intercourse | 15.9 | 15.7 | 15.6 | 16.0 | 16.7 | 15.7 | 16.9 |
| Family attitudes at $t_1$ | 14.5 | 14.5 | 14.3 | 14.4 | 14.9 | 14.5 | 14.1 |
| Mother's age at first birth | 20.2 | 19.6 | 20.1 | 19.8 | 21.0 | 20.9 | 20.4 |
| Total kids expected at $t_1$ | 2.3 | 2.6 | 2.3 | 2.2 | 2.2 | 2.5 | 2.0 |
| Total kids at "$t_{20}$" | 2.5 | 2.9 | 2.7 | 2.5 | 2.0 | 2.0 | 1.9 |
| Minority (%) | 24.5 | 39.1 | 28.0 | 18.9 | 12.3 | 18.1 | 8.2 |
| Mother's education | 10.3 | 10.0 | 10.2 | 10.5 | 10.5 | 9.5 | 10.9 |
| Self esteem (in 1980) | -0.39 | -0.48 | -0.42 | -0.45 | -0.21 | -0.18 | -0.48 |
| **High School Diploma:** | | | | | | | |
| Age at first birth | 23.5 | 19.0 | 18.8 | 21.0 | 23.6 | 26.4 | 30.9 |
| Age at first intercourse | 17.3 | 16.5 | 16.3 | 17.1 | 17.7 | 17.6 | 18.0 |
| Family attitudes at $t_1$ | 14.1 | 14.5 | 14.2 | 14.2 | 14.1 | 14.4 | 13.6 |
| Mother's age at first birth | 20.7 | 20.3 | 19.8 | 20.3 | 20.9 | 20.9 | 21.7 |
| Total kids expected at $t_1$ | 2.5 | 2.5 | 2.3 | 2.4 | 2.4 | 2.4 | 2.3 |
| Total kids at "$t_{20}$" | 2.2 | 2.7 | 2.5 | 2.5 | 2.2 | 1.9 | 1.7 |
| Minority (%) | 19.2 | 33.1 | 39.6 | 21.5 | 12.9 | 9.1 | 7.6 |
| Mother's education | 11.0 | 10.7 | 10.6 | 10.8 | 11.2 | 11.2 | 11.6 |
| Self esteem (in 1980) | -0.06 | -0.24 | -0.16 | -0.10 | -0.07 | 0.02 | 0.09 |
| **Some College Education:** | | | | | | | |
| Age at first birth | 25.5 | 19.2 | 20.3 | 23.0 | 25.4 | 28.1 | 32.0 |
| Age at first intercourse | 17.9 | 16.0 | 17.1 | 17.7 | 18.6 | 18.7 | 17.8 |
| Family attitudes at $t_1$ | 13.2 | 12.7 | 13.1 | 13.0 | 13.3 | 13.7 | 12.9 |
| Mother's age at first birth | 22.0 | 21.0 | 20.7 | 21.8 | 21.7 | 22.8 | 23.3 |
| Total kids expected at $t_1$ | 2.5 | 2.3 | 2.2 | 2.5 | 2.6 | 2.6 | 2.6 |
| Total kids at "$t_{20}$" | 2.3 | 2.6 | 2.5 | 2.5 | 2.2 | 2.1 | 1.8 |
| Minority (%) | 30.6 | 74.3 | 57.8 | 26.7 | 20.7 | 20.8 | 24.0 |
| Mother's education | 12.0 | 11.0 | 11.4 | 12.0 | 12.5 | 12.2 | 12.2 |
| Self esteem (in 1980) | 0.26 | 0.39 | 0.35 | 0.34 | 0.22 | 0.40 | -0.06 |
| **College Degree+:** | | | | | | | |
| Age at first birth | 28.9 | 19.6 | 22.2 | 25.5 | 27.7 | 30.7 | 33.8 |
| Age at first intercourse | 18.7 | 15.7 | 18.9 | 19.0 | 19.2 | 18.5 | 18.0 |
| Family attitudes at $t_1$ | 12.4 | 14.3 | 13.1 | 12.6 | 12.3 | 12.4 | 11.8 |
| Mother's age at first birth | 23.6 | 20.6 | 21.2 | 24.4 | 22.9 | 24.3 | 23.9 |
| Total kids expected at $t_1$ | 2.4 | 2.6 | 2.6 | 2.6 | 2.4 | 2.4 | 2.3 |
| Total kids at "$t_{20}$" | 2.2 | 2.5 | 2.6 | 2.5 | 2.3 | 2.1 | 1.8 |
| Minority (%) | 8.6 | 41.7 | 13.9 | 11.7 | 7.0 | 6.5 | 7.1 |
| Mother's education | 13.3 | 12.3 | 13.3 | 13.6 | 13.0 | 13.2 | 13.7 |
| Self esteem (in 1980) | 0.42 | 0.20 | 0.37 | 0.42 | 0.49 | 0.35 | 0.42 |

See notes to Table 1.

