Dads and Daughters: The Changing Impact of Fathers on Women's Occupational Choices*

Judith K. Hellerstein

Department of Economics and MPRC
University of Maryland, College Park
and NBER
Email: hellerst@econ.umd.edu

Melinda Sandler Morrill

Department of Economics
North Carolina State University
Email: melinda_morrill@ncsu.edu

This version: April 22, 2010

*We are grateful for the comments provided by Seth Sanders, Mark Duggan, Suzanne Bianchi, Bill Evans, Jonah Gelbach, Julie Zissimopolous, two anonymous referees, and seminar participants at the University of Maryland, the University of Wisconsin, the NBER, the PAA, the University of Rochester, the University of Kentucky, the University of Arizona, and the Federal Reserve Bank of New York. We are especially grateful for the efforts of Megan Brooks early in this project. Morrill acknowledges the support of the Training Program at the Maryland Population Research Center.

Abstract

We examine whether women's rising labor force participation led to increased intergenerational transmission of occupation from fathers to daughters. In order to motivate our empirical strategy, we develop a model of investment of fathers in daughters, where fathers invest in human capital that is specific to their occupations. Our model generates an empirical test where we compare the trends in the probabilities that a woman works in her father's versus her father-in-law's occupation. Because of assortative mating, the father-in-law's occupation serves as a kind of counterfactual for the set of occupations that each particular woman might have chosen absent transmission from her father. Using data from birth cohorts of women born between 1909 and 1977, our results indicate that the estimated difference in these trends accounts for at least 13 to 20 percent of the total increase in the probability that a woman enters her father's occupation. Thus, over the birth cohorts studied, daughters are increasingly likely to enter their fathers' occupation, above and beyond what would have been expected from the secular rise in women entering traditionally male-dominated occupations.

1 Introduction

Over the last century, the labor force participation of women in the United States has risen threefold.[1] In addition, there has been a tremendous increase in the amount of integration of women in the labor market, so that women are far more likely now to work with men than in previous generations. Although the exact mechanisms for these changes remain somewhat elusive,[2] the fact that more women enter the labor market now and work in the same occupations as men has profound implications for many dimensions of the U.S. economy.

Interestingly, there is virtually no previous research examining how rising labor force participation and labor market integration of women have affected intergenerational transmission to daughters. In this paper, we examine changing intergenerational transmission from fathers to daughters by focusing on one key dimension of this change: the increasing probability across birth cohorts that a woman enters her father's occupation. The first contribution of this paper is to document the steady and large rise in the United States across birth cohorts of the 20th

[1] See, for example, Goldin (1991) for a detailed analysis of the U.S. experience. Rising female labor force participation is not a phenomenon unique to the United States of course. Mincer (1985) provides early evidence of this across twelve industrialized countries.

[2] See, for example, Acemoglu, Autor, and Lyle (2004) and Goldin and Katz (2002).

century in the probability that a woman works in her father's occupation. This suggests that the occupations of women's fathers may have played an increasingly important role in determining women's occupational choices, but does not directly imply that the transmission of what we call "occupation-specific human capital" has increased over this period. Because a woman born in a recent cohort is more likely to work in any traditionally male-dominated occupations relative to a woman in an older cohort, she is more likely than a woman born in an earlier cohort to enter her father's occupation even absent any changes in the transmission of occupation specific human capital from fathers to daughters. As a result, in order to provide evidence that intergenerational transmission of occupation-specific human capital between fathers and daughters really has increased, we demonstrate empirically that the positive trend in the probability that a woman works in her father's occupation is larger than that which would be predicted just from the fact that women are more likely to enter men's occupations more generally.

The difficulty in demonstrating this is in figuring out the counterfactual for a given woman: What occupations might the woman have entered absent any transmission of occupation-specific human capital from her father? Clearly, for any given woman, some occupations are likely to be closer substitutes than others (and, in particular, closer substitutes to her father's occupation). Moreover, this set of close substitutes may differ across women and across birth cohorts in ways that would require us to place some structure on the substitutability of occupations.

Consider the case of a woman whose father is a doctor. If her father is a country doctor who treats (and lives among) many farmers, the daughter may be more likely all else equal to become a farmer (versus, say, a college professor) because her father's labor market and social network puts her in contact with farmers. In contrast, the daughter of a city doctor who treats college professors is more likely to become a college professor than a farmer. But for both daughters, if the transmission of her father's occupation-specific human capital increases, she is more likely to become a doctor than to become either a farmer or a college professor.

Of course no data set is going to contain information on the relevant alternative occupations for any given woman. In the end, then, we draw on an identification assumption about assortative mating in order to identify the set of alternative occupations that a woman might enter, thereby distinguishing the country doctor's daughter from the city doctor's daughter. The basic idea is that because of assortative mating, a woman's father-in-law is likely to be working in the set of occupations that a woman might choose to work in, given her general human capital and given her preferences–preferences that may have been shaped by her social network growing up, her father's labor market network, her religious background, etc. To complete the example, the daughter of the country doctor is more likely to marry the son of a farmer than the son of a college professor, whereas the opposite is true for the daughter of the city doctor. In a sense, the father-in-law is like a "counterfactual" father. He represents the same male generational influences of that time period and, because of assortative mating, is from a similar social class or social group. However, the father-in-law generally does not directly transmit human capital to his daughter-in-law. Using the father-in-law's occupation as a comparison group allows us to model how the probability a woman enters her father's occupation changes over time, accounting for the secular trends in women's occupational upgrading and increased labor force participation.

In order to formalize these ideas, we develop a model that combines features of intergenerational job-specific human capital transmission with an occupational choice model. The comparative statics of the model explicitly motivate the use of the information on the occupation of a woman's father-in-law to generate an empirical test of whether daughters have become more likely to enter their fathers' occupations, conditional on the general economic forces that have led to women increasingly entering men's occupations. The resulting empirical test compares the rate of increase in the probability that a woman works in her father's occupation to the rate of increase in the probability that a woman works in her father-in-law's occupation. Within the framework of the model we also are able to examine the empirical implications that arise if assortative mating is not perfect across fathers' occupations. We demonstrate that as long as assortative mating by occupation has not gone down across subsequent birth cohorts, our results are robust. In our empirical work, we confirm that assortative mating has not decreased by showing that there has been no change in the probability that a woman's husband works in the same occupation as her father (his father-in-law).

In our model, we follow the common tradition in economics of treating intergenerational human capital transmission as arising from parental investments in children. We show, however, that similar empirical predictions would arise from a model of purely increased returns to heredity. As a result, in our empirical work and in most of our discussion outside of the model, we simply refer to the "transmission" of human capital between fathers and daughters, be it via investments or heredity.

Using data from the General Social Survey (GSS), the Survey of Income and Program Participation (SIPP), and the Occupational Changes in a Generation (OCG), we document changes in occupation-specific human capital transmission between fathers and daughters spanning birth cohorts from 1909 to 1977.[3] There is clear evidence in the data of an increase in the probability that a woman works in her father's occupation over time. For example, with the baseline definition of occupation that we use, just under 6 percent of women born in our earliest birth cohorts work in their fathers' occupations, while around 20 percent of women born most recently work in their fathers' occupations. We also document an increase over birth cohort in the fraction of women working in the same occupations as their fathers-in-law. However, the increase in the probability over birth cohorts that a woman works in her father's occupation is larger than the increase in the probability that she works in her father-in-law's occupation, and the difference in these trends is statistically significant. In our baseline full sample results, we estimate that around 13 to 20 percent of the total increase in the probability a woman enters her father's occupation over our sample period can be attributed to an increase in the transmission of occupation-specific human capital between fathers and daughters, an estimate that we argue is likely a lower bound.

