

THE NARROWING OF THE U.S. GENDER EARNINGS  
GAP, 1959-1999: A COHORT-BASED ANALYSIS

Catherine Weinberger  
Peter Kuhn

Working Paper 12115  
<http://www.nber.org/papers/w12115>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2006

This material is based upon work supported by the National Science Foundation under Grant No. 0120111. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We would like to thank Solomon Polachek for extensive comments on an earlier draft of this paper, and Francine Blau for a very useful discussion. This paper also benefited from the comments of Deborah Cobb-Clark, Heidi Hartmann, Kevin Lang, Paul Latreille, and from participants in presentations at the Society of Labor Economists' Annual Meetings, the National Science Foundation, and at the Universities of British Columbia, Washington, and Western Ontario. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2006 by Catherine Weinberger and Peter Kuhn. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Narrowing of the U.S. Gender Earnings Gap, 1959-1999: A Cohort-Based Analysis  
Catherine Weinberger and Peter Kuhn  
NBER Working Paper No. 12115  
March 2006  
JEL No. J7

**ABSTRACT**

Using Census and Current Population Survey data spanning 1959 through 1999, we assess the relative contributions of two factors to the decline in the gender wage gap: changes across cohorts in the relative slopes of men's and women's age-earnings profiles, versus changes in relative earnings levels at labor market entry. We find that changes in relative slopes account for about one-third of the narrowing of the gender wage gap over the past 40 years. Under quite general conditions, we argue that this provides an upper bound estimate of the contribution of changes in work experience and other post-school investments (PSIs) to the decline of the gender wage gap.

Catherine Weinberger  
Department of Economics  
University of California, Santa Barbara  
Santa Barbara, CA 93106  
[weinberg@isber.ucsb.edu](mailto:weinberg@isber.ucsb.edu)

Peter Kuhn  
Department of Economics  
University of California, Santa Barbara  
Santa Barbara, CA 93106  
and NBER  
[pjkuhn@econ.ucsb.edu](mailto:pjkuhn@econ.ucsb.edu)

## 1. Introduction

After several decades of remarkable stability, the U.S. gender wage gap began to narrow after about 1980 (e.g. Blau and Kahn 2000). A well known explanation of this fact is based on changes in the rate at which women accumulate labor market experience over their lifetimes (e.g. Goldin 1989; O'Neill and Polachek 1993). According to this “experience” hypothesis, even though women might enter the labor market on a par with men, women’s earnings should grow more slowly with age than men’s, because women experience more career interruptions. However, as women’s commitment to the labor force has increased, the rate at which women’s earnings “fall behind” men’s should be less severe among more recent cohorts of women. Thus, a change across cohorts in the relative slopes of women’s (versus men’s) age-earnings profiles is a central, testable feature of the experience hypothesis.

In this paper, we argue that surprisingly little is currently known about whether such a change in slopes has actually occurred. This is because existing analyses of the evolution of the gender wage gap have concentrated their attention on the earnings premium associated with potential (or actual) experience across a series of cross-section regressions.<sup>1</sup> As is well known, however, cross-sectional patterns can sometimes give a misleading picture of what happens to a cohort of workers over time. In one well-known example, Borjas (1985) showed that what appears in cross-section data to be a difference in earnings slopes between two groups (in his case, immigrants versus natives) could instead be a difference across cohorts in immigrants’ earnings levels at labor market entry (i.e. “cohort quality”).<sup>2</sup> To avoid such problems, this paper explicitly examines women’s *within*-cohort relative wage growth using data from the 1960

---

<sup>1</sup> See for example O'Neill (2003), who presents dramatic evidence that the women’s cross-sectional return to potential experience has increased.

<sup>2</sup> Relatedly, but distinctly, this paper argues that what appears in repeated cross-sections to be a difference *across generations* in the slopes of women’s relative lifetime earnings profiles could be caused purely by a difference across generations in women’s entry-level relative wages.

through 2000 Censuses and the 1964 through 2004 Current Population Surveys.<sup>3</sup> Our analysis deviates from previous work on the gender earnings gap in another important way. Instead of dividing changes in the gender gap into a portion explained by changes in women's human capital and a portion explained by changes in "discrimination", we focus (as noted) on distinguishing changes in earnings slopes versus levels. This new dichotomy is distinct from the more traditional one, and enables us to glean new information from existing data.

Our main results are as follows. First, in contrast to a simple reading of the experience hypothesis, women's earnings do not monotonically "fall behind" men's in *any* cohort as it is followed over time in our data. Instead, women's relative wages follow a U-shaped pattern with age over the life of a typical cohort, falling behind at first, but recovering after that.<sup>4</sup> Second, once we allow for a U-shaped "baseline" age effect, our data do exhibit growth across cohorts in the relative slope of women's age-earnings profile, consistent with the experience hypothesis. Third, we estimate that this growth in relative slopes accounts for about one-third of the narrowing of the gender wage gap over the past 40 years, with the remainder due to a declining wage gap at the time of labor market entry. Finally, we provide evidence of other changes that are likely to have contributed to the growth in relative slopes, suggesting that one third represents an upper bound on the importance of rising experience and other post-schooling investments to the narrowing of the gender gap.

---

<sup>3</sup> Some existing studies of the gender gap do, of course, present descriptive statistics that allow cohorts to be followed over time (e.g. O'Neill and Polachek 1993, Table 2; Blau and Kahn 2000, Table 1). To our knowledge, however, formal statistical analyses of trends in within-cohort rates of women's relative wage growth have not been conducted.

<sup>4</sup> A U-shaped pattern of women's relative wages is predicted by human capital investment models in which women can avoid depreciation costs if they defer human capital investments until after the childbearing period (Polachek 1975; Weiss and Gronau, 1981). This empirical pattern is also consistent with models in which the level of discrimination falls toward the end of the life cycle, relative to levels prevailing at the time of cohort entry (Blau and Kahn 2000).

Section 2 of the paper introduces a simplified version of the “experience” hypothesis. It also describes a simple competing explanation of the declining gap based purely on cohort effects on wage levels, and demonstrates that these two models *cannot* be distinguished by comparing changes over time in the slope of cross-sectional age-earnings profiles. Sections 3 and 4 present our results from Census and CPS data respectively. Section 5 discusses the interpretation of those results and Section 6 concludes.

## 2. The Implications of Post-Schooling Investments: A Simple Model

To be more precise about the implications of gender differences in work experience and other post-schooling investments (PSIs) for women’s earnings profiles, we begin with an illustrative example. Suppose that, for men, the log wage rate commanded by individual  $i$  is described by:

$$y_{it} = Y^0 + .03e_{it} + h_i + v_{it} \quad (1)$$

where  $Y^0$  gives the average log earnings of men at labor market entry,  $e_{it}$  is a measure of potential labor market experience for individual  $i$  at time  $t$ ,  $h_i$  captures time-invariant person-specific productivity (incorporating variation in both pre-market schooling and ability, with mean zero across the full population of men), and  $v_{it}$  is an *iid* error term.<sup>5</sup>

Under this model, on average, men with  $e$  years of potential work experience earn:

$$y(e) = Y^0 + .03e \quad (2)$$

Under the experience hypothesis, women accumulate fewer years of actual experience (and make fewer other earnings-enhancing investments, such as changing locations to enhance

---

<sup>5</sup> In the empirical implementation, men’s earnings are assumed to be quadratic in experience. However, including an experience squared term here would simply complicate the example, with no change in its qualitative features.

their careers) per year of potential experience than men. This might lead to the following observations on women's average annual earnings:

$$y_c(e) = (Y^0 - G_c) + .03(\lambda_c e) \quad (3)$$

This formulation is consistent with a model in which, for each year of potential experience, women of cohort  $c$  accumulate only a fraction  $\lambda_c$  of the actual experience (or other PSIs) gained by men.<sup>6</sup> The constant  $G_c$  captures any gender differential in cohort  $c$ 's average earnings that are present upon labor market entry. In a pure human capital interpretation,  $G_c > 0$  might reflect a negative mean of  $h_i$  across the population of cohort  $c$  women.<sup>7</sup>

Mathematically, the gender wage gap ( $GG(e,c)$ ) faced by the average cohort  $c$  woman at age  $e$  is the difference between equations 2 and 3<sup>8</sup>:

$$GG(e,c) = G_c + .03(1-\lambda_c)e \quad (4)$$

Women's relative wages ( $RW(e,c)$ ) are defined as:

$$RW(e,c) = 1 - GG(e,c) = 1 - G_c - .03(1-\lambda_c)e \quad (5)$$

Note that in our example, for any given level of  $e > 0$ , the gender gap will be smaller (and women's relative wages will be higher) if  $\lambda_c$  is larger (closer to 1). The heart of the experience hypothesis is the argument that  $\lambda_c$  is growing with the entry of each new cohort, and that the average gender gap is shrinking as a result.

Two properties of the gender gap under this experience hypothesis are easy to establish:

- 1) At the time of labor market entry ( $e=0$ ), the relative wages of women in cohort  $c$  are independent of  $\lambda_c$ .<sup>9</sup>

---

<sup>6</sup> Empirical evidence consistent with this model includes the observation that women's earnings tend to be more highly correlated with actual than with potential experience (Mincer and Polachek, 1974).

<sup>7</sup> In order to simplify this example, the assumption that wages reflect productivity is maintained here. The implications of changes in discrimination are explored later in the paper.

<sup>8</sup> Formally, the log wage difference in (4) is of course only an approximation of the gender gap. A more precise measure of the gender gap is  $1 - \exp(-GG(e,c))$ .

- 2) Within any given cohort, women's relative wages will decline more slowly with age if women's relative rate of PSI investment ( $\beta_c$ ) is higher.<sup>10</sup>

In Figure 1, these properties are depicted graphically, assuming that each cohort of women enters the labor market at a 20 percent wage disadvantage relative to men ( $G_c = .20$ , for all  $c$ ), and that  $\beta_c$  grows by .1 in each successive cohort:  $\beta_c = .2 + .1c$ .<sup>11</sup> Figure 1 illustrates the following key features of the experience model: first, despite the fact that women's relative wage at labor market entry is constant across generations of women, the average level of women's relative wages in any particular year (given by a weighted average of the vertical array of points in any particular year) is rising over time; this entire increase is accounted for by the changing slope coefficients (the growing  $\beta$ s). Second, the gender gap in the estimated "returns" to potential experience estimated from a cross-section regression also declines across years. This is illustrated by the fact that the vertical array of points is closer together in 2000 than in 1960.