# D Influence of Sample Selection Criteria

Among NLSY79 mothers that I observe through at least age 40, Table A-2 compares the women included in my sample to those excluded. Grouping women by their timing of first birth ($K_1$), the top panel reports the percentage of mothers included in the sample (women who never work are included in the earliest group). The middle panel explores the reasons why the excluded women are lost from the sample, and the bottom panel compares the characteristics of the excluded and included women.

The top panel of Table A-2 shows that the proportion of women captured in my sample varies strongly by first-birth timing. Whereas I capture at least 81 percent of mothers with $K_1 \in \{-3, 11\}$, the sample captures only 52 percent of those with their first birth after $t_{11}$, and 43 percent of those with their first birth at least 4 years before $t_1$. The bottom panel shows that in each timing group, the women excluded are observed for fewer years beyond $K_1$, although even among the latest mothers, we see these women on average for at least another 10 years. We also see that in each group with $K_1 > 0$, those excluded take more time off from the labor market after having their first child.

<u>Women Not Observed to "$t_{20}$":</u>

The low proportion of women captured among the earliest mothers is worrisome if these women are systematically different from later $pt_1$ mothers. The second panel of Table A-2 shows that 17 percent of these excluded women are lost because they never enter the labor force. A much larger 64 percent are lost because we do not observe them through at least $t_{19}$ or $t_{20}$, the earliest years used for "$t_{20}$." (By comparison, among women with $K_1 > 0$, at most 11 percent are lost because they are not observed sufficiently long beyond $t_1$.)

In practice, the necessity of observing women through at least $t_{19}/t_{20}$ means that I lose all women with their first birth more than 15 years before $t_1$. This restriction is mild for women with at least a high school diploma at $t_1$, but for high school dropouts, among those who ever work, 7 percent of mothers have their first child before this point.

Table A-2: Comparison of Individuals Included Versus Excluded From the Analysis Sample