We perform a number of robustness checks and confirm our key empirical findings. In particular, our results are robust to changing the definition of occupation, accounting for (1) changes in labor force participation rates over time, (2) changes in the age at first marriage and the age of retirement for women, (3) changing educational attainment of women, and (4) the changing composition of male employment. Our key implications are also robust to using alternative definitions of occupation, including one that defines "occupations" to be industries. Finally, in contrast to what we find for daughters, we find no increase over time in the fraction of sons working in their fathers' occupations, nor any evidence that there has been an increased amount of specific human capital transmission over time between fathers and sons. Our results document an experience unique to women.

[3] As we discuss below, our data cover cohorts where many mothers were not in the labor force, and where for much of the time information was not even collected on occupations of mothers who were in the labor force. Therefore we abstract completely from mother-daughter transmission, although obviously we are keenly aware of its existence and potential importance.

2 Background and Related Literature

2.1 Estimates of Intergenerational Transmission

Research on intergenerational transmission between parents and children has a long and rich history across multiple disciplines, going all the way back to Galton's work (1889) on the heritability of height. Becker and Tomes (1979, 1986) present an economic model where the utility of parents is a function of current consumption and the utility of a child. Because the utility of the child is itself a function of the child's general human capital, the parents optimize by choosing between consumption and investments in children. The model in its simplest form generates a straightforward, empirically testable relationship that specifies that the log of the income of the child will be a linear function of the log income of the parent. Actually testing the model empirically, however, is harder than it first appears given measurement error in income, (Solon, 1992), with perhaps the best current estimate of a stable intergenerational income parameter in the United States between fathers and sons standing at 0.6 (Mazumder, 2005).[4] Because these estimates just measure a correlation across generations, they cannot distinguish between a simple model of genetic heritability of traits associated with income and an economic model of investments parents make in children. This point has been made and examined in detail by Mulligan (1999) and Grawe and Mulligan (2002), who derive tests aiming to distinguish between economic models and models of heritability and find some evidence weakly consistent with investments. The model we specify below follows in the tradition of treating intergenerational human capital transmission as arising from parental investments in children, although in our discussion of the empirical results we simply refer to transmission of human capital, be it via investments or heredity.[5]

In sociology, one common tradition has been to estimate intergenerational measures of "occupational prestige" and "occupational mobility." Sons may enter their fathers' occupations because of investments that fathers make in sons, because of heritable aspects of occupation-specific skills that lead sons to have comparative advantages in their fathers' occupations, or

[4] There has been much less research devoted to intergenerational income transmission between pairs other than father-son. One notable exception is Chadwick and Solon (2002). Fernandez et al. (2004) discuss preference formation in an intergenerational transmission framework between mothers and sons.

[5] In discussing our model, we do briefly outline the empirical implications that would arise under a model of pure heritability, where the main empirical implication would be that we are testing for an increase over time in the return to occupation-specific heritability.

because of barriers to movement out of a father's occupation. Contingency tables (transition matrices) can be utilized to measure the extent of occupational mobility, where the cells of the contingency table are determined by fathers' occupations and sons' occupations. Occupational mobility can then calculated as the probability or odds of a son not entering his father's occupation (see, e.g., Ferrie, 2005 and Mosteller, 1968).

Measurements of the intergenerational transmission of occupational prestige involve rankings of occupations along some index, usually determined as functions of average income in occupations, and estimating the correlation in occupational prestige across generations. A few of these studies do examine intergenerational generational transmission between fathers and daughters (see, e.g., DiPrete and Grusky, 1990). While the exact specification of the occupational prestige index may be subject to criticism, using average incomes in an occupation may mitigate some of the problems associated with noisy measures of permanent income that have plagued some of the estimates of intergenerational income transmission in the economics literature.

2.2 Changes in Intergenerational Transmission

Sociologists have long been interested in changes in intergenerational transmission across generations, as seen in early work discussed by Featherman and Hauser (1978). Economists have only recently begun to examine this issue, partially as a result of the increasing availability of panel data with enough years of data to estimate changes in the transmission parameter over birth cohorts. Evidence on the extent of change in intergenerational transmission of income between fathers and sons in the United States is mixed (see Fertig, 2003, Lee and Solon, 2009, and the references therein) and depends to a large extent on the data sets used, how income is measured, and the time span considered. Partially as a result of this, we are careful in this paper to combine data from three different data sets collected over a time span of 29 years to ensure the robustness of our results across data sets that differ in the timing of data collection, the wording of questions, and sampling schemes.

When estimating changes in occupational mobility over (at least) two different generations, researchers distinguish between changes in "prevalence" and changes in "association." Changes in prevalence refer to changes across generations in the marginal distributions of the rows and columns of the contingency tables, whereas changes in association refer to changes that are left over once marginal distributions of contingency tables have been adjusted to be equal. It is changes in association that are generally referred to as changes in occupational mobility over time.[6]

What is absent from much of the empirical investigation into changes in intergenerational transmission is an investigation of underlying changes in behavior. For example, when contingency tables are adjusted for prevalence, so that only changes in association are used to quantify changes in occupational mobility over generations, there is no consideration given for why it may be that the marginal distributions of occupations have changed across the generations. In the case of women and their fathers, the fact that women have become more likely over time to be in male-dominated occupations may be a function of changes in investments made by their fathers. Adjusting contingency tables so that the marginal distribution of women's occupations for recent cohorts looks like that of much older cohorts may adjust away the important changes in the impact of fathers on the occupation choices of their daughters. As a result, we take an entirely different approach.

3 The Illustrative Model

In this section, we develop an illustrative model to motivate how fathers' incentives to invest in daughters change as women's labor market opportunities change. To do this, we combine a model of intergenerational transmission and a model of occupation choice. We then use the model to illustrate how daughters' fathers-in-law can be used to control for changes in the marginal distributions of occupations over time. The comparative statics generate an empirical test of changes over time in the transmission of occupation-specific human capital from fathers to daughters.

The model consists of an occupational choice decision nested within a model of human capital investments in children. First, the father chooses the amount of the consumption good to purchase and the amount of investment to make in his daughter's general human capital H and job-specific human capital S, given his income I. The father can only invest in job-specific human capital for his own occupation.

[6] For a summary of statistical methods to adjust for differences in prevalence across contingency tables, see Little and Wu (1991). For a recent study of changes in occupational mobility, see Ferrie (2005) who concludes that occupation mobility in the United States has fallen over the 20th century. For an analogy between these methods and estimation techniques more commonly used by economists, see Hellerstein and Imbens (1999).

The daughter chooses her occupation conditional on paternal investments that have been made, and may decide to remain out of the labor force. We begin with the daughter's occupation decision.