Next, for purposes of comparison, Figure 2 depicts a contrasting model of the rise in women's relative wages. In this "pure cohort effects" model, the slope of women's and men's age-earnings profile remains constant across all cohorts; however women's relative wages at labor market entry are changing as new cohorts enter.<sup>12</sup> As in Figure 1, this model also has

---

<sup>9</sup> In Section 6 we discuss an extended experience model in which a cohort's *expected* experience can affect its average wage at the time of labor market entry. In the standard human capital model, this should lead women's entry wages to *fall*, rather than rise, as their expected labor market attachment rises when pre-market investments are held constant.

<sup>10</sup> Of course, as Weiss and Gronau (1981) point out in a continuous-time model, sharp declines in the rate of investment, can, under some circumstances, lead to periods of faster earnings growth for demographic groups *expecting* future career interruptions. But the broad conclusion that mean earnings differences between early and later phases of the life cycle should be smaller for groups making fewer investments is a robust and widely-accepted prediction of almost any standard model of PSI.

<sup>11</sup> For convenience the time frame for these examples was chosen to mirror exactly our Census data, which follow eight ten-year birth cohorts ranging from 1897-1906 to 1967-1976 over the five census years 1959 through 1999. Cohorts in Figure 1 and subsequent figures are labeled by the year in which the cohort's median age was 27; thus for example cohort 1 --our oldest--, was born in 1897-1906, and was between the ages of 23 and 32 in 1929).

<sup>12</sup> Of course, another possible alternative to the "experience" hypothesis is a pure "year effects" model rather than a pure "cohort effects" model. We discuss the performance of such a model, and the identification issues that arise in distinguishing year effects, cohort effects and (changing) age effects in Section 5 of the paper.

rising women's wages over time; however in this case none of the increase is explained by changes in the returns to potential experience. More importantly, *despite the fact that every cohort's age-earnings profile has the same slope*, estimates from cross-section regressions show that the gender gap in returns to potential experience declines across years (the points are closer together vertically in 2000 than in 1960).<sup>13</sup> We conclude that caution is required in drawing conclusions from time trends in the slope of cross-sectional experience-wage profiles alone. Only the pattern observed in Figure 1, with slopes changing between cohorts, is consistent with the experience hypothesis.

### 3. Results: Census Data

#### a. Estimating Gender Gaps

Our Census samples comprise U.S. born, full-time, full-year white workers aged 23-62 in the years 1959, 1969, 1979, 1989 and 1999. Simple descriptive statistics for these samples are provided in Appendix 1; their main features are well known.<sup>14</sup> In our analysis, gender earnings differentials are estimated for four birth cohorts in any given year, corresponding to workers who attain the age ranges 23-32, 33-42, 43-52 and 53-62 in that year. Altogether, the analysis includes at least one year of data for each of eight ten-year cohorts with birth dates ranging from 1897-1906 for the oldest cohort to 1967-1976 for the youngest. In what follows, we refer to the oldest cohort as number 1, followed by cohorts 2 through 8 in turn.

Coefficients from cross-sectional earnings regressions using the above data are reported in Table 1. Since the sample is restricted to full-time, full-year workers and detailed hours controls

---

<sup>13</sup> Mathematically, this results from Figure 2's assumption that the rate of decline in the gender wage gap is *decelerating* across cohorts. This would be the case, for example, if women's cohort-specific relative wage at labor market entry asymptotically approaches one from below, which strikes us as a plausible scenario.

<sup>14</sup> The difference in the 1999 mean gender gap between the Census and CPS data is due to measurement error, and disappears when data from surrounding CPS years (1995-2003) are aggregated.

are included, the dependent variable should be interpreted as an hourly rate of pay. In addition to these hours controls, the Table 1 regressions include standard (and comparable) controls for education and region, plus a quadratic in age (to capture the life-cycle pattern of men's wages). Finally, to allow women's wages to evolve differently over the life cycle than men's in as flexible a manner as possible, we include four gender-age interaction terms in each Census year. By construction, the gender coefficients along the diagonals of Table 1 thus describe the gender gap faced by a given cohort of women as those women are followed over time.

Several patterns in Table 1 are immediately obvious: Gender coefficients tend to be larger among older cohorts than among younger cohorts observed in the same year (vertical); gender coefficients fall if a given age group is followed over time (horizontal); but are surprisingly constant when a given cohort is followed over time (diagonal). These gender coefficients are also depicted graphically in Figure 3, with observations from the same cohort connected by lines.<sup>15</sup> On closer inspection of the within-cohort trends, we also see that gender wage gaps widen for every cohort as its median age rises from 27 to 37, and *fall* for every cohort between the (median) ages of 47 and 57, reflecting the nonmonotonic pattern described earlier. Between the ages of 37 and 47, we see a widening of the gender wage gaps for the two oldest cohorts observed in those age ranges, but a narrowing for the two youngest. This more subtle pattern suggests an increase across cohorts in the overall slope of women's relative age-wage profile. This first glance at the data reveals patterns consistent with at least some role for the "flattening slopes" hypothesis, as depicted in Figure 1.

---

<sup>15</sup> For ease of interpretation, Figures 3 and 4 transform the gender gap (GG) coefficients of Tables 1 and 2 into female relative wages (RW) (by adding one to the absolute value of the coefficient). This makes them directly comparable to Figures 1 and 2.

## b. Modeling the Evolution of the Gender Wage Gap

To quantify the role of age-cohort interactions in explaining the recent decline in the gender wage gap, we now use the 20 age- and year-specific gender wage gaps estimated in Table 1 (and depicted in Figure 3) as data points in some simple aggregate regressions.<sup>16</sup> In particular, we estimate the following empirical model of women's relative wages:

$$RW_j = \beta + \sum_{a=1}^3 \beta_a A_j(a) + \sum_{c=2}^8 \beta_c C_j(c) + \sum_{c=5}^7 \beta_c \text{age}_j C_j(c) \quad (6)$$

In this model,  $RW_j$  describes women's relative wages in each of the 20 age x cohort cells. The effects of age on relative earnings within cohorts 1-4 are captured by  $\sum_{a=1}^3 \beta_a A_j(a)$ , where  $A_j(a)$  is an indicator for  $j$  belonging to age group  $a$ . Pure cohort effects on women's relative earnings are captured by the term  $\sum_{c=2}^8 \beta_c C_j(c)$ , where cohort 1 (the oldest—born in 1897-1906) is the reference category, and  $C_j(c)$  is an indicator for  $j$  belonging to cohort  $c$ . Finally, for each of cohorts 5 through 7, the interaction term (modeled as the product of age and the cohort indicator) allows the effect of age on women's relative earnings to differ from its effect in cohorts 1 through 4.<sup>17</sup> In the regressions,  $\text{age}_j$  is scaled to measure potential experience in decades elapsed since the first observation of the cohort, so  $\text{age}_j \in \{0,1,2,3\}$ . If each successive cohort of women after cohort 4 had a steeper age-relative wage profile, we would observe that  $0 < \beta^5 < \beta^6 < \beta^7$ .<sup>18</sup>

<sup>16</sup> We could, of course pool the microdata from all years, and estimate the influence of cohort-age interactions on the gender wage gap in a one-stage procedure. We prefer the current approach because it allows for an intuitive and visual analysis of the 20 gender-wage gaps that interest us here. Since the samples from which these 20 data points are estimated are so large, the consequences of using estimated coefficients as regressors are minimal.

<sup>17</sup> For practical reasons, we chose to group cohorts 1-4 together to estimate the baseline profile. Figure 4 shows little change in the age-relative wage profile within this group of cohorts.

<sup>18</sup> Since cohorts 1 and 8 are observed only once in these data, separate age effects on earnings cannot be estimated for them.

OLS estimates of equation (6) using the 20 estimated values of  $RW_j$  in Table 1 as data are presented in column 1 of Table 2. Overall, these results provide support for both the changing levels and changing slopes models. As predicted by the changing levels model, cohorts of women born later earn significantly more, relative to men, than cohorts born earlier, even at labor market entry. And, as predicted by the experience hypothesis, younger cohorts of women exhibit a higher rate of age-related relative wage growth than older cohorts. Together, the “composite” model in column 1 fits our data on gender wage gaps almost perfectly, with an adjusted  $R^2$  of .96. But how can we assess the quantitative contribution of changes in the slope of age-earnings profiles across cohorts to the recent narrowing of the gender wage gap?

One way to pose this question is to ask how well we can explain recent trends without recourse to steepening age-wage profiles at all. To that end, column 2 of Table 2 restricts all the cohort-experience interaction terms of column 1 to equal zero ( $0 = \beta^5 = \beta^6 = \beta^7$ ). According to this model, there has been a large and statistically significant increase in women’s relative wages upon labor market entry across cohorts. As might be expected, the growth in estimated cohort effects is somewhat larger in this model than in column 1. Somewhat less expected, though, is the fact that relatively little explanatory power is lost by dropping the interaction terms from the model: adjusted  $R^2$  only falls to .91 from .96. *Thus, a parsimonious model that fits the past four decades of data on the U.S. gender wage gap surprisingly well has each successive cohort of women entering the labor market at a higher wage relative to men, with each cohort having the same rate of wage growth, relative to men, as every other cohort.*

For comparison, column 3 of Table 2 estimates a “pure changing slopes model” (as depicted in Figure 1) where women’s relative entry wages are constrained to be the same across all cohorts (and years) but their relative rate of age-related wage growth can differ across cohorts.

Compared to the composite model of column 1, the cohort-experience interactions are now much stronger, as the model attempts to fit the declining gender gap across cohorts using slope terms only. Clearly, however, with an adjusted  $R^2$  of .72 this model does a considerably worse job of fitting the data than either the composite or the pure cohort model. We conclude that, if an analyst had to choose only one of these two polar case models to describe the evolution of the gender wage gap over the last 40 years, he or she should choose the changing levels model over the changing slopes model.