| **Timing of First Birth ($K_1$):** | $\leq -4$ | | | $-3-0$ | | | $1-3$ | | | $4-7$ | | | $8-11$ | | | $12+$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sample Size:** | | | | | | | | | | | | | | | | | | |
| Observed till age 40: | 791 | | | 573 | | | 609 | | | 760 | | | 460 | | | 378 | | |
| (%W) | (15.7) | | | (12.7) | | | (17.1) | | | (25.4) | | | (16.3) | | | (12.8) | | |
| Sample: | 324 | | | 486 | | | 528 | | | 638 | | | 368 | | | 189 | | |
| (%W) | (9.2) | | | (14.7) | | | (20.2) | | | (28.9) | | | (18.0) | | | (8.9) | | |
| **Captured in sample** (%W): | **42.9** | | | **85.1** | | | **87.0** | | | **83.7** | | | **81.2** | | | **51.5** | | |
| **Among excluded, reason: (%W):** | | | | | | | | | | | | | | | | | | |
| Never entered labor force: | 17.4 | | | 0.0 | | | 0.0 | | | 0.0 | | | 0.0 | | | 0.0 | | |
| $K_1 > t_{17}$ or age at $K_1 > 35$: | 0.0 | | | 0.0 | | | 0.0 | | | 2.3 | | | 2.0 | | | 74.0 | | |
| Missing $w_{20}$ because: | | | | | | | | | | | | | | | | | | |
| Not observed thru $t_{19}/t_{20}$ | 63.7 | | | 13.8 | | | 11.1 | | | 8.7 | | | 3.6 | | | 0.0 | | |
| Did not work during "$t_{20}$" | 16.6 | | | 50.5 | | | 57.9 | | | 63.1 | | | 70.4 | | | 22.6 | | |
| Missing wages during "$t_{20}$" | 2.2 | | | 35.7 | | | 31.1 | | | 25.8 | | | 24.0 | | | 3.4 | | |
| **Sample size:** | **467** | | | **87** | | | **81** | | | **122** | | | **92** | | | **189** | | |
| | **In** | | **Out** | **In** | | **Out** | **In** | | **Out** | **In** | | **Out** | **In** | | **Out** | **In** | | **Out** |
| **Characteristics Driving Inclusion:** | | | | | | | | | | | | | | | | | | |
| Fertility timing ($K_1$) | -6.7 | ** | -12.2 | -1.5 | + | -1.7 | 2.2 | | 2.3 | 5.4 | | 5.4 | 9.2 | | 9.3 | 13.5 | ** | 16.3 |
| Yrs observed beyond $K_1$ | 29.6 | ** | 27.2 | 28.5 | ** | 26.3 | 25.3 | * | 23.5 | 21.9 | ** | 19.4 | 18.1 | ** | 15.9 | 14.7 | ** | 10.2 |
| Time off at $K_1$ (mos) | - | | - | - | | - | 17.6 | | 22.3 | 19.6 | | 28.8 | 14.4 | | 18.3 | 7.4 | + | 14.1 |
| Hours worked during "$t_{20}$" | 5172 | ** | 171 | 7463 | ** | 1445 | 7524 | ** | 1187 | 6794 | ** | 936 | 6881 | ** | 902 | 6898 | ** | 4144 |
| **Background Characteristics:** | | | | | | | | | | | | | | | | | | |
| AFQT score | 31.5 | ** | 24.2 | 37.7 | | 31.9 | 47.3 | | 51.8 | 56.1 | | 60.5 | 57.4 | | 60.6 | 56.0 | | 59.2 |
| Rotter score | 0.22 | | 0.24 | 0.11 | | -0.09 | -0.01 | | -0.07 | -0.17 | | -0.11 | -0.19 | | -0.21 | -0.07 | | -0.25 |
| Expected schooling ('79) | 12.8 | | 12.5 | 13.3 | | 13.3 | 13.9 | | 14.2 | 14.1 | * | 14.6 | 14.3 | * | 15.0 | 14.1 | * | 14.6 |
| Education at $t_1$ | 11.8 | | 11.7 | 12.0 | | 11.8 | 12.8 | | 12.9 | 13.3 | * | 13.8 | 13.3 | ** | 14.0 | 12.8 | ** | 13.6 |
| Starting wage ($2000) | 7.08 | | 7.34 | 7.18 | | 7.23 | 8.43 | | 8.16 | 8.90 | | 9.49 | 8.63 | + | 9.40 | 7.90 | ** | 9.22 |
| Age at first intercourse | 16.0 | + | 16.3 | 16.3 | | 16.2 | 17.4 | | 17.7 | 18.0 | | 18.0 | 17.9 | + | 18.5 | 17.9 | | 17.8 |
| $k_1^{pred}$ | - | | - | - | | - | 3.9 | | 3.7 | 4.9 | * | 4.0 | 5.6 | | 5.3 | 5.6 | | 5.7 |
| Mother's education | 10.6 | * | 10.0 | 10.7 | | 11.0 | 11.4 | | 11.6 | 11.9 | | 12.3 | 12.0 | * | 12.8 | 12.2 | | 12.1 |
| **Sample Size:** | **324** | | **467** | **486** | | **87** | **528** | | **81** | **638** | | **122** | **368** | | **92** | **189** | | **189** |

**NOTES:** "%W" reflects proportions calculated using the original 1979 sample weights. The bottom panel likewise reports weighted means. Total hours worked during "$t_{20}$" only reported for those who are observed at least to $t_{19}/t_{20}$, the first year used as "$t_{20}$".

The bottom panel of Table A-2 also shows that in this earliest group, there are large differences in AFQT scores between those included and excluded from the sample. These differences remain if you consider the comparison only within education groups. Both among high school dropouts and high school graduates (which, combined, represent 82 percent of this group), those excluded from the sample have significantly lower AFQT scores.[14] Thus overall we see that my sample systematically loses women who take a long time to enter the labor force after having their first child, and that those women are in turn negatively selected.

This raises the question of the influence of this selection on my estimates of $\gamma$ and $\delta$. To consider this, I compare the starting- and 10th-year wages of $pt_1$ mothers, between those who I do and do not observe through $t_{19}/t_{20}$.[15] This lets me consider whether the women that I lose start at a systematically different wage level, or show a different wage growth pattern over their first 10 years in the labor force, as insight into whether their 20th-year wages are also likely to be systematically different.