3.1 The Daughter's Problem

The daughter gains utility from working or not working, given the investment made by her father. We think of the daughter's choice of occupation as arising from the maximization of her latent utility y_j^* over four possible occupations. Occupation 1 is her father's occupation. Occupations differ in their closeness to each other, where occupations 1 and 2 are "closer" in utility terms to each other than occupation 3. Occupation 4 represents the choice of the daughter to remain out of the labor force. The woman's utility in each of the four occupations is represented as:

$$y_1^* = \alpha + \beta H + \gamma S + \epsilon_1$$

$$y_2^* = \alpha + \beta H + \epsilon_2$$

$$y_3^* = \beta H + \epsilon_3$$

$$y_4^* = \beta_o H + \epsilon_4.$$

In this formulation, occupation 2 is close to occupation 1 in the sense that they have the same (presumably positive) intercept shift α in utility. We think about this as reflecting something about a woman's taste for a set of occupations, as defined by her father's occupation, so that occupation 2 is a closer substitute to occupation 1 in utility than is occupation 3. These tastes may reflect factors, such as the social class in which the woman was raised or labor market networks defined by her father's occupation, that change the woman's preferences directly or decrease costs of entry in some occupations more than others. To again use the example we discussed in the Introduction, a woman whose father is a doctor treating farmers may find herself, all else equal, more likely to become a farmer than to become a professor.[7]

Note that general human capital pays the same return to a woman in the labor market regardless of what occupation she chooses, but a different return (β_o) if she is out of the labor

[7]At least conceptually, our discussion of occupations 1 and 2 (versus 3) bears similarity to a distinction in sociology that has been made between "big classes" (corresponding roughly to occupations 1 and 2 in our formulation) or "micro-classes" (corresponding roughly to our concept of occupation 1 alone). See, e.g., Weeden and Grusky, 2005.

market. Specific human capital only has a payoff if the woman enters her father's occupation. Alternatively and without loss of generality, y_4^* could instead represent the latent utility in an entirely female occupation. In that case the occupations 1 - 3 represent occupations in which men do work. The error terms ϵ_j can represent differences across occupations in a woman's underlying ability (comparative advantage) in that occupation, or preferences for that occupation.[8]

The daughter will choose the occupation j which yields the maximum value of y^*. If one has data from a sample of daughters that contains information on their occupational choices and those of their fathers, one can formulate an empirical test of whether $\gamma > 0$ without having actual data on S. In particular, one can estimate a discrete choice model of occupation by making functional form assumptions on the ϵ's, and by making the distinction between occupation 2 on the one hand, and occupations 3 and 4 on the other; that is, to model the closeness of occupations.

As is often done in models of occupational choice, assume the ϵ's are i.i.d from a Type I Extreme Value Distribution. Conditional on the H and S with which she has been endowed, the probability that a woman enters her father's occupation is:

$$\Pr(y_1 = 1) = \frac{e^{\alpha+\beta H+\gamma S}}{e^{\alpha+\beta H+\gamma S} + e^{\alpha+\beta H} + e^{\beta H} + e^{\beta_o H}}. \tag{1}$$

The probability that a woman enters occupation 2 is

$$\Pr(y_2 = 1) = \frac{e^{\alpha+\beta H}}{e^{\alpha+\beta H+\gamma S} + e^{\alpha+\beta H} + e^{\beta H} + e^{\beta_o H}}. \tag{2}$$

Note that if $S = 0$, so that the father has not made any investments in the daughter, then $Pr(y_1 = 1) = Pr(y_2 = 1)$. Therefore, if one does not have data on the investments of fathers, and in particular no data on S, an empirical test of whether fathers are making any specific human capital investments in their daughters can instead involve testing whether $Pr(y_1 = 1) = Pr(y_2 = 1)$. This test presumes no knowledge of S, but it does presume the ability to distinguish occupation 2 from the other occupations. Given a sample of women and a mapping of real occupations (e.g., doctor, farmer, professor, and out of the labor force) to occupations 1-4, this test is operationalized by estimating a multinomial logit regression over the four occupations. It requires including one dummy variable to represent α for the father's occupation, occupation 1, and the occupation that is close to it, occupation 2. A second dummy variable would be

[8] The closeness of occupations could also be modeled through assumptions on the covariance of the ϵ_j's, as for example in a nested multinomial logit model. For our purposes, modeling the closeness explicitly in the utility function through an intercept shift α is sufficient.

included for the father's occupation alone, where an estimated coefficient greater than zero on this dummy variable implies the existence of occupation-specific transmission between fathers and daughters.

Identifying which occupations are close together requires assumptions, however. First, one would have to make the assumption that closeness is the same across all women (at least in unobservable ways). Second, the nature of occupations and women's roles in those occupations themselves has changed over time in ways that may reflect changes in the closeness of occupations. To the extent that we are ultimately interested in measuring changes over time in the transmission of occupation-specific human capital between fathers and daughters, we would need to model changes in the closeness over time. One simple way we could implement this is to assume that α is equal to zero and estimate a standard multinomial logit model of occupation choice, including a dummy variable for father's occupation and, to estimate changes in the probability of entering a father's occupation over time, an interaction between a time (birth cohort) trend and father's occupation. As we discuss below, we have assembled a sample of women for whom we do not have complete information on father's occupation. Nonetheless, for the subset of our estimation sample for whom we do have information on father's occupation we have estimated this kind of simple multinomial logit model. The assumption of the Independence of Irrelevant Alternatives is strongly rejected by the data, substantiating that we do need to model the substitutability of occupations.

Therefore, we instead define which occupations are "close" to each other in utility terms by assuming that assortative mating occurs along the dimension of α, so that a woman whose father is in occupation 1 will always choose a husband whose father is in occupation 1 or occupation 2. This is true regardless of what occupation 2 happens to be for that woman, and whatever its origin (social networks, labor market networks across occupations, etc.). Repeating our example, the daughter of a doctor who treats farmers may be more likely all else equal to become a farmer (versus, say, a professor) because her father's social and labor market networks puts her in contact with farmers. The strict assortative mating assumption implies that this woman would meet and marry a man whose father is either a doctor or a farmer. In contrast, the daughter of a doctor who treats professors is more likely to become a professor (rather than a farmer) and will meet and marry a man whose father is either a doctor or a professor. In our view, a taxonomy of closeness defined by something like social or labor market networks that influence occupational choice and influence assortative mating is much less mutable over time than closeness as defined by the tasks performed in given occupations, for example. The assumption that assortative mating is perfect along occupations 1 and 2 in the model is obviously a strict assumption. Below we discuss the implications of relaxing both the assumption of strict assortative mating and the (less restrictive) assumption of constant assortative mating over time.

Given this assortative mating assumption, the probability that a woman is in the occupation of her father-in-law is:

$$\Pr(y_{father-in-law} = 1) = \Pr(y_1 = 1)\sigma + \Pr(y_2 = 1)(1 - \sigma). \tag{3}$$

We assume that the father-in-law works in either occupation 1 or 2. Therefore σ can be thought of as the number of men in the woman's father-in-law's cohort who are in occupation 1 divided by the number of men who are in either occupation 1 or 2 (i.e., the probability the father-in-law is in occupation 1). Similarly, $(1-\sigma)$ is the fraction of men in this population in occupation 2. Note again that if the father makes no specific human capital investments in his daughter so that $S = 0$, then $Pr(y_1 = 1) = Pr(y_{father-in-law} = 1)$. This generates an empirical test of whether $S = 0$, which would involve testing whether $Pr(y_1 = 1) = Pr(y_{father-in-law} = 1)$.

3.2 The Father's Problem

We assume that the father gains utility from his own consumption and from the utility of his daughter. The father has a finite level of income I to allocate between his own consumption, general human capital investment (e.g., schooling) in his daughter, and job specific human capital investment in his daughter. We assume that the father can only invest in occupation-specific human capital S for his own occupation. There are many possible forms these specific investments may take. For example, a father could make explicit investments in teaching his daughter his trade (either through teaching her himself or spending money to have her trained by others). He could also spend more time with her, and through their time together, demonstrate the value of working in his occupation. A father could invest in lowering barriers to entry for his daughter in his own occupation. Or, he could give his daughter monetary transfers and a taste for his occupation that she could use to invest in the skills necessary to work in his occupation.

The father's problem takes the following form:

$$\max_{H,S} \left\{ E[u^F(C, max_j y_j^*(H,S))] \right\} \tag{4}$$

$$s.t.\ I = C + p_H H + p_S S. \tag{5}$$

where $u^F(C, y_j^*(H,S))$ represents the father's utility, a function of own consumption, C, and daughter's utility, $max_j y_j^*$ as determined by her occupation choice, and where p_H is the cost of investments in general human capital and p_S is the cost of investments in occupation-specific human capital.