### **c. Decomposing the decline in the gender wage gap**

An alternative way to quantify the effects of steepening relative age-wage profiles on the narrowing of the gender wage gap uses the coefficients estimated in the “composite” model of column 1, Table 2 to predict what the 1999 gender wage gap might have been in the absence of these effects. By comparing the actual change in the gender wage gap with the change that would have occurred under the counterfactual assumption of no changes in slopes, we can estimate the relative importance of changing slopes to the narrowing of the gender wage gap.

The first two columns of Table 3 are similar to the regressions of Table 1, columns 1 and 5, except that only a single gender coefficient is estimated for women of all ages in 1959 (column 1) and 1999 (column 2). The estimated log wage differentials are -0.540 in 1959, and -0.278 in 1999. The third column is similar, but the dependent variable is the actual 1999 log wage minus the potential experience\*cohort interactions estimated in Table 2, column 1 (zero for men and for the youngest women, .068 for women age 33-42, .067\*2 for women age 43-52, and .051\*3 for women age 53-62). Under the counterfactual, the 1999 log wage differential is

-0.361 . In other words, about one-third of the closing of the gender wage gap between 1959 and 1999 ( $(.361-.278)/(.540-.278)=.32$ ) can be attributed to changes in slopes.<sup>19</sup>

#### 4. Results: CPS data

To assess the robustness of our results to the data source, in this section we replicate the main aspects of our analysis using March CPS data.<sup>20</sup> The main advantage of these data over the Census is that they allow us to estimate an annual series of gender wage gaps, disaggregated by exact year of age rather than aggregated age group. The key disadvantage is fewer observations per year, and a later initial observation. To overcome the absence of early observations, we simply used the 1960 Census data as a proxy for 1960-1963 CPS observations, so that we can compare CPS with Census results over the same 1960-2000 observation window.<sup>21</sup>

CPS estimates are based on 41 years (1960-2000) and 40 ages (23-62), broken down into 10-year, 5-year, or 1-year cohorts and age groups. Rather than run dozens of cross-section regressions with between 4 and 40 gender\*age coefficients apiece, we use the other common two-stage procedure to compute the gender gaps. In the first stage, male-only regressions are run for each year, with controls identical to those used in the Table 1 regressions. Gender gaps for each individual woman are computed as the difference between her actual earnings and the earnings predicted if she were paid as much as a man with the same observable characteristics.

---

<sup>19</sup> The contribution of changing age distributions was estimated to be negligible. When the 1959 women were reweighted to match the 1999 age distribution, the estimated gender coefficient fell by only 0.002 . When *pe\*cohort* interactions were estimated from a single stage regression (gender interacted with age, cohort, and the *pe\*cohort* interactions; year interacted with education, region, and hours per week), the estimated contribution of changing slopes fell from 0.32 to 0.29 .

<sup>20</sup> As with the Census, we observed annual earnings for the year preceding the survey date. These data were collected in 1964-2004, so we have earnings observations for 1963-2003.

<sup>21</sup> The assumption was made that the wage *structure* remained stable during these early years, i.e. that the gender gap for a 39 year old in 1960 was a good proxy for a 39 year old in 1962. In sensitivity testing, the later CPS observations (2001-2004) are also included.

The data points used in the second stage regressions are mean estimated log wage differentials for each age group\*cohort cell, estimated from male-only regressions.<sup>22</sup>

Table 4 reports the results of the second-stage regressions. In column 1, the Census data are subjected to this procedure, with the same 10-year age and cohort groups as before. In column 2, the CPS data from the five Census years are used, with very similar results. Column 3 still uses 10-year age and cohort groups, but includes data from between-Census years. (Here, the age-group controls indicate the proportion of the cell in the age group). Column 4 does the parallel analysis using 5-year cohorts and age groups. Column 5 is similar to columns 1 and 2, but with 1-year cohort and age groups, for a total of 1640 1-year cohort\*age cells. This final specification includes 40 age levels and 80 cohort-specific fixed effects. The estimated potential experience-cohort (*pe\*cohort*) interaction effects are similar in all five specifications.

All specifications produce comparable estimates of the *pe\*cohort* interactions, and find U-shaped age-relative earnings profiles with upswing after age 47.<sup>23</sup> To help assess the magnitudes of the experience effects estimated from each of the 5 specifications, the analog of the Table 3, column 3 counterfactual log wage differential is included at the bottom of each column.<sup>24</sup> In every case, the CPS estimates of the importance of changing slopes to the overall change in the gender wage gap are very close to the corresponding Census estimate (based on male coefficients). These estimates do not change much as the within-cell age-range shrinks, indicating that the larger groupings do not obscure important changes in slope. All estimates,

---

<sup>22</sup> To maintain comparable units between specifications, we estimate potential experience\*cohort interactions using the mean of (age/10) in the cell, minus the mean of (age/10) at the first observation of the cohort. In specifications where cohorts do not always fit neatly into a single age group, the age group indicator is replaced with a vector indicating the proportion of the cell in that age group.

<sup>23</sup> The column 5 specification contains 39 age coefficients, too many to report in Table 4. These are monotonically increasing from age 47 through 62.

<sup>24</sup> Here, we subtract the *pe\*cohort* interactions estimated in the corresponding specification from women's actual 1999 log earnings, then estimate the 1999 log wage differential using this counterfactual as the dependent variable.

both Census and CPS, attribute 30-40 percent of the change in the log wage differential to changing slopes.

Additional sensitivity testing finds that these results are robust to variations in specification. For example, the estimated importance of potential experience\*cohort effects is virtually unchanged if we use an estimate of age minus education (rather than simply age) as our measure of potential experience.<sup>25</sup> Similarly, there is very little change if the oldest workers are eliminated from the sample.<sup>26</sup> If the analysis is restricted to college graduates only, changes in slopes play a somewhat larger role when the Census data are used, but a somewhat smaller role when the CPS data are used. Although we expected work experience to play a much greater role in the closing of the college graduate gender gap, our data do not provide clear support for this hypothesis.

Only one experiment led to a substantial change in the results. When the window of observation was shifted by two or four years (1962-2002, or 1964-2004), the estimated potential experience\*cohort effects fell substantially, implying an even smaller role for the experience hypothesis.<sup>27</sup> For example, when the Table 4, column 5 specification was run with the window shifted by two years, the estimated importance of changing slopes fell from 39 percent of the closing to only 24 percent of the closing. When the window was shifted by four years, it fell even further, to only 12 percent. This sensitivity to the time frame chosen suggests that a more complete model should incorporate a time trend-- a possibility we explore in the following section.

---

<sup>25</sup> Since Census questions do not always allow us to estimate total years of education received, and since men and women had similar education levels throughout our sample period, we simply used age-22 as our potential experience measure. When this is changed to age-education-6, estimated *pe\*cohort* effects were virtually unchanged, but slightly smaller.

<sup>26</sup> Estimated *pe\*cohort* effects were similar when the age range was 23-52 rather than 23-62.

<sup>27</sup> The two- or four-year shift included the start and end dates of the sample, and the birthdates of cohorts used to calculate the *pe\*cohort* interactions.

## 5. Interpreting the Changes in Slope

In our discussion so far, we have made two assumptions that considerably simplify the interpretation of our empirical results. One of these is the (implicit) assumption that the population of women employed at any point in time (for whom we have wage data to analyze) is representative of all women who ever participate in the labor force. How can we interpret our estimates if this is not true? To address this question, we note first that --since our interest is in changes over time-- nonrandom selection of women into work will not affect our estimates of the importance of PSIs if the nature of selection is constant over time. But what if the nature of women's selection into work has become increasingly positive (with respect to unobserved ability) over the past 40 years, as both Blau and Kahn (2004) and Mulligan and Rubinstein (2005) have recently argued? Of course, if changing selection into work has been purely a cohort-specific phenomenon, its effects would be completely absorbed by our cohort coefficients; in this case our estimates of cohort-age interactions, and thus the importance of PSIs would remain unbiased.

Thus, the only way that changing selection of women into work could lead our procedure to underestimate the contribution of PSIs to women's relative wage growth would be if (a) there were *within-cohort trends* towards women's positive selection into work *and* (b) these trends were weaker within more recent cohorts (thus tending to flatten women's age-wage profiles across cohorts). While this is possible, the opposite seems more likely to us: that the recent trend towards positive selection into work has had both a between- and a within-cohort component. In this case, the changes in slope that we estimate here will include within-cohort

selection effects, and will thus *overstate* the true contribution of changing post-schooling investments to the narrowing of the gender wage gap.<sup>28</sup>

A second key assumption in our analysis so far is the absence of any pure *year effects* in our models of the gender wage gap. In addition to some well-known reasons related to model identification (discussed below), we formulated our baseline model in terms of cohort and age effects (and their interactions) because in some sense year effects are the hardest to interpret in this context: cohort effects naturally capture pre-market investments and cultural differences across generations, while age effects (and the changes therein across cohorts) naturally capture human capital factors such as post-schooling investments. Aside from a decline in discrimination, it is harder to think of reasons why the relative price employers are willing to pay for female versus male labor of given age and experience would rise purely as a function of calendar time across all cohorts.

That said, however, suppose now that women's relative wages *are* also affected by a simple time trend in the relative price of female labor (caused by an increase in relative demand for female labor, such as a decline in discrimination). How would the presence of such a time trend affect our conclusions? Of course it is impossible to estimate a regression that adds a complete set of year effects to equation 6: year, cohort and age effects are not separately identified because year equals cohort plus age (see for example Deaton and Paxson 1994 for an excellent exposition). Still, it is important to ask whether the presence of such a trend would affect our quantitative assessment of the contribution of cohort-age interactions to the decline in

---

<sup>28</sup> Additional insights on changing patterns of selection into work could of course be gained from examination of long, nationally representative panel data sets. While certainly of interest, this is beyond the scope of the current paper.

the gender wage gap.<sup>29</sup> We address this question formally in Appendix 2; the results are as follows. First, if there is a true, positive linear time trend in women's relative wages, both the age effects ( $\beta$ 's) and the cohort effects ( $\gamma$ 's) presented in all our Tables so far will be overestimates, because any true year effects will be absorbed into these age and cohort coefficients. Importantly, however, our estimates of cohort-year interactions (the  $\delta$ 's) will remain unbiased. More importantly, it follows that *the share of the year-to-year decline in the gender gap attributed by our model to cohort-experience interactions (versus cohort and year effects together) is invariant to the magnitude of any linear time trend affecting women's wages.*

Unfortunately, as Appendix 2 goes on to show, the above reasoning does not extend to nonlinear time trends; in such cases our estimates of the relative importance of cohort-year interactions (versus all other factors together) will be biased. In particular, when the time trend is accelerating, our estimated changes in earnings slopes will overstate the true trend in cohort-year interactions; when the time trend is decelerating, changes in slope understate the trend in cohort-year interactions. Given this information, is it possible to place bounds on the true contribution of PSIs to the declining gender gap?