In each education level, I find no statistical difference in starting wages.[16] Furthermore, this holds whether or not I control for characteristics observable at $t_1$, thus among the high school dropouts and high school graduates, the significant differences in AFQT scores do not translate into different starting wages.

For the set of $pt_1$ women for whom I observe a "$t_{10}$" wage, controlling for characteristics at $t_1$, within each education group I likewise find no difference in wage growth over this first 10 years. In combination, since the starting and "$t_{10}$" wages of the $pt_1$ mothers who do not reach $t_{19}/t_{20}$ are statistically equivalent to those who do, this suggests that their "$t_{20}$" wages would likewise be similar. Thus it is unlikely that my estimates of $\gamma$ or $\delta$ are affected by the systematic exclusion of women who I do not observe sufficiently long to reach "$t_{20}$."

[14] These differences remain when I exclude the women who never work.

[15] I build "$t_{10}$" wages using any observed wage for years $t_7$ through $t_{12}$, starting as close to $t_{10}$ as possible. Among the $pt_1$ women who ever work, I observe a "$t_{10}$" wage for 87 percent of high school dropouts, 90 percent of high school graduates, and 92 percent of women with some college.

[16] I do not consider college graduates, where the number of $pt_1$ women is small.

Women with No Observed Wage During "$t_{20}$":

Returning to Table A-2, consider now the low proportion of women captured in the sample among those with $K_1 \geq 12$. In the middle panel we se that this is driven primarily by women with a first birth beyond $t_{17}$ or after age 35. In the bottom panel, we also see that the excluded women have systematically higher education and starting wages. This arises primarily from the educational distribution; although the proportion of women with a first birth after $t_{17}$ is very similar across education levels, the proportion with their first birth after age 35 rises with schooling. Testing these mean differences instead within education groups, the only remaining significant difference is among women with some college, where the excluded women have systematically more education.

The greater concern among later mothers are the women who are missing from the sample because they never worked during the 5-year stretch used for "$t_{20}$." In most timing groups, roughly 50 to 70 percent of those excluded are lost for this reason. Among women with a first birth between $K_1 = \{-3, 11\}$, the remaining 25 to 35 percent excluded are lost because I do not observe a wage within this five-year stretch, even though these women worked at some point during that timeframe.[17]

This again raises the question of whether this exclusion influences my coefficient estimates. To consider this, for women with $K_1 > 0$, I compare the women excluded from my sample to those included on the following four dimensions: $(i)$ AFQT scores, $(ii)$ starting wages, $(iii)$ pre-motherhood wage growth to approximately $t_5$, and $(iv)$ wage level after first birth. For women with $K_1 \leq 0$, I run the same comparison for all but the third factor.

In particular, I start with the set of women with their first birth by $t_{17}$ and age 35, whom I observe through at least $t_{23}/t_{24}$ – the end of the range used for "$t_{20}$." To consider whether the excluded women show systematically different wage growth before children ($g_1$),

[17] Among women who I observe through the full five-year stretch of "$t_{20}$," those who work during this period but for whom I lack $w_{20}$ work on average 3,200 hours (or approximate 650 hours per year), compared to 7,400 hours among those captured in my sample (or approximately 1,500 hours per year).

I build a "$t_5$" wage for the set of women who remain childless by that point.[18] To test for differences in long-run wage growth among $pt_1$ mothers, and in post-birth wages among non-$pt_1$ women, I also build an alternate post-$K_1$ wage using data for any year between $t_9$ and $t_{34}$ (exclusive of the "$t_{20}$" period).[19]

The first three tests let me consider whether the women lost from my sample are systematically positively or negatively selected, in terms of ability, starting wage ($w_0$), and, among women with $K_1 > 0$, pre-birth wage growth ($g_1$). For the latter, I also consider whether the incidence of being missing from the sample is correlated with $k_1$. Any selection will only influence my estimate of $\theta$ if this incidence is in turn correlated with fertility timing.

The fourth test lets me consider whether heterogeneity in $\theta$ influences the probability that women remain working after children. One would suspect that the women who face an especially large drop in their wage growth after $K_1$ are more likely to drop out of the labor force. Using the alternate post-$K_1$ wage lets me consider whether the women who are unobserved during "$t_{20}$" have a systematically different $\theta$, although note that I cannot include in this test those women who never return to work.