The father calculates expected utility knowing β, γ, and only the distribution of the ϵ's in the daughter's optimization problem. One could make functional form assumptions about the form of the father's utility function, but this is unnecessary for our purposes.[9]

3.3 Comparative Statics, Heredity vs. Transmission, and Empirical Strategy

In the model, a father must make predictions about the actions of his daughter and decide on the levels of general and specific human capital investments to make in his daughter in order to maximize his utility. A father's investment decision will change with exogenous changes in the parameters of the model. We focus on changes in β, the return to general human capital, where we can think of an increase in β as representing an overall rise in the return to female labor market participation. Because we have no direct data on investments of H or S that fathers make in daughters, we cannot directly examine what happens to these investments over time. Instead, we derive a comparative static that shows that if a father's investment in S increases with β, then the probability that a woman will enter her father's occupation increases relative to the probability a woman enters her father-in-law's occupation.

From the daughter's problem, we derive the following comparative static from considering how the probability a woman enters her father's occupation changes with respect to β:

$$\frac{\partial \ln[Pr(y_1 = 1)]}{\partial \beta} = \gamma \frac{\partial S}{\partial \beta} + \frac{\partial \ln[Pr(y_2 = 1)]}{\partial \beta}, \tag{6}$$

This shows that if $\frac{\partial S}{\partial \beta} > 0$, the rate of change at which the daughter enters her father's occupation, occupation 1, due to a rise in β is larger than the rate of change at which she enters

[9] An obvious functional form assumption to make is that $u^F(C, y_j^*(H,S)) = \phi \ln(C) + (1-\phi)E[max_j\{y_j^*\}]$. This, coupled with the assumption that $\beta_0 = 0$ would lead to a closed form solution of $u^F(C, y^*(H,S)) = \phi \ln(C) + (1-\phi)\ln[e^{\alpha+\beta H+\gamma S} + e^{\alpha+\beta H} + e^{\beta H} + 1] + E$, where E is Euler's constant (see, e.g., McFadden, 1981).

occupation 2.[10]

It is also the case that,

$$\frac{\partial \ln[Pr(y_{father-in-law} = 1)]}{\partial \beta} = \frac{\sigma\gamma\frac{\partial S}{\partial \beta}e^{\gamma S}}{(\sigma e^{\gamma S} + (1-\sigma))} + \frac{\partial \ln[Pr(y_2 = 1)]}{\partial \beta}, \tag{7}$$

so that the rate of change at which she enters her father-in-law's occupation is also positive if $\frac{\partial S}{\partial \beta} > 0$. The difference between these two comparative statics is:

$$\gamma\frac{\partial S}{\partial \beta}(1 - \frac{\sigma e^{\gamma S}}{(\sigma e^{\gamma S} + (1-\sigma))}), \tag{8}$$

which is positive as long as $\frac{\partial S}{\partial \beta}$ is positive, and zero if $\frac{\partial S}{\partial \beta}$ is zero.

Note that these comparative statics focus on the effects of changes in β on investments in S. Our empirical work, however, examines changes over time, not changes in β. Given the extensive empirical evidence in the literature, we assume that $\frac{\partial \beta}{\partial t} > 0$. Therefore, these comparative statics generate an empirical test of whether fathers' specific human capital investments in daughters have increased over time that can be cast as examining the difference between Equation 6 and Equation 7. Moreover, because the last term in parentheses in Equation 8 is positive, the difference between Equation 6 and Equation 7 actually provides a lower bound estimate for the rate at which changes in S increase the probability that a woman works in her father's occupation, occupation 1, relative to occupation 2.[11]

Until this point, the comparative statics assume that $\frac{\partial \gamma}{\partial t} = 0$, and focus only on the effects of changes in β (arising from $\frac{\partial \beta}{\partial t} > 0$) on investments in S. It is worth considering what happens if in addition or instead, $\frac{\partial \gamma}{\partial t} > 0$, so that over time there has also been a rise in the return to

[10] Given the functional form of the utility function suggested in footnote 9, an interior solution for the optimal level of specific human capital S is $S = \frac{\ln[p_S\beta]+\ln[1+e^{-\alpha}]-\ln[p_H\gamma-p_S\beta]}{\gamma}$. One of the requirements for an interior solution is therefore the sensible condition that $\frac{\gamma}{\beta} > \frac{p_S}{p_H}$, or that the relative return to investing in specific rather than general human capital is greater than the relative price. The relevant comparative static is $\frac{\partial S}{\partial \beta} = \frac{\gamma p_H}{\beta(\gamma p_H - \beta p_S)}$ which must be greater than zero.

[11] We recognize that this model is simple in many ways and potentially could be extended along a number of interesting dimensions. For example, it incorporates no dynamics of the form of increasing β leading to increasing H and S which lead to further changes in the returns to H and S (similar in spirit to Fernandez et al., 2004). It would also be interesting to expand our model to incorporate a search model of marriage with the intergenerational transmission of human capital framework. Ermisch et al. (2006) contains a model of general human capital investment and marriage. It would also be interesting to consider investments that vary by family structure (e.g., family size, marital status of parents). We do not have data that would allow us to investigate these differences empirically, so we do not consider them further.

occupation-specific human capital. In that case, the difference over time in the probability that a woman enters her father's occupation versus her father-in-law's occupation is:

$$[\gamma\frac{\partial S}{\partial t} + S\frac{\partial \gamma}{\partial t}][(1 - \frac{\sigma e^{\gamma S}}{(\sigma e^{\gamma S} + (1-\sigma))})], \tag{9}$$

If this is positive, then it is impossible to distinguish whether this arises because $\frac{\partial S}{\partial t} > 0$ or because $\frac{\partial \gamma}{\partial t} > 0$. This is the manifestation in our context of the more general problem of disentangling heredity from investments that is common to much of the empirical work on human capital transmission. While we find increases in investment to be a more compelling interpretation of our empirical test than one solely of increasing returns, especially within the context of economic models of utility maximization, we have no direct evidence on whether or not $\frac{\partial \gamma}{\partial t} > 0$. In the empirical work, therefore, we refer to changes over time in the more general phenomenon of occupational transmission from fathers to daughters, rather than explicitly tying our empirical results to increasing paternal investments in daughters.

3.4 Imperfect Assortative Mating

Recall that our strict assortative mating assumption requires that a woman's father-in-law be chosen from occupation 1 or occupation 2 and not from occupation 3. Of course, in reality assortative mating is not perfect along the dimensions of sets of occupations as defined in the model with the parameter α, and so it is important to understand how violations of strict assortative mating might impact our results. One obvious way in which imperfect assortative mating could occur is if there is some probability that the woman will marry a man whose father is in occupation 3. If this probability is unchanging over time, however, this will simply lead to an intercept shift down in the probability that the woman is in her father-in-law's occupation, and more specifically will not alter the comparative statics as β rises. Alternatively, assortative mating could be stronger than we assume, so that the woman may be more likely to marry a man whose father is in her father's occupation than in any other. To the extent that this is true, the equality of the two sides of Equation 7 will no longer hold, and instead the right-hand-side of the equation will be less than the left-hand-side. This will lead us to further underestimate the extent to which increased specific human capital investments have induced women to enter their fathers' occupations.[12] Finally, it is possible that, in reality, assortative mating patterns

[12] This is also the reason why we use the probability that a woman is in her father-in-law's occupation rather than in her husband's occupation in our empirical test. If, for example, successive cohorts of men

themselves have changed over time. To the extent that women are more likely than previously to marry a man whose father is in her own father's occupation, this again will cause us to underestimate the extent to which father's specific human capital investments have increased. Thus, the key identification assumption with respect to assortative mating in our model is that assortative mating by occupation cannot have gone down across birth cohorts.