To do this, we first argue that our best empirical evidence about time trends, net of PSI effects, comes from observation of gender gaps among very young workers. These are depicted graphically in Figure 5, using both Census and CPS data for ages 23-27. These series are flat between 1960 and 1970, then increase at a constant rate until 1990 or 1995 before flattening out again. If *any* part of this nonlinear trend reflects year rather than cohort effects, then the changes

---

<sup>29</sup> As is very well known, when an outcome variable (say  $RW$ ) is modelled as a linear function of age, cohort and year, then (because  $\text{year} = \text{age} + \text{cohort}$ ) a regression of  $RW$  on any two of these three variables is numerically equivalent to a regression on any other two. Less well known is the fact that this observational equivalence breaks down when age, cohort and year are each represented as a vector of fixed effects. To see this most quickly, suppose (as is the case in this paper) the the number of years defining a time period is the same as the number defining a cohort. Then in *any* data set, the number of time periods *cannot* equal the number of cohorts. In data (like ours) collected as a series of cross sections, there will be more cohorts than years (8 versus 4 in our case). In data sets that follow a group of cohorts over time (such as the NLS series) there will be more years than cohorts.

in earnings slopes we estimate in Tables 2 and 4 incorporate *both* PSI effects and these year effects.

Next, we perform a series of simulations, all of them variants of the following experiment: Assume that a fraction  $\lambda$  of the time trend in Figure 5 is due to year effects, while the remainder is due to cohort effects. Use this assumption to create a trend-adjusted version of the relative wage variable, taking out both the assumed year-specific and cohort-specific trends.<sup>30</sup> Then estimate the  $pe*cohort$  interactions from the de-trended dependent variable. The results of this exercise are reported in Table 5, with the fraction of the trend attributed to year effects ( $\lambda$ ) varying from zero to one.

In column 1 of Table 5, the entire time trend is attributed to cohort effects.<sup>31</sup> In columns 2-5, a gradually increasing proportion of the time trend is attributed to year effects. As expected (recall that year effects were accelerating in the early part of our sample period), the estimated  $pe*cohort$  interaction falls for the 1937-1946 birth cohort as increasing importance is assigned to year effects. The effect is substantial. A significant drop is also apparent in the next entering cohort. For the youngest (1957-1967) cohort, which entered the labor market as the time trend began decelerating, the estimated  $pe*cohort$  interaction grows as year effects are introduced, but only slightly. Overall, as we simulate a larger demand shift towards female labor, the contribution of PSI effects (i.e. cohort-age interactions) to the narrowing of the gender wage gap tends to fall, but not dramatically (see Figure 6, or the last row of Table 5). For example, if half

---

<sup>30</sup> The exact specification can be found in Appendix 3.

<sup>31</sup> This differs slightly from the Table 4, Column 1 specification because the cohort fixed effects are replaced by the more constrained functional form described in Appendix 3. Similarly, the Column 5 result differs only slightly from that obtained with unconstrained year fixed effects.

of the trend is attributed to year effects, the contribution of cohort-age interactions falls from one-third to one-quarter.<sup>32</sup>

We conclude our discussion of Table 5 by noting two important senses in which the experience hypothesis and the declining discrimination hypothesis are truly complementary, rather than substitutes. First, adding some simulated year effects attenuates one of the most puzzling findings: the U-shape of the age-relative wage profile.<sup>33</sup> While it is possible to construct variants of a pure experience model in which women's wages grow more rapidly than men's throughout a substantial portion of the life cycle, it seems more natural to imagine that at least some of this empirical regularity has been due to a common time trend affecting women of all ages. Second, the introduction of even a small time trend induces the set of  $pe*cohort$  interaction terms to increase monotonically, as the experience model predicts it should (Table 5). Finally, we note that in the presence of such a time trend, the estimated contribution of women's rising post-schooling investments to the narrowing of the gender wage gap is unambiguously less than our "baseline" estimate; i.e. *less* than one-third.

In sum, both of the factors considered in this section --changing selection into work and pure time effects on women's relative wages-- provide reasons to suspect that our baseline estimates of the contributions of experience and other PSIs to the narrowing of the gender wage gap are overestimates. Once these factors are accounted for, the contribution of rising post-school investments (including experience) by women to the recent decline in the gender wage gap is likely less than the one third estimated earlier in the paper.

---

<sup>32</sup> This is not an exhaustive set of simulations. For example, suppose relative demand grows more quickly for women at the entry level, while older women remain on previously established career tracks, at least for a while. Then a portion of the time trend will be captured by cohort effects, rather than by changes in slope, and the original PSI estimates will contain less bias than they do in the cases we simulate.

<sup>33</sup> This is to be expected, as the simulations attribute some of the wage growth among older women in the baseline cohorts to the year effects.

## 6. Discussion

A widely cited potential explanation of the recent decline in the gender wage gap has focused on changes across cohorts in the rate at which women make post-schooling investments in their earnings capacity, such as accumulating work experience. In this paper we have assessed the contribution of changing post-school investments to the recent decline in the gender wage gap by decomposing the decline into components associated with the slopes versus levels of women's relative wage profiles across eight cohorts of women observed over the past 40 years.

In our cohort-based analysis of women's wage trends, we find that the gender gap does tend to widen during the earliest years of the career, but then actually *narrows* substantially during much of the life cycle for all cohorts of U.S. workers in our data. Some cross-cohort increases in women's relative rate of age-related wage growth are observed; taken together these increases can account for about one-third of the narrowing of the gender wage gap in a simple baseline model. When that model is expanded to allow for changing selection of women into work over time, and for pure time effects on women's relative wages, the likely contribution of rising post-school investments to the narrowing gender wage gap falls to under one third. Thus, while our analysis provides some support for the popular "experience"-based explanation of the decline in the U.S. gender wage gap, it also documents the existence of large, unexplained wage differences across cohorts that are already present at the start of women's working lives. Further, the reduction in these entry-level differentials accounts for the majority of the narrowing of the overall gender wage gap. What underlying factors might explain this larger portion of the change?

Obviously, one set of factors that might account for the large cohort effects in our data is unmeasured changes in *pre-market* investments. In other words, while our wage gap estimates

hold years of education constant, trends in the type or quality of human capital women bring to the labor market (Polachek 1978; Brown and Corcoran 1997; Weinberger 1998, 1999, 2001) could account for the large cohort effects we estimate here. Using information on the detailed college majors of a panel of college-educated workers of all ages from 1989 to 1999, Weinberger (2005) examines this hypothesis in a companion paper. Perhaps surprisingly, this analysis finds that controlling for detailed college major does little to attenuate the large cohort effects.<sup>34</sup> If no effect is found within a panel of college graduates, it seems unlikely that trends in unobserved pre-market human capital investments can account for much of the “unexplained” decline in the gender wage gap in the population of women as a whole.<sup>35</sup>

A second possible explanation of the decline in the gender gap at labor market entry extends the human capital model of equation (1) by allowing cross-cohort differences in *expected* labor force attachment to affect women’s entry-level earnings. Of course, in the standard general training model (e.g. Blau et al. 1997, chapter 6), this makes it even harder to explain the cohort effects we estimate here: controlling for pre-market investments (whose effects are discussed above), the standard model predicts that early-career wages should actually be *lower* for persons who expect to be more committed to the labor market, while we observe *rising* entry-level wages over time. For inter-cohort differences in expected labor force attachment to explain the large cohort effects in our data, one would thus need early career investments to take a different form from what is usually assumed. For example, suppose that -- rather than taking time away from production-- training investments take the form of increased hours or effort (beyond the level that would be optimal based on the worker’s current

---

<sup>34</sup> Another relevant finding of this paper is that the U-shape is even more pronounced, and swings upward earlier, when a fixed sample of college graduates is followed over time, compared to the corresponding synthetic cohort analysis.

<sup>35</sup> As mentioned earlier, the importance of cohort effects is surprisingly similar at different levels of educational attainment.

productivity alone). In contrast to the standard model, these factors should *raise* earnings during the training period. That said, we note that our main findings include detailed controls for work hours; thus we are skeptical that a model based on changing hours or effort across cohorts of young women can explain the large cohort effects in our data.<sup>36</sup>

Finally, of course, as noted there may have simply been a time trend in the relative price employers are willing to pay for female labor of a given level of education, training and expected future work attachment; it is hard to think of a label for such a trend other than declining discrimination. As shown in Section 5, such a time trend would appear in our baseline model as upward bias in our cohort effects. Thus, declining discrimination could also account for the large and declining estimated cohort effects in our baseline model. In this regard, we also note that, in two key senses, introducing *some* pure year effects actually improves the performance of the “experience” hypothesis in our data: First, (because year effects also bias our estimated *age* effects upward) a decline in discrimination can account for our (unexpected) finding that women’s wages actually rise more rapidly with age than men’s after middle age (about 47) in *every* cohort in our data. Second, when year effects are introduced, our estimated trend in age-cohort interactions becomes monotonically positive --as predicted by the experience model. In this important sense, then, the “experience” and “discrimination” hypotheses may be truly complementary, rather than substitutes, as possible explanations of the recent decline in the U.S. gender wage gap.

---

<sup>36</sup> An alternative modification to the basic model would be to make training firm-specific. For example, suppose, as in Kuhn (1993), that returns to specific training are shared between workers and firms, and that entry-level wages are determined by a zero-expected-profit condition for firms given each demographic group’s probability of remaining with the firm after training is complete. Now, because workers are paid some of their expected post-training productivity “up front,” an increase in the expected labor force attachment of a cohort of women can, under reasonable conditions, *raise* the starting wages of that cohort. While this is an important possibility, we note that it can only apply to firm-specific components of on-the-job training.