Starting with women with $K_1 > 0$, for both high school dropouts and graduates I find no evidence of either positive or negative selection out of my sample. I find no statistical differences in AFQT scores, whether I consider only the set of women who never work during "$t_{20}$," or the full set of women with no observed $w_{20}$. I likewise find no difference in starting wages, and for high school graduates, no difference in pre-$K_1$ wage growth to "$t_5$."

Lastly, when I use the alternate post-$K_1$ wage, for high school graduates I find no difference in the estimates of $\hat{\theta}$ for those women included and excluded from my sample.

[18] I use any pre-$K_1$ wage for $t_3$ through $t_7$ (as close to $t_5$ as possible). Because the sample uses wages for as early as $t_3$, it captures the vast majority of women with $K_1 \geq 4$. (Of all non-$pt_1$ mothers, 73 percent of high school graduates, 78 percent of women with some college, and 83 percent of college graduates have $K_1 \geq 4$.) (I do not consider high school dropouts because only 33 percent have their first birth this late.)

[19] Among non-$pt_1$ women with missing $w_{20}$, I observe a post-$K_1$ wage for 86, 76, 71, and 60 percent, respectively, across the education levels. The proportion is smallest for college graduates in largest part because they enter motherhood slightly later, and so fewer have returned to work by their last year observed in the survey.

Among high school dropouts, $\hat{\theta}$ for the excluded women is instead significantly smaller than the estimate for the included women, although the combined coefficient remains statistically indistinguishable from zero, as in Table 2. In combination, these results suggest that my estimates of $\theta$ for women with a high school diploma or less are largely unaffected by the exclusion of women for whom I lack $w_{20}$.

Now considering women with at least some college education at $t_1$, mothers with missing $w_{20}$ have systematically *higher* AFQT scores. Yet these higher scores do not translate into significantly higher starting wages, and among those for whom I observe a pre-$K_1$ wage at $t_5$, women with missing $w_{20}$ have significantly *lower* wage growth by this point.

If this pattern continues, missing women will likewise have lower wages at "$t_{20}$," thus among women with some college, the mean $w_{20}$ within my sample may overstate the value in the population as a whole. Yet the incidence of missing values is not correlated with $k_1$, suggesting that this negative selection out of the sample may not affect my estimate of $\theta$.

When I use the alternate post-$K_1$ wage to directly test whether $\theta$ varies, however, I find that $\hat{\theta}$ for the missing women is significantly smaller. Calculating a combined coefficient, the estimate of $\hat{\theta}$ is 25 percent smaller than the estimate when calculated only on the set of women captured in my sample. Thus this evidence suggests that my results in Table 2 may overstate $\theta$ for women with some college at $t_1$.[20]

By contrast, for college graduates I find mild evidence of positive selection out of my sample, whether I consider only those who never work during "$t_{20}$," or also those who do. These women have higher AFQT scores, and although they enter the labor force with similar starting wages, I find that they likewise have somewhat greater wage growth by $t_5$. Although the *incidence* of being missing is uncorrelated with $k_1$, suggesting that this positive selection out of my sample should not influence my estimate of $\theta$, when I regress the alternate post-$K_1$ wage on first-birth timing, my estimate $\hat{\theta}$ for the missing women is significantly *larger*.

[20]See footnote 43 of the main text, however, for evidence that measurement error is simultaneously attenuating this coefficient towards zero.

Results from combining the two samples suggest that the coefficient reported in Table 2 for college graduates may therefore be as much as 30 percent too small.

Turning last to the $pt_1$ women, among high school dropouts and graduates, I find no differences between those for whom I do and do not observe a "$t_{20}$" wage, either in terms of ability, starting wage, or their "long-run" wage (the wage observed between $t_9$ and $t_{34}$).[21] (For women with at least some college, there are too few $pt_1$ women with missing $w_{20}$ to consider the influence of their exclusion.) Thus, just as with the case of women who do not reach $t_{19}/t_{20}$, I conclude that my estimates of $\gamma$ and $\delta$ are unlikely to be affected by the exclusion of these women with missing "$t_{20}$" wages.

[21] For both groups I observe a "long-run" wage for roughly 85 to 90 percent of the women for whom I lack a "$t_{20}$" wage.