In the context of our model, the assortative mating assumption can be thought of as occurring along the dimension of "α." In words, it implies that a woman's father-in-law works in an occupation that is "close" to her father's occupation. The identifying assumption necessary for our empirical test to be valid is simply that assortative mating along α has not decreased over time. In practice, one implication of this assumption is that the probability that a father and father-in-law work in the same occupation is not declining over time. Because data constraints prevent us from testing this directly, we instead estimate a series of regressions where we test whether there has been a change across birth cohorts in the probability that a woman's husband works in her father's occupation (his father-in-law). This approach is similar in spirit to a test of assortative mating used by Lam and Schoeni (1994) who examine the extent of correlations in incomes of fathers-in-law (that is, women's fathers) and sons-in-law (women's husbands) as a measure of assortative mating.[13] We detail the findings of this exercise in Section 5. To preview, we find that the probability that a man works in the same occupation as his wife's father is high (around 25 percent), and there is no evidence in any of our results of a decline across birth cohorts in this probability. That is, we find no evidence suggesting that assortative mating by fathers' occupation has decreased over time.

---

are increasingly more (less) likely to enter occupation 1, we will understate (overstate) the importance of transmission between fathers and daughters if we use husbands rather than fathers-in-law in our empirical test. By using fathers-in-law as the "counterfactual" for fathers, we draw men from the same cohort as fathers with the same underlying distribution of occupations.

[13]Lam and Schoeni (1994) compare the intergenerational income correlation between fathers and sons and fathers-in-law and sons-in-law in the United States and Brazil. The father-son correlation is higher than the father-in-law-son-in-law correlation in the United States, but the opposite is true in Brazil. They argue that in Brazil assortative mating is so strong as to match husbands to fathers-in-laws who are more similar to them than the husband's own fathers, but that this is not true in the United States. For other papers examining the extent of assortative mating and its change over time see, e.g., Mare (1991) and references therein.

4 Data and Summary Statistics

4.1 The Data Sets

To create our sample, we combined data from three sources: the 1973 Occupational Changes in a Generation (OCG), the General Social Survey (using years 1975-2002), and the Survey of Income and Program Participation (1986-1988, Wave II). In the Data Appendix we provide an explanation of how the main variables of interest, labor force participation and occupation, were defined.

We chose to focus on more than one survey, and on these three surveys in particular, for a few reasons. First, these surveys are similar in that they are cross-sectional in nature and all ask information about a wife's occupation and the occupation of at least her father or her father-in-law, and sometimes both. Both the SIPP and OCG samples were specifically designed to capture intergenerational information, and the GSS has the advantage that because it consists of repeated cross-sections, it helps us to separate age and cohort effects. Second, together the three samples comprise a large enough sample to allow us to obtain precise estimates. Third, because we use data spanning the years 1973 to 2002 and focus on individuals between the ages of 25 and 64, we are able to estimate effects for birth cohorts spanning a long time period: 1909 to 1977. Finally, using multiple data sets allow us to examine the robustness of our estimates. This is important given the heterogeneous findings in research on intergenerational income transmission for men. That said, we can only compare results across these data to the extent that the cross-sectional data sets do not confound age and cohort effects, something we return to below.[14]

The GSS has the distinct advantage of being drawn from a series of nationally representative cross-sectional data sets over a long period of years. Because of this, when a series of GSS's are linked together, there are observations on individuals at different ages who were born in the same birth cohort, allowing analyses that separately identify age and cohort effects. This is vital in our context because our aim is to identify how the relationship between fathers' occupations and daughters' occupations has changed across birth cohorts, conditional on the age of the women in the sample. This analysis obviously cannot be done conditional on age with only one

[14] We considered adding information from longitudinal data sets like the NLS and PSID, but they are different enough in structure and do not contain information on in-laws of sample individuals or families that we chose not to utilize them in this paper.

cross-section of data. The GSS does have a few shortcomings, however. First, it is a small data set, even when surveys are pooled over multiple years. Second, the unit of observation in the GSS is an individual and not a household, so while information is collected on the occupation of the respondent, the respondent's father, and the respondent's spouse there is no information on the occupation of the respondent's father-in-law. As a result, one cannot use data from the GSS to estimate a full-blown multinomial occupational choice model, where one estimates whether a given woman is more likely to go into her father's occupation and her father-in-law's occupation, relative to other occupations. Third, until 1992, no information was collected on the occupation of the respondent's mother if she was in the labor force. So it is not possible to do an analysis of changes across multiple birth cohorts in occupational transmission between mothers and daughters. Notwithstanding these shortcomings, the GSS data can be used for our empirical test of changing transmission bewteen fathers and daughters. We utilize data from the GSS surveys of 1975-2002.[15]

The 1973 OCG, while no longer a particularly well known data set, is an obvious candidate survey for this paper because it was a large survey that was designed specifically to capture intergenerational relationships (see Featherman and Hauser, 1978, for more information). Because we combine data from the OCG with later surveys, we concorded the 1970 occupation codes that are used in the OCG to 1980 so that the occupations would be comparable. More details on this are given in the Data Appendix.[16] The OCG was conducted as a supplement to the Current Population Survey (CPS) in March 1973. Questionnaires were mailed out to (only) male CPS respondents, specifically asking information about their family and their background, including the occupation of their father when they were 16 and, for married respondents, the occupation of their wife's father when their wife was 16. These responses, combined with the occupation responses and other background variables given as part of regular CPS survey, allow us to have for our sample the key information that we need to conduct our analysis: occupations of the husband and wife, the occupations of each of their fathers, and the ages of the husband and wife. Like the earlier years of the GSS, the OCG did not collect information on the occupations of respondent's mothers, nor of their mothers-in-law. Similarly, since the survey itself was sent only to men, the OCG data only contain information for married women (i.e, women who were

[15] 1975 was the first year that the GSS employed standard probability sampling

[16] There was a 1962 OCG survey as well, but we have chosen not to use it because it would have required yet another concordance, of 1960 occupations to 1980 occupations.

married to the male respondents).

The SIPP Personal History Topical Modules in 1986, 1987, and 1988 were designed to mimic the OCG in many ways. They contain similar information to the OCG, but the information was collected at the level of a responding family, and without regard to the sex of the family member(s) completing the survey. Because these SIPP topical modules were all conducted in Wave II, there is very little of the attrition that sometimes plagues studies that use the SIPP. Unfortunately, these particular topical modules have not been repeated for more recent years.

Because our analysis relies on using information on the occupation of fathers-in-law, we necessarily restrict the data from the GSS and SIPP to married women (the OCG, as described above, is already restricted to married women). We further restrict the sample to only whites, so as not to confound occupational changes that are unique to women with those that are due to changing opportunities for blacks. Finally, we restrict our baseline sample to those between the ages of 25 and 64 in order to obtain information on individuals during their prime working years. But because the age at first marriage has risen over time and the age of retirement has declined over time, we examine the robustness of our results to limiting the age range to those between the ages of 35 and 55.[17]

4.2 The Definition of Occupations

Until this point we have been vague as to what we mean by an occupation and how to operationalize it. Following standard practice, we define occupation using Census definitions. In our baseline results, we use the six major occupation groups as defined by the 1980 Census Occupation Codes: Managerial and Professional Specialty; Technical, Sales, and Administrative Support; Service; Farming, Forestry, and Fishing; Precision Production, Craft, and Repair; and

[17] It is worth noting that because we are combining data across multiple surveys, we have a limited amount of demographic information that is consistently measured across surveys. This limits the scope of questions we can explore. As described above, the limited amount of information on the occupations of mothers prevents us from examining changing transmissions between mothers and daughters or how mothers' investments interact with those of fathers. Additionally, we would expect that changes in family structure over time should have affected the transmission of human capital from fathers to daughters. Rising divorce rates should have reduced the average transmission over time from fathers to daughters if divorce leads daughters to have less exposure to fathers. On the other hand, declining family size may have increased a father's transmission to the focal daughter, since there is less competition for resources from siblings. However, we do not know consistently in our data whether a woman was raised in a household with her father present, nor do we know how many siblings she had (nor their gender composition). What we do estimate, then, is an average trend across family structures in the rate of transmission of occupation-specific human capital between fathers and daughters over time.