## References

- Blau, Francine D., Marianne Ferber, and Anne Winkler. The Economics of Women, Work and Pay (3<sup>rd</sup> edition) Upper Saddle River, NJ: Prentice-Hall, 1997.
- Blau, Francine D. and Lawrence M. Kahn. "Gender Differences in Pay". Journal of Economic Perspectives 14(4) (Fall 2000): 75-100.
- Blau, Francine D. and Lawrence M. Kahn . "The U.S. Gender Pay Gap in the 1990s". NBER working paper no. 10853, October 2004.
- Borjas, George. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants", Journal of Labor Economics 3 (1985): 463-489
- Brown, Charles and Mary Corcoran. "Sex-Based Differences in School Content and the Male-Female Wage Gap." Journal of Labor Economics 15 (1997):431-465.
- Deaton, Angus S. and Christina H. Paxson. "Saving, Growth and Aging in Taiwan". In David A. Wise, ed. Studies in the Economics of Aging. Chicago: University of Chicago Press, 1994, pp. 331-361.
- Goldin, Claudia. "Life-Cycle Labor Force Participation of Married Women: Historical Evidence and Implications". Journal of Labor Economics 7(1) (January 1989): 20-47.
- Kuhn, Peter. "Demographic Groups and Personnel Policy", Labour Economics 1 (June 1993): 49-70.
- Mulligan, C., and Y. Rubinstein. "Selection, Investment and Women's Relative Wages Since 1975". NBER working paper no. 11159, February 2005.
- Mincer, Jacob and Polachek, "Family Investments in Human Capital: Earnings of Women," Journal of Political Economy 82(2,part 2)(March-April 1974): s76-S108.
- O'Neill, June. "The Gender Gap in Wages, circa 2000". American Economic Review 93(2) (May 2003): 309-314.
- O'Neill, June and Solomon Polachek. "Why the Gender Gap in Wages Narrowed in the 1980's". Journal of Labor Economics 11(1) (Jan. 1993): 205-228.
- Polachek, S. "Potential Biases in Measuring Male-Female Discrimination," Journal of Human Resources, 10(2) (Spring 1975): 205-229.
- Polachek, S. "Sex Differences in College Major" Industrial and Labor Relations Review July 1978; 31(4): 498-508.

Weinberger, Catherine J. "Race and Gender Wage Gaps in the Market for Recent College Graduates." Industrial Relations 37(January, 1998):67-84.

Weinberger, Catherine J. 1999. "Mathematical College Majors and the Gender Gap in Wages." Industrial Relations 38 (July, 1999):407-13.

Weinberger, Catherine J. "Is Teaching More Girls More Math the Key to Higher Wages?" in Squaring Up: Policy Strategies to Raise Women's Incomes in the U.S., edited by Mary C. King. Forthcoming June 2001, University of Michigan Press.

Weinberger, Catherine J. 2005. "In Search of the Glass Ceiling: Cohort Effects among U.S. College Graduates in the 1990's" UCSB Working Paper.

Weiss, Y. and R. Gronau. "Expected Interruptions in Labour Force Participation and Sex-Related Differences in Earnings Growth" Review of Economic Studies 48(4) (October 1981): 607-619.

**Table 1: Gender Earnings Gaps by Cohort and Year, Cross-sectional Census Data Regressions**

|                           | (1)       | (2)       | (3)       | (4)       | (5)       |
|---------------------------|-----------|-----------|-----------|-----------|-----------|
| Year                      | 1959      | 1969      | 1979      | 1989      | 1999      |
| <b>female*(age 23-32)</b> | -0.432    | -0.424    | -0.329    | -0.216    | -0.198    |
|                           | (0.004)** | (0.004)** | (0.002)** | (0.002)** | (0.003)** |
| <b>female*(age 33-42)</b> | -0.558    | -0.571    | -0.509    | -0.350    | -0.263    |
|                           | (0.003)** | (0.004)** | (0.003)** | (0.002)** | (0.002)** |
| <b>female*(age 43-52)</b> | -0.595    | -0.600    | -0.587    | -0.460    | -0.334    |
|                           | (0.003)** | (0.003)** | (0.003)** | (0.003)** | (0.003)** |
| <b>female*(age 53-62)</b> | -0.560    | -0.539    | -0.537    | -0.458    | -0.336    |
|                           | (0.005)** | (0.004)** | (0.004)** | (0.004)** | (0.004)** |
| Age                       | 0.026     | 0.029     | 0.040     | 0.041     | 0.039     |
|                           | (0.000)** | (0.000)** | (0.000)** | (0.000)** | (0.000)** |
| (age-22) squared          | -0.000    | -0.000    | -0.001    | -0.001    | -0.001    |
|                           | (0.000)** | (0.000)** | (0.000)** | (0.000)** | (0.000)** |
| 35-39 hours/week          | 0.018     | 0.007     | -0.069    | -0.121    | -0.158    |
|                           | (0.003)** | (0.003)** | (0.003)** | (0.003)** | (0.004)** |
| 41-48 hours/week          | -0.015    | 0.050     | 0.089     | 0.121     | 0.133     |
|                           | (0.002)** | (0.002)** | (0.002)** | (0.002)** | (0.002)** |
| 49+ hours/week            | -0.059    | 0.049     | 0.115     | 0.188     | 0.245     |
|                           | (0.003)** | (0.003)** | (0.002)** | (0.002)** | (0.002)** |
| Observations              | 238937    | 296103    | 393118    | 497703    | 538469    |
| R-squared                 | 0.33      | 0.35      | 0.32      | 0.34      | 0.32      |

Robust standard errors in parentheses

\* significant at 5%; \*\* significant at 1%

Other controls: census region, education level

**Table 2- Testing alternative models of women's relative earnings**

|                                    | (1)                 | (2)                | (3)                 |
|------------------------------------|---------------------|--------------------|---------------------|
| Cohort 2<br>(born 1907-1916)       | 0.034<br>(0.031)    | 0.037<br>(0.049)   |                     |
| Cohort 3<br>(born 1917-1926)       | 0.042<br>(0.030)    | 0.054<br>(0.048)   |                     |
| Cohort 4<br>(born 1927-1936)       | 0.069<br>(0.031)    | 0.100<br>(0.047)   |                     |
| Cohort 5<br>(born 1937-1946)       | 0.073<br>(0.042)    | 0.179<br>(0.047)** |                     |
| Cohort 6<br>(born 1947-1956)       | 0.183<br>(0.043)**  | 0.291<br>(0.051)** |                     |
| Cohort 7<br>(born 1957-1966)       | 0.287<br>(0.045)**  | 0.380<br>(0.055)** |                     |
| Cohort 8<br>(born 1967-1976)       | 0.305<br>(0.045)**  | 0.390<br>(0.063)** |                     |
| Potential Experience<br>* Cohort 5 | 0.051<br>(0.014)*   |                    | 0.063<br>(0.021)**  |
| Potential Experience<br>* Cohort 6 | 0.067<br>(0.022)*   |                    | 0.146<br>(0.035)**  |
| Potential Experience<br>* Cohort 7 | 0.068<br>(0.041)    |                    | 0.286<br>(0.077)**  |
| Age 33-42                          | -0.115<br>(0.022)** | -0.063<br>(0.027)* | -0.229<br>(0.047)** |
| Age 43-52                          | -0.139<br>(0.027)** | -0.060<br>(0.029)  | -0.279<br>(0.046)** |
| Age 53-62                          | -0.057<br>(0.029)   | 0.028<br>(0.031)   | -0.204<br>(0.045)** |
|                                    |                     |                    |                     |
| R squared                          | 0.99                | 0.96               | 0.81                |
| Adjusted R squared                 | 0.96                | 0.91               | 0.72                |

Sample size for all regressions is 20 age-year cells estimated in Table 1. Omitted categories are Age 23-32, and Cohort 1 (born 1897-1906). Potential experience is measured as *decades* elapsed since the cohort was aged 23-32. Cohort-specific earnings growth rates are estimated relative to women born before 1937, and cannot be estimated for Cohort 8 since we have only one year of data for this cohort.

**Table 3: Contribution of Changing Slopes to Changes in Gender Earnings Gaps  
(from Cross-sectional Census Data Regressions )**

|                    | (1)              | (2)              | (3)                                                                                      |
|--------------------|------------------|------------------|------------------------------------------------------------------------------------------|
| Dependent Variable | 1959<br>Earnings | 1999<br>Earnings | 1999<br>Earnings<br>MINUS<br>Estimated<br>Effects of<br>Changing<br>Slopes <sup>37</sup> |
| <b>female</b>      | -0.540           | -0.278           | -0.361                                                                                   |
|                    | (0.002)**        | (0.002)**        | (0.002)**                                                                                |
| Age                | 0.023            | 0.035            | 0.031                                                                                    |
|                    | (0.000)**        | (0.000)**        | (0.000)**                                                                                |
| (age-22) squared   | -0.000           | -0.001           | -0.001                                                                                   |
|                    | (0.000)**        | (0.000)**        | (0.000)**                                                                                |
| 35-39 hours/week   | 0.019            | -0.161           | -0.163                                                                                   |
|                    | (0.003)**        | (0.004)**        | (0.004)**                                                                                |
| 41-48 hours/week   | -0.017           | 0.133            | 0.133                                                                                    |
|                    | (0.002)**        | (0.002)**        | (0.002)**                                                                                |
| 49+ hours/week     | -0.061           | 0.244            | 0.243                                                                                    |
|                    | (0.003)**        | (0.002)**        | (0.002)**                                                                                |
| Observations       | 238937           | 538469           | 538469                                                                                   |
| R-squared          | 0.33             | 0.32             | 0.34                                                                                     |

Robust standard errors in parentheses

\* significant at 5%; \*\* significant at 1%

Other controls: census region, education level

<sup>37</sup> Estimated Effects of Changing Slopes are based on Table 2, column 1 estimates of *pe\*cohort* interactions: zero for men and for the youngest women, .068 for women age 33-42, .067\*2 for women age 43-52, and .051\*3 for women age 53-62.

**Table 4—Robustness to Changes in Data Source and Cell-Size**

|                                                           | (1)                 | (2)                 | (3)                 | (4)                 | (5)                |
|-----------------------------------------------------------|---------------------|---------------------|---------------------|---------------------|--------------------|
| Data Source                                               | Census              | CPS                 | CPS                 | CPS                 | CPS                |
| Number of Cohort Controls                                 | 8                   | 8                   | 8                   | 16                  | 80                 |
| Number of Age Group Controls                              | 4                   | 4                   | 4                   | 8                   | 40                 |
| Potential Experience * (born 1937-1946)                   | 0.055<br>(0.016)*   | 0.062<br>(0.020)*   | 0.066<br>(0.007)**  | 0.051<br>(0.005)**  | 0.044<br>(0.003)** |
| Potential Experience * (born 1947-1956)                   | 0.073<br>(0.024)*   | 0.065<br>(0.030)    | 0.068<br>(0.011)**  | 0.066<br>(0.008)**  | 0.065<br>(0.005)** |
| Potential Experience * (born 1957-1966)                   | 0.073<br>(0.044)    | 0.038<br>(0.055)    | 0.046<br>(0.025)    | 0.053<br>(0.016)**  | 0.068<br>(0.009)** |
| Potential Experience *(born 1967-1976)                    |                     |                     |                     | 0.003<br>(0.077)    | 0.015<br>(0.030)   |
| Age 28-32                                                 |                     |                     |                     | -0.047<br>(0.012)** |                    |
| Age 33-42 (33-37 in Col 4)                                | -0.131<br>(0.025)** | -0.118<br>(0.031)** | -0.133<br>(0.012)** | -0.121<br>(0.012)** |                    |
| Age 38-42                                                 |                     |                     |                     | -0.150<br>(0.013)** |                    |
| Age 43-52 (43-37 in Col 4)                                | -0.158<br>(0.030)** | -0.142<br>(0.037)** | -0.148<br>(0.013)** | -0.169<br>(0.014)** |                    |
| Age 48-52                                                 |                     |                     |                     | -0.133<br>(0.014)** |                    |
| Age 53-62 (53-57 in Col 4)                                | -0.069<br>(0.031)   | -0.068<br>(0.039)   | -0.061<br>(0.015)** | -0.085<br>(0.014)** |                    |
| Age 58-62                                                 |                     |                     |                     | -0.011<br>(0.015)   |                    |
| Observations                                              | 20                  | 20                  | 128                 | 296                 | 1640               |
| R squared                                                 | 0.99                | 0.98                | 0.97                | 0.95                | 0.90               |
| Adjusted R squared                                        | 0.96                | 0.93                | 0.96                | 0.95                | 0.89               |
| Counterfactual Log Wage Differential (LWD)                | -.365               | -.352               | -.358               | -.364               | -.378              |
| Proportion of Change in Gender Gap Due to Changing Slopes | .34                 | .29                 | .31                 | .33                 | .39                |

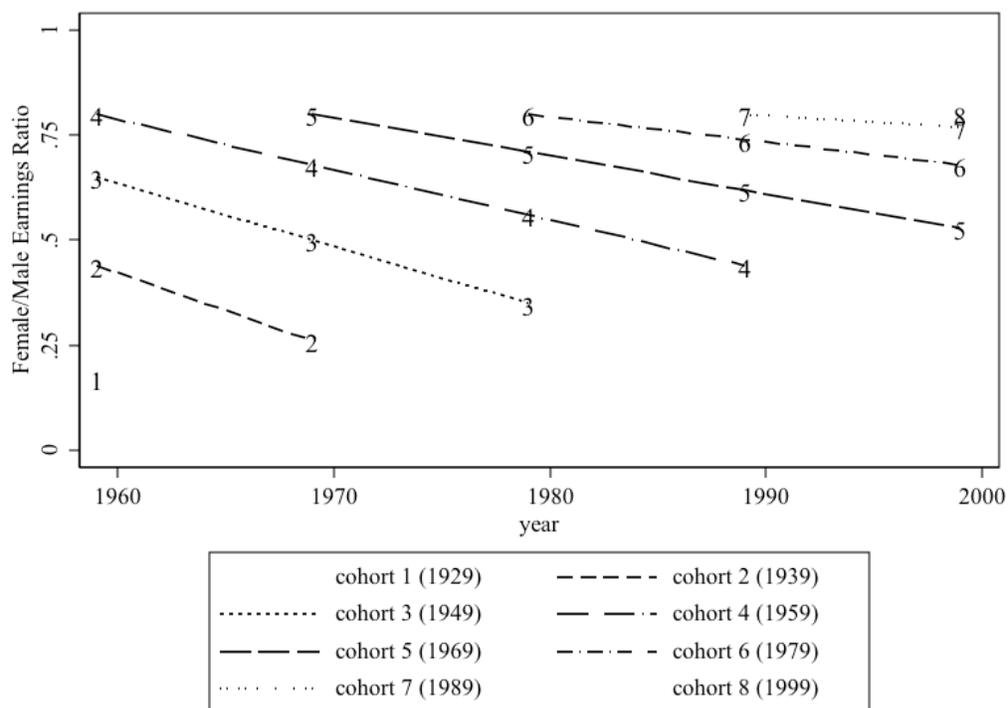
Column 1 & 2 regressions use data from the 5 Census years, columns 3-5 include annual data spanning the same four decades. In columns 1-4 the youngest age group is the omitted category. Potential experience is measured in *decades*. Cohort-specific earnings growth rates are estimated relative to women born before 1937. Last row computed relative to actual mean LWD from male coefficients: -.277 in 1999 and -.539 in 1959.

**Table 5—Simulated Experience Effects Under Alternative Assumptions about Year Effects**

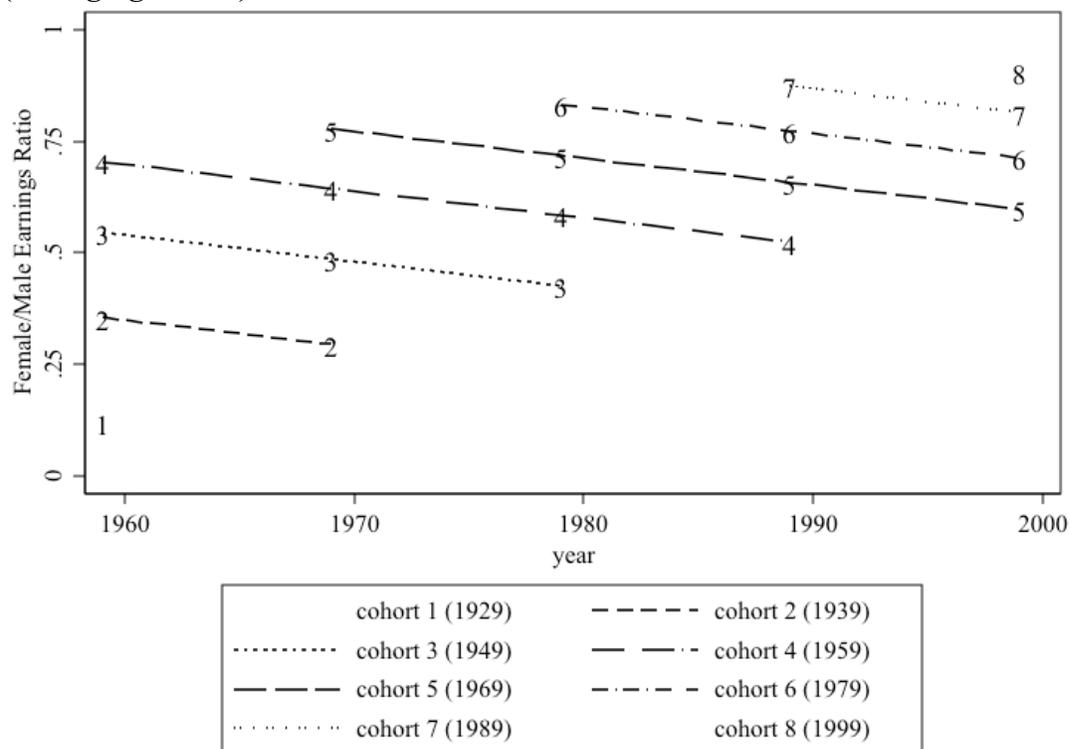
|                                                                                                             | (1)            | (2)            | (3)            | (4)            | (5)            |
|-------------------------------------------------------------------------------------------------------------|----------------|----------------|----------------|----------------|----------------|
| <b>Experimental Treatment:</b><br>Proportion of Entry-Level<br>Narrowing Attributed to<br>Pure Year Effects | <b>//=0.00</b> | <b>//=0.25</b> | <b>//=0.50</b> | <b>//=0.75</b> | <b>//=1.00</b> |
| Potential Experience                                                                                        | 0.053          | 0.040          | 0.028          | 0.016          | 0.004          |
| * (born 1937-1946)                                                                                          | (0.007)**      | (0.005)**      | (0.006)**      | (0.008)        | (0.011)        |
| Potential Experience                                                                                        | 0.070          | 0.062          | 0.054          | 0.046          | 0.038          |
| * (born 1947-1956)                                                                                          | (0.011)**      | (0.008)**      | (0.009)**      | (0.013)**      | (0.017)*       |
| Potential Experience                                                                                        | 0.071          | 0.074          | 0.077          | 0.080          | 0.083          |
| * (born 1957-1966)                                                                                          | (0.020)**      | (0.015)**      | (0.017)**      | (0.024)**      | (0.032)*       |
| Age 33-42                                                                                                   | -0.141         | -0.147         | -0.154         | -0.160         | -0.166         |
|                                                                                                             | (0.018)**      | (0.014)**      | (0.015)**      | (0.021)**      | (0.029)**      |
| Age 43-52                                                                                                   | -0.170         | -0.182         | -0.193         | -0.205         | -0.216         |
|                                                                                                             | (0.018)**      | (0.014)**      | (0.015)**      | (0.021)**      | (0.029)**      |
| Age 53-62                                                                                                   | -0.091         | -0.109         | -0.127         | -0.145         | -0.163         |
|                                                                                                             | (0.018)**      | (0.014)**      | (0.015)**      | (0.021)**      | (0.028)**      |
| R-squared                                                                                                   | 0.92           | 0.95           | 0.94           | 0.90           | 0.85           |
| Adjusted R squared                                                                                          | 0.89           | 0.93           | 0.91           | 0.86           | 0.78           |
| Counterfactual Log Wage<br>Differential (LWD)                                                               | -.362          | -.353          | -.344          | -.335          | -.327          |
| Simulated Proportion of<br>Change in Gender Gap<br>Due to Changing Slopes<br>(net of demand shifts)         | .32            | .29            | .26            | .22            | .19            |

Data: 20 Census observations, trend-adjusted as described in Appendix 3.