Operators, Fabricators, and Laborers. As in our model, for women we also include a seventh occupation group, Out of the Labor Force, which includes women who are not working, are not in school full time, and are not unemployed or looking for work. As part of our robustness checks we disaggregate the list of occupations further, to 13 occupations listed as subheadings of the three-digit 1980 Occupation Codes.[18] Clearly, the more we disaggregate, the less power we have to detect changes in father-daughter occupation transmission, so we do not consider levels of occupational disaggregation below this. Perhaps more importantly, as mentioned above, the theoretical notion of occupation-specific human capital does not map directly to Census occupation classifications. For example, just as the literature on job-specific human capital can be recast to be about industry-specific human capital (see, e.g., Neal, 1995), so our definition of occupation can be recast to map into industries. We therefore also present our main results using an indicator of a woman being in the same industry as her father or father-in-law.

Table 1 contains summary statistics for our pooled sample, as well as for each data set. The statistics cover the occupational breakdown of women, fathers, and fathers-in-law, as well as age and birth year of women in our sample. The pooled data set contains information on every woman in our estimation sample, where we observe either the occupation of her father, the occupation of her father-in-law, or both.[19] In our pooled data set, almost half (46.2 percent) of women are out of the labor force, with the next most populated occupation being Technical, Sales, and Administrative Support, comprising 22.8 percent of the sample. By comparing the proportions in each occupation across data sets, and particularly by comparing women in the OCG and women in the SIPP, one can clearly see how the occupational distribution of women has changed over time. In the OCG, 57.0 percent of women are coded as out of the labor force, whereas only 37.3 percent of women are in the SIPP. Moreover, conditional on labor force participation, women in the SIPP are more likely than their earlier counterparts to be either managerial and professional occupations or in technical, sales, and administrative support (46.1 percent in the SIPP versus 27.7 percent in the OCG).

The occupation distributions of fathers and fathers-in-law are very similar within each data set, as they should be absent non-random sampling or differential response rates by occupation of parents and in-laws.[20] Over time for these fathers and fathers-in-law there are also changes

[18]See Appendix Table A.2 for a mapping between the six and thirteen occupation category groupings.

[19]The distribution of occupation for women is very similar when women for whom we have no information on fathers or fathers-in-law are included.

[20]The one exception is the somewhat greater likelihood that fathers-in-law will be farmers (relative

in the occupational distribution; for example, these men are less likely to be in farming in the SIPP relative to the OCG. Because of this, and because fathers in different occupations may transmit different amounts of specific human capital to children, we show results below with and without occupation controls for fathers and fathers-in-law.

Below the distributions of occupations in Table 1 we present summary statistics on the fraction of women who are in their father's and father-in-law's occupations in each data set. Overall, 10.7 percent of women in the data work in their father's occupation, and 9.9 percent work in their father-in-law's occupation. While these differences are not large in absolute terms, they are in percentage terms. Moreover, across data sets, it becomes clear that the differences grow over the birth cohorts in our sample: in the OCG, where the mean birth year of women is 1931, the difference between the two means is 0.2 percentage points, whereas in the GSS, where the mean birth year is 1946, the difference is 1.8 percentage points.

5 Empirical Implementation and Results

Our basic empirical strategy is to compare the trends over birth cohorts in the probability that a woman works in her father's occupation relative to the probability that a woman is in her father-in-law's occupation. We do this via a single regression equation:

$$Prob(same = 1)_i = \delta_0 + \delta_1 * DIL_i + \delta_2 * D_i * Y_i + \delta_3 * DIL_i * Y_i + \delta_4 * D_i * A_i + \delta_5 * DIL_i * A_i + \varepsilon_i. \quad (10)$$

Recall that for (only) some of the women in our sample, we observe information on her occupation and that of both her father and her father-in-law. For other women, we only observe information on her occupation and that of her father, or information on her occupation and that of her father-in-law. An observation i in the regression is defined as a woman paired either with her father, or with her father-in-law. Women for whom we have information on both her father and her father-in-law comprise two separate observations for the purposes of estimation: one observation for her and her father, and another one for her and her father-in-law. In these cases, clustered standard errors for the two observations from a single woman take care of their obvious dependence.

to fathers). This is due to the fact that women tend to marry somewhat older men, so that fathers-in-law are somewhat older than fathers, and farming has been a continuously declining industry. The empirical results we present later are robust to excluding women who are married to much older men, and excluding women whose fathers-in-law or fathers are farmers.

We define DIL to be an indicator that takes the value of one if an observation contains information on a woman and her father-in-law. Similarly, D is an indicator for an observation containing information on a woman and her father. In other words $D = 1 - DIL$. In defining the dependent variable, *same* is an indicator which, for observations on a woman paired with her father, equals one if the woman is in the same occupation as her father. If the observation is of a woman paired with her father-in-law, *same* equals one if the woman is in the same occupation as her father-in-law. Y is the birth year of the woman, and A is the age of the woman.

The coefficient δ_2 measures the annual change in the probability that a woman enters her father's ocupation (conditional on the other covariates), while the coefficient δ_3 measures the annual change in the probability that a woman enters her father-in-law's occupation. Note that the empirical prediction of the theoretical model suggests that we should be comparing the rates of change in the probabilities over birth cohorts of women enter their fathers' occupations and that of their fathers-in-law, rather than absolute changes. But, as we show below, the estimate of δ_1 is small, and, when statistically significant, is positive. This indicates that the baseline probability for fathers and daughters to be in the same occupation is the same or lower as that for fathers-in-law and daughters-in-law. Therefore, our key empirical test of increasing transmission over birth cohorts of women of occupation-specific human capital from fathers to their daughters is simply a straightforward test of whether δ_2 is greater than δ_3.

Controlling for age (when possible) is important because women may transition into their "final" occupations as they gain experience in the labor market and, more importantly, as women move in and out of the labor force when they have children. Theoretically, it is quite possible for the coefficients on age, δ_4 and δ_5, to be different if, for example, a woman moves into her father's occupation as she gains experience in the labor market.

5.1 Results

In Table 2 we report basic results for this regression specification for all three data sets together and then for the three data sets separately. We present results for linear probability models. Marginal effects from logit models (not reported) are virtually identical. Because we cannot separately identify age and cohort effects in the OCG and SIPP, we do not include separate controls for age in this table. Column 1 contains results for the full sample. The estimated coefficient on the daughter's birth year, δ_2, is statistically significant and indicates that the probability that a woman enters her father's occupation increases by 0.27 percentage points per year. To put this in perspective, the fraction of women in their father's occupation born over the first decade of our sample (1909-1919) is only 0.058, so that we estimate each year thereafter leads to a very large 4.59 percent increase in the probability that a woman works in her father's occupation.

The coefficient estimate on the daughter-in-law's birth year is 0.21 and is also statistically significant.[21] The fact that this point estimate is also large in magnitude, a finding repeated throughout the empirical results to follow, highlights the importance of controlling for overall trends in women's labor market entry when teasing out the distinct impact of the change in the extent of occupation-specific human capital transmission between fathers and daughters. We estimate nonetheless that $\delta_2 - \delta_3$, the annual change in the probability of a daughter being in her father's occupation *relative* to the equivalent change for a daughter-in-law/father-in-law pair, is a statistically significant 0.05 percentage points, or a difference of 1.19 percent. This difference, a measure of the impact of increased transmission in specific human capital on the shift toward women working in their fathers' occupations, accounts for approximately one fifth of the overall change over time in the probability that a woman works in her father's occupation.

Figure 1a is the graphical representation of Table 2, column (1), except that instead of using linear regression, we generate these results using locally weighted least squares. There are two important things to take away from this figure. First, the probability that a woman is in her father's occupation is very slightly below that of fathers-in-law and daughters-in-law early in the period, but grows over the period of our sample to be above that of fathers-in-law and daughters-in-law. Second, the time trends in both of these probabilities are indeed close to linear, as we model them in equation 5. (Estimated coefficients on higher order terms on birthyear when added to the regression are statistically insignificant.) Figure 1b presents the same results as in Figure 1a, but in terms of rates of change rather than absolute changes. Here, trends show the probability a woman is in her father's (father-in-law's) occupation relative to the fraction of women in their father's (father-in-law's) occupation in a baseline cohort defined to be the first decade of our sample, 1909-1919. As in Table 2, on average, each successive decade leads to an approximately 50 percent increase in the probability that a woman works in

[21] The fraction of women in their father-in-law's occupation born over the first decade of our sample (1909-1919) is 0.063, which is actually statistically indistinguishable from the fraction of daughters in their fathers' occupations.

her father's occupation.

Column 2 of Table 2 shows results for only the OCG sample of women who were born between 1909 and 1948. Of the women in this sample, 57 percent are recorded as being out of the labor force in 1973. The gradient of the probability of a woman working in her father's occupation is relatively flat over this period, with a precisely estimated increase of 0.09 percentage points every year. The estimated increase in the fraction of women entering their father-in-law's occupation is lower, 0.06 percentage points. The difference between these two is not statistically significant. This is not surprising given that women born in these years largely remained out of the labor force.

Column 3 shows the baseline results for the GSS for women in birth cohorts spanning 1911 to 1977 (although with very few observations for women at the tails of this distribution). The point estimate on the increased probability of father-daughter occupation transmission is 0.33 percentage points per year. Relative to the baseline over the 1909-1919 period, this represents almost a 6 percent increase in this probability per year of the sample. The increase in the probability that a woman works in her father-in-law's occupation is smaller, at 0.27 percentage points. Finally, the relative difference between these two is 0.06 percentage points per year, which while not statistically significant, is 17.8 percent of the overall increase in the probability that a woman works in her father's occupation. In column 4 we report results for the SIPP sample, representing women born 1921-1963. These results are similar to those for the GSS.

In Table 3 we examine results for various specifications of the model in the full sample of pooled data. Column 1 replicates the baseline results of Table 2 but includes controls for the survey from which the observation comes. If survey questions differ in a way that might affect the baseline probability of a woman being in a man's occupation, the inclusion of these controls should account for that.[22] The point estimates of the father-daughter and father-in-law-daughter-in-law trends are somewhat smaller than in the previous table, but still are large and statistically significant. Moreover, the difference in the trends between fathers and daughters and fathers-in-law and daughters-in-law again is 0.06 percentage points per year and is statistically significant. There are a few other things to note in this specification. First, there are statistically significant differences in the constant terms across data sets. As we show below, however, this result does not hold up in other specifications when we include more controls. Second, the

[22]For example, the GSS asks the respondent to report the occupation of her father while she was growing up, while the SIPP and OCG ask for the occupation of her father when she was 16 years old.

dummy variable for the daughter-in-law equation constant, δ_1 in the regression equation, has a coefficient of 0.02 and is statistically significant. This result is also not robust.

In column 2 of Table 3 we include variables for age separately for daughters and daughters-in-law, as in Equation 5. The estimates of the coefficients on the age variables are highly significant, almost identical (0.25 and 0.23 percentage points), and statistically indistinguishable from one another. They imply that every for 10 years that the woman ages the probability that a woman enters her father or father-in-law's occupation increases by over a healthy 2 percentage points. Given this result, and given that birth year and age are negatively correlated in these data, the inclusion of age into the model should cause the coefficients on the birth year trend variables to go up. Indeed, the estimates more than double, indicating large changes between birth cohorts in the probability that a woman works in both the occupations of her father and her father-in-law. The coefficient on the birth year of a daughter rises to 0.44 percentage points (from 0.22), while the coefficient on the birth year of a daughter-in-law rises to 0.37 percentage points. The former result can be interpreted as a 7.6 percent increase per year relative to the baseline probability, while the latter yields a 5.9 percent increase. The estimate of the relative difference in the two trend variables is 0.07 percentage points and again is statistically significant.

The fact that the dummy variables for the survey of origin and the dummy variable for the daughter-in-law are all statistically significant in column 2 leads to the specification in column 3, where we interact the survey of origin dummy variables with the daughter-in-law dummy variable. The point estimates on birth year are slightly closer together and slightly less precise, so while the estimate of the relative difference between the two is very close to the previous specifications (0.05 percentage points), it is not statistically significant. The full set of interactions between the survey dummies and the daughter-in-law constant leads to small and insignificant differences across the board in these coefficients. Because they are small and statistically insignificant, we drop the interaction terms in the columns that follow, in order to gain more power in estimating the difference in the trends between daughters and daughters-in-law. Similarly, we constrain the coefficient on daughter's age, δ_4, to equal that on daughter-in-law's age, δ_5. These results are presented in column 4, where the point estimates on the birth year coefficients are virtually unchanged, but more precise, so that the estimated difference of 0.05 percentage points between the two coefficients is statistically significant (standard error of 0.02). This difference accounts for 13.1 percent of the overall change in the probability that a woman works in her father's occupation.

Because the distribution of the occupations of fathers and fathers-in-law has changed over time as well (see Table 1) in ways that may affect the probability that a woman is in one of these men's occupations, and because we can only estimate the impact of average investments made by fathers across different occupations, in column 5 we include a full set of controls for the occupations of fathers and fathers-in-law. To the extent that it is fathers transmitting occupation-specific human capital rather than fathers-in-law, there is no reason to expect the coefficients on these dummy variables to be the same for these two groups, and indeed (in results not shown) they are not. Including these dummy variables reduces the point estimates on the birth year variables and the variable for women's age. The coefficient on the birth year of daughters is 0.28 percentage points, a 4.8 percent increase per year over the baseline father-daughter probability, while that of the daughter-in-law is 0.23 percentage points. Both remain highly statistically significant. The difference between these two is 0.05 and is again statistically significant. This represents 16 percent of the overall increase in the probability that a woman works in her father's occupation and, once again, implies that there has been an substantial increase in the transmission of occupation-specific human capital between fathers and daughters.

The last two columns of Table 3 replicate the specifications in columns 4 and 5, but they also include controls for the educational attainment of women. Education partially determines occupational choice, and educational attainment is also correlated with the occupations of fathers and fathers-in-law. Therefore, a woman's educational attainment may represent an important omitted variable in the regressions. Of course, education is also endogenously chosen and we have no way to account for this, so we view these results as a robustness check. As anticipated, the coefficient estimates on the probability a woman is in her father's or father-in-law's occupation decline with the inclusion of education controls, but the difference in these trends is still significant and is similar to the specification without education controls. In both the specification without and with the controls for father's and father-in-law's occupation, the difference in trends represents 19 percent of the increase in the probability a woman is in her father's occupation.