**Figure 1: Hypothetical Female Relative Wage Profiles: Pure Experience Model (Changing Slopes)**

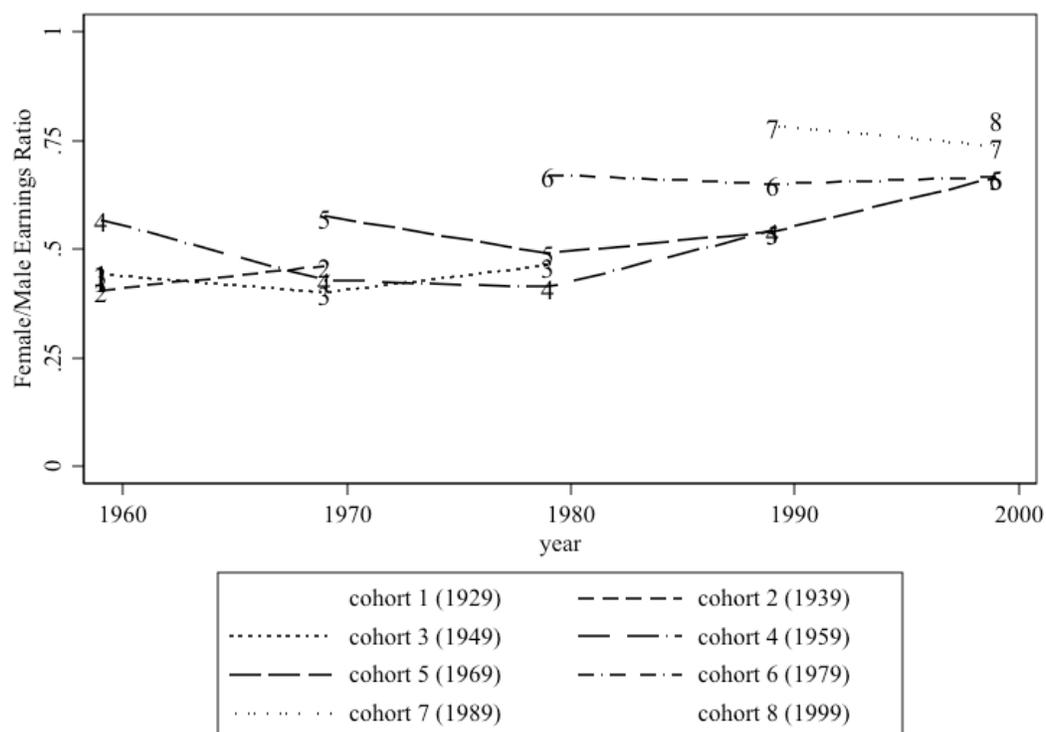


**Figure 2: Hypothetical Female Relative Wage Profiles: Pure Cohort Model (Changing Levels)**

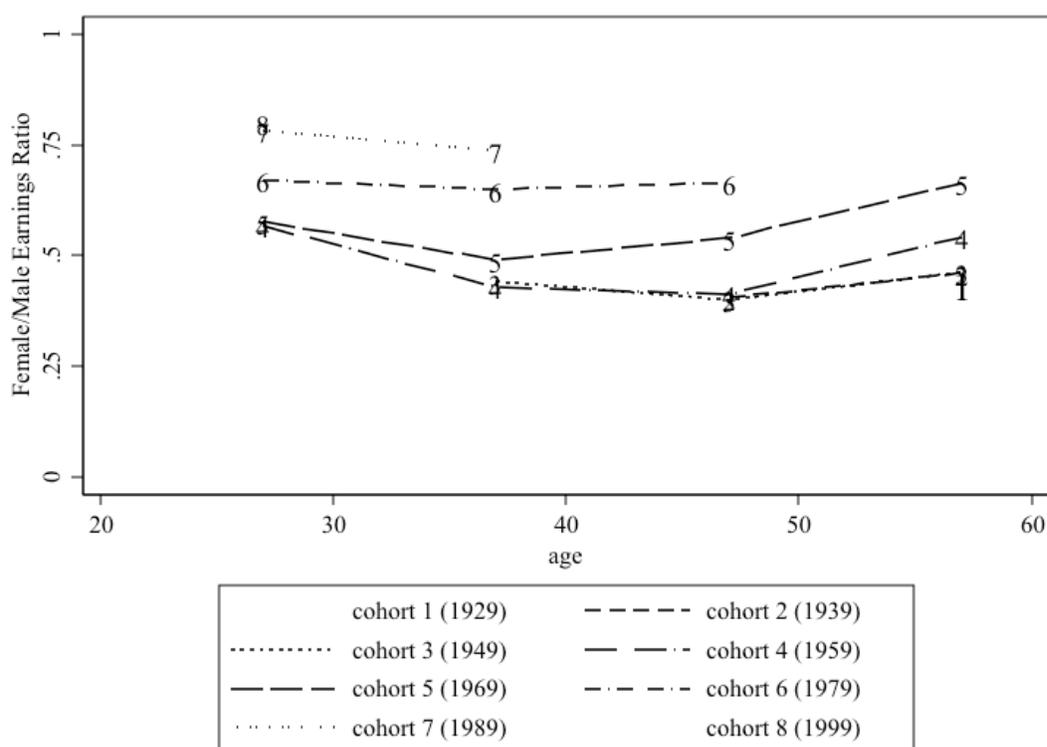


Note: Cohorts are labeled by the year in which their median age was 27

**Figure 3: Actual Female Relative Wage Profiles, by year and cohort**

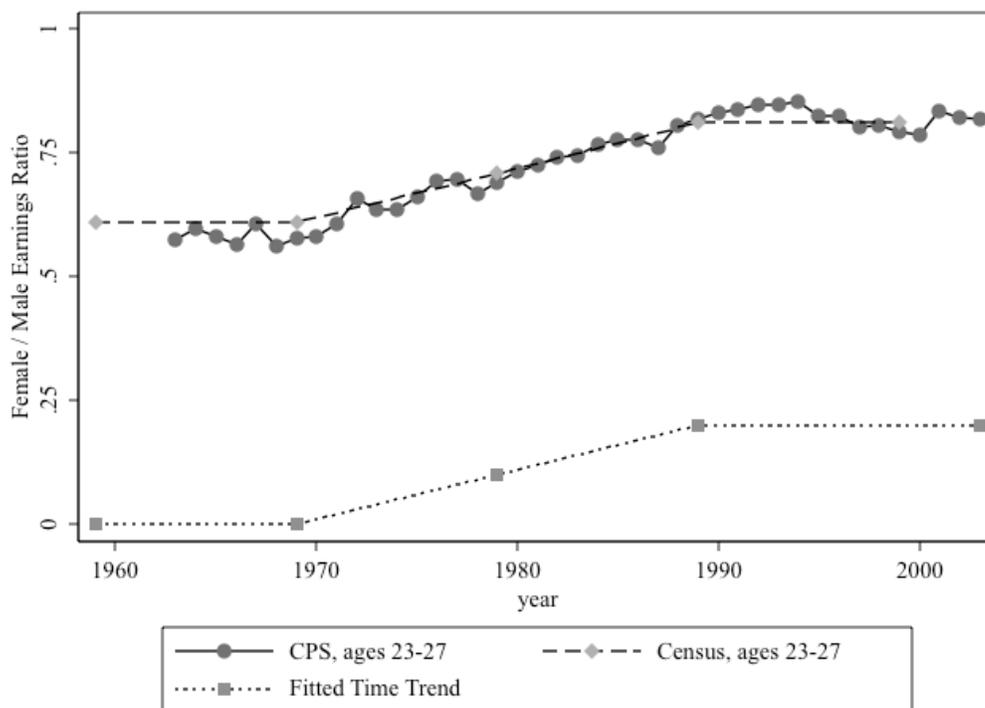


**Figure 4: Actual Female Relative Wage Profiles, by age and cohort**



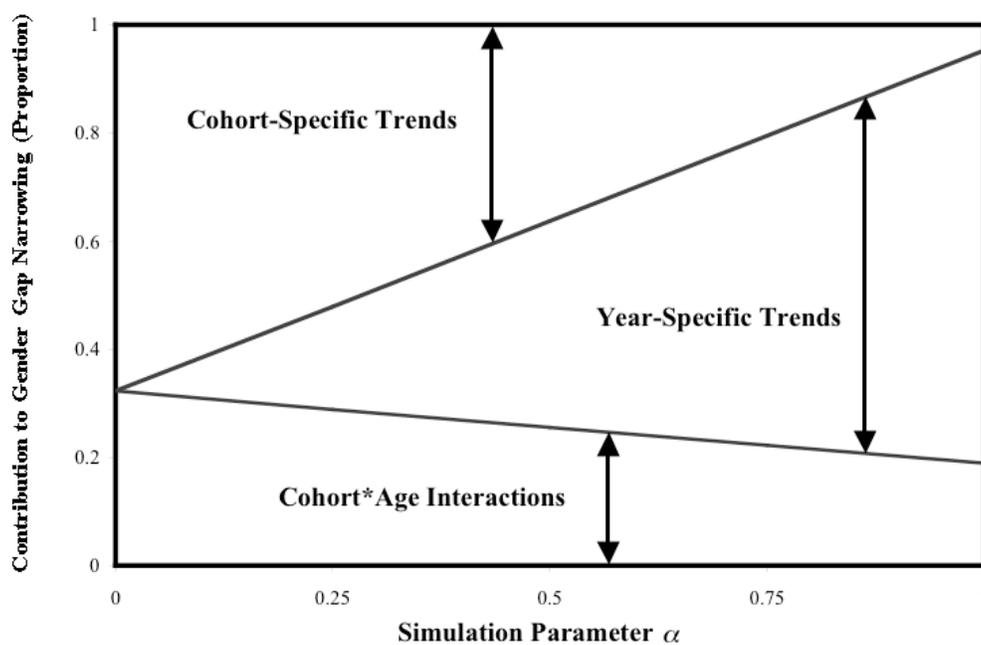
Note: Cohorts are labeled by the year in which their median age was 27.

**Figure 5: Relative Wages at Labor Market Entry, and Simple Fitted Time Trend.**



Note: The fitted time trend is a piecewise linear function, with kinks at 1969 and 1989.

**Figure 6: Sensitivity of Estimated Cohort and Cohort\*Age Effects to Simulated Time Trend in the Demand for Women**



Note: Simulations are described, and  $\alpha$  is defined, in Appendix 3.

## APPENDIX 1:

## Sample means of selected Census variables, by gender (standard deviation in parentheses)

| Year                                            | 1959           | 1969            | 1979             | 1989             | 1999             |
|-------------------------------------------------|----------------|-----------------|------------------|------------------|------------------|
| <b>All Men (age 23-62):</b>                     |                |                 |                  |                  |                  |
| Proportion employed full time, full year        | 0.64           | 0.67            | 0.67             | 0.67             | 0.67             |
| N                                               | 342443         | 365602          | 433780           | 500352           | 519909           |
| <b>All Women (age 23-62):</b>                   |                |                 |                  |                  |                  |
| Proportion employed full time, full year        | 0.17           | 0.21            | 0.29             | 0.39             | 0.43             |
| N                                               | 355566         | 385058          | 448945           | 516213           | 531392           |
| <b>Men employed full time, full year:</b>       |                |                 |                  |                  |                  |
| Income                                          | 6200<br>(3200) | 10100<br>(5800) | 20600<br>(11500) | 35100<br>(22900) | 48800<br>(32800) |
| Age                                             | 40<br>(10)     | 41<br>(11)      | 39<br>(11)       | 39<br>(10)       | 41<br>(10)       |
| Hours < 40                                      | .04            | .05             | .04              | .03              | .02              |
| Hours > 50                                      | .22            | .23             | .22              | .29              | .34              |
| College graduate                                | .13            | .16             | .26              | .29              | .31              |
| N                                               | 183797         | 220156          | 267601           | 309713           | 321517           |
| <b>Women employed full time, full year:</b>     |                |                 |                  |                  |                  |
| Income                                          | 3600<br>(1500) | 5600<br>(2800)  | 11700<br>(5800)  | 22300<br>(13200) | 33700<br>(22000) |
| Age                                             | 42<br>(10)     | 43<br>(11)      | 39<br>(12)       | 39<br>(10)       | 41<br>(10)       |
| Hours < 40                                      | .16            | .16             | .15              | .12              | .10              |
| Hours > 50                                      | .05            | .05             | .06              | .11              | .16              |
| College graduate                                | .07            | .09             | .16              | .23              | .29              |
| N                                               | 55140          | 75947           | 125517           | 188010           | 217073           |
| <b>Census Gender Gap from Male Coefficients</b> |                |                 |                  |                  |                  |
| CPS Gender Gap from Male Coefficients           | (1963): -0.54  | -0.54           | -0.48            | -0.34            | -0.30            |

## APPENDIX 2: Introducing Year Effects

In order to examine the effects of introducing a time trend into our estimated model, we first generalize equations (2) and (3) to let post-schooling investments (PSIs) be a cumulative, non-linear, gender- and cohort-specific function of potential labor market experience. To that end, assume the following for men:

$$y(e) = Y^0 + \int_0^e PSI^M(a) da, \quad (2)$$

and for women of cohort  $c$ :

$$y_c(e) = Y^0 + G_c + \int_0^e PSI_c^F(a) da, \quad (3)$$

implying the relative wage function:

$$RW(e, c) = 1 + G_c + \int_0^e (PSI_c^M(a) - PSI^F(a)) da \quad (4)$$

Letting  $c=0$  be the reference cohort, it follows that:

$$RW(e, c) = (1 + G_0) + (G_0 - G_c) + \int_0^e (PSI_0^M(a) - PSI_0^F(a)) da + \int_0^e (PSI_c^F(a) - PSI_0^F(a)) da \quad (5)$$

The first term,  $(1-G_0)$ , in equation 5 corresponds to the intercept,  $\beta$ , in our estimating equation, (6). The second term,  $(G_0 - G_c)$ , corresponds to the cohort effects ( $\gamma_c$ ) in (6); the third to the “baseline” age effects on women’s relative wages. Finally, the parameters  $\lambda_c$  in equation (6) approximate the average value (over all ages) of  $(PSI_c^F(a) - PSI_0^F(a))$ , i.e. the cohort-age interactions at the heart of the “experience” hypothesis. This allows us to test the argument underlying the experience hypothesis, i.e. that  $PSI_c^F(a)$  is higher, on average, for recent cohorts of women, relative to earlier cohorts at the same age.

We are now in a position to add a time trend,  $trend(t)$ , to equation (5'), and to explore its impact on our estimates of PSI effects. To do this, we normalize units of measurement so that time ( $t$ ) is measured in decades, and  $e+c=t$ . After adding the trend, equation 5 becomes:

$$RW(e,c) = (1 - G_0) + (G_0 - G_c) \int_0^e (PSI_0^M(a) - PSI_0^F(a)) da + \int_0^e (PSI_c^F(a) - PSI_0^F(a)) da + trend(e+c) \quad (A1)$$

Suppose first that the trend is a linear function, i.e. ( $trend(t)=bt=be+bc$ ), with  $b > 0$ . In this case, (A1) can be rewritten:

$$RW(e,c) = (1 - G_0) + (G_0 - G_c + bc) + be \int_0^e (PSI_0^M(a) - PSI_0^F(a)) da + \int_0^e (PSI_c^F(a) - PSI_0^F(a)) da \quad (A2)$$

Thus, our estimated cohort effects,  $(G_0 - G_c + bc)$ , will overstate the true cohort effects,  $(G_0 - G_c)$ .

Likewise, the estimated baseline age effect,  $be \int_0^e (PSI_0^M(a) - PSI_0^F(a)) da$ , will understate

the extent to which women's PSI accumulation tends to fall behind men's,

$\int_0^e (PSI_0^M(a) - PSI_0^F(a)) da$ . Critically, however, a linear trend will not affect our estimated

PSI effects ( $pe*cohort$  interactions) at all. In sum, the presence of a true, positive linear time trend (obviously) affects the relative role of time versus cohort effects in explaining the

narrowing of the gender wage gap (raising the importance of time and reducing that of cohort).

It also biases our estimated age coefficients upward relative their true values. That said, the

presence of such a time trend leaves the total relative wage change attributable to pure time and

cohort effects *together* unchanged, and leaves the contribution of cohort-year interactions *relative* to these two alternative mechanisms unchanged as well.

Finally, assume a more general case where the trend is not necessarily linear. The effect of any monotonically increasing trend can be seen most clearly if we decompose the trend in the following way:

$$trend(e + c) = trend(c) + trend(e) + [trend(e + c) - trend(c) - trend(e)] \quad (A3)$$

Where  $trend(c)$  represents the value of the trend at cohort  $c$  entry, and  $trend(e)$  represents the value of the trend when the baseline cohort reaches age  $e$ , relative to the value at  $t=0$ .

Then:

$$RW(e, c) = (1 - G_0) + (G_0 - G_c + trend(c)) + \int_0^e trend(e) \int_0^e [PSI_0^M(a) - PSI_0^F(a)] da + \int_0^e [trend(e + c) - trend(c) - trend(e)] + \int_0^e [PSI_c^F(a) - PSI_0^F(a)] da \quad (A4)$$

Again, both cohort and baseline age effects are upward biased. But here, the estimated PSI effects ( $pe * cohort$  interactions) will be biased when the trend is not linear —upward when the trend is accelerating, and downward when the trend is decelerating.<sup>38</sup>

<sup>38</sup> To see this simply recall that acceleration means that  $f(x + \Delta) > f(x) + f(\Delta)$ .

### APPENDIX 3: Simulations

According to Appendix 2, our estimates of  $pe*cohort$  interactions in Tables 2 and 4 are biased upward when there is an accelerating time trend in women's relative wages, and downward when the trend is decelerating. Further, the data on entry-level wages in Figure 5 shows acceleration in the early years of our data and deceleration later. Unless this trend is purely cohort-specific, it follows that our  $pe*cohort$  interactions are biased upward for old cohorts, and downward for more recent cohorts.

The simulations presented in Table 5 estimate the size of this bias under a range of hypotheses about the nature of the true time trend. Based on the empirically observed time trend among entry-level workers (depicted in Figure 5), define:

$$T(t) = \begin{cases} 0.0 & | \text{year} \in \{1959, 1969\} \\ 0.1 & | \text{year} = 1979 \\ 0.2 & | \text{year} \in \{1989, 1999\} \end{cases}$$

Based on the relative wage equation (A1), evaluated at  $e=0$ ,  $T(c)$  approximates the time trend in relative wages at labor market entry:

$$T(c) \approx RW(0,c) - RW(0,0) = (G_0 - G_c) + trend(c) \quad (A5)$$

Since it is impossible to identify the relative contributions of true cohort-specific effects ( $G_0 - G_c$ ) and year-specific time trends ( $trend(c)$ ) to  $T(t)$ , we vary the relative importance of each in a series of simulations. Suppose that (for some value of  $\lambda \in [0,1]$ ):

$$(G_0 - G_c) = (1 - \lambda)T(c) \quad (A6)$$

Then:

$$trend(c) \approx \lambda T(c) \quad (A7)$$

In this specification,  $\lambda$  describes the proportion of  $T(c)$  that is due to year effects, rather than cohort effects.

Returning to relative wage equation (A1), the dependent variable for observations at all ages can be "de-trended" using the following transformation:

$$\begin{aligned}
 RW(e,c) - trend(e+c) - (G_0 - G_c) &= RW(e,c) - \lambda T(e+c) - (1-\lambda)T(c) \\
 &= (1-\lambda)G_0 - \int_0^e (PSI_0^M(a) - PSI_0^F(a)) da + \int_0^e (PSI_c^F(a) - PSI_0^F(a)) da
 \end{aligned} \tag{A8}$$

If the correct value of  $\lambda$  is known, regression of this de-trended dependent variable on age controls and  $pe$ \*cohort interactions will yield unbiased estimates of age and PSI effects. Of course, we do not know the true value of  $\lambda$ . The best we can do is to see how the PSI estimates (and age-earnings profiles) change as different values of  $\lambda$  are simulated. Columns 1-5 of Table 5 report the results of simulations where  $\lambda$  takes values 0, 0.25, 0.50, 0.75, and 1.