The different specifications reported across columns in Table 3 vary the set of covariates included in the model. In Table 4 we vary other aspects of the model, presenting results that parallel those in columns 4 and 5 of Table 3. In columns 1 through 4 of Table 4 we vary the samples over which we estimate the model. Recall that the age range of women in our baseline sample is 25-64. By necessity all of these women are married. Rising age at first marriage therefore may have an effect on the composition of women across birth cohorts who are included in our sample. That said, the fraction of white women ever marrying has not changed substantially across birth cohorts (Goldstein and Kenney, 2001, and Martin, 2004). Indeed, using successive public use micro-samples of the Decennial Census from 1910 through 2000[23] we find that the fraction of white women ages 35 to 55 who have ever been married has stayed remarkably constant across time. Therefore, to gauge whether rising age at first marriage has an effect on our results, in columns 1 and 2 of Table 4 we restrict the age range of our sample to 35 to 55. Note that we restrict the upper age range to 55 to account for the falling age of retirement.[24] Restricting the sample to women ages 35 to 55 reduces our sample size considerably, from 63,076 to 34,544.

The specification in column 1 of Table 4 mimics that of Table 3, column 4, and therefore includes separate controls for survey and constrains the effect of age to be the same for daughters and daughters-in-law. The coefficient on the birth year variables in this column (0.46 percentage points for daughters and 0.40 percentage points for daughters-in-law) are very similar to those in the previous table and are again statistically significant. The estimate of the difference between the two, 0.06 is also statistically significant. Column 2 adds controls for the occupations of fathers and fathers-in-law, paralleling the specification of Table 3, column 5. Here, the point estimates are again similar to those of Table 3, column 5, but the difference between the two of 0.03 has a larger standard error than in Table 3, presumably because the sample size has been cut almost in half. In total, we interpret these specifications as providing evidence that our results are robust to these sample selection issues.

The theoretical model in Section 3 differentiates between women who are out of the labor force and women who are in a set of occupations in which men work. As mentioned above,

[23]See Ruggles et al. (2008) for information on the public use micro-samples (PUMS).

[24]Differential sample selection by birth cohort due to divorce may remain an issue even for women ages 35 to 55, however. To the extent that divorce rates have risen over most of the birth cohorts in our sample (see, e.g., Stevenson and Wolfers, 2007), more women from recent cohorts should be selected out of the sample due to divorce. Because these divorced women are more likely than married women to participate in the labor force (and therefore to work in their father's or father-in-law's occupation), selection out of the sample for divorced women from more recent birth cohorts should actually lead us to understate the results we report. Note again that in one of our three data sets, the OCG, which comprises many of the observations on women from the earlier birth cohorts, no information at all was collected from non-married women. As a result, we cannot test formally for changing selection into or out of marriage that might affect estimates of the changing probability that married women work in their father's occupation.

we could recast the occupation of women who are out of the labor force in our model (which we labeled occupation 4) to be traditionally female occupations where men never (or almost never) work, such as nursing. The model would yield the same implications. Moreover, our model suggests that if investments in specific human capital S increases as the return to general human capital in the labor market increases, we should see an increase in the probability that a woman works in her father's occupation, relative to that of her father-in-law, even conditional on labor market participation.[25] In columns 3 and 4 we therefore explore this empirically by restricting the sample to include only women who participate in the labor force. This again causes the sample to fall by almost one half. Column 3 replicates the specification of column 4, Table 3. Because so much of the change over time in the probability a woman enters the occupation of her father or father-in-law is due to labor force entry, the coefficients on the birth year trend are lower when we restrict the sample in this way. The coefficient on birth year of daughters is 0.33 percentage points and that of daughters-in-law is 0.25. Both are statistically significant and, importantly, the difference between the two is 0.08 and is statistically significant.

In column 4 we add controls for father and father-in-law occupations. While the birth year coefficients themselves become small and statistically insignificant, the relative difference between the two remains of the same magnitude as the full sample result at 0.05. This estimate is only marginally significant, due to the much smaller sample size. In summary, we interpret the results in columns 3 and 4 as showing that our full sample results are not being driven solely by entry into the labor market, and, as our model suggests, that occupational changes over time by women in the labor market are affected by the transmission of specific human capital between fathers and daughters.

In the last four columns of Table 4 we vary the definition of "occupation." In Columns 5 and 6 we refine the definition of occupation to consist of 13 occupation categories (rather than 6).[26] Perhaps the most important distinction between the categorizations is that the two broad occupations in which most women work conditional on labor market participation, "Managerial and Professional Specialty" and "Technical, Sales, and Administrative Support," are each broken up. In column 5, the estimate on the coefficient on birth year for daughters is 0.19 percentage points and is statistically significant. The estimate on the father-in-law trend

[25]The same would be true if, in our model, the return to specific human capital investments increases over time with no changes in investment.

[26]For a list of the 6 and 13 occupation groupings see Appendix Table A.2.

coefficient is smaller, at 0.15 percentage points, and the difference between the two is 0.04 and is statistically significant. Column 6 includes father's and father-in-law's occupation controls. While the coefficients on the trends fall and the difference between the two falls to 0.03, it is still statistically significant. Again, this implies there has been an increase in the transmission of occupation-specific human capital from fathers to daughters.

Finally, as mentioned previously, while our results thus far have used Census occupation codes, this may not correspond to our theoretical notion of occupation-specific human capital. Therefore, in Columns 7 and 8 of Table 4 we report results from specifications paralleling Columns 4 and 5 of Table 3, now using industry to classify the notion of specific human capital.[27] The magnitudes of the estimates in these columns are similar to those using the more disaggregated occupation categories. The baseline probabilities that a woman is in her father's and father-in-law's occupation are .046 and .049 respectively. Again, we calculate the baseline probability as the fraction of women born between 1909 and 1919 that are in the same occupation as their father (father-in-law). Therefore the results in Column 7 should be interpreted as indicating that for every decade there is a 48.4 percent increase in the probability that a woman works in her father's industry. The difference in trends between a woman entering her father's versus father-in-law's industry is .04 percentage points, representing almost 19 percent of the increase in probability a woman enters her father's industry. Though the trends are flatter in Column 8, when controls are added for father's and father-in-law's industry, the difference is still statistically significant and large. This evidence is consistent with increased specific-human capital transmission from fathers to daughters.

In total, the results are remarkably robust across specifications and samples. There has been a large increase over time in the probability that a woman enters her father's occupation. Moreover, this increase is not due simply to changes in the marginal distribution of women's occupations, but is due, at least partially, to increased transmission of occupation-specific human capital from fathers to daughters. Our results imply that the increase in the probability a woman is in her father's occupation is about 13 to 20 percent larger than the increased probability that

[27]In order to make the definition of industry consistent over time, we collapse the 15 major industries categories from the 1980 Census into 13 categories: (1) Agriculture, Forestry, and Fishing, (2) Mining, (3) Construction, (4-5) Manufacturing (combined Nondurable and Durable Goods), (6) Transportation, Communications, and Other Public Utilities, (7-8) Wholesale Trade (combined Durable and Nondurable Goods), (9) Retail Trade, (10) Finance, Insurance, and Real Estate, (11) Business and Repair Services, (12) Personal Services, (13) Entertainment and Recreational Services, (14) Professional and Related Services, and (15) Public Administration.

a woman will enter her father-in-law's occupation, and this is a lower bound estimate of the impact of increased transmission.

In the last two columns of Table 3 we explored the robustness of our results to the inclusion of education controls for women, noting that occupation and education are correlated. In Figure 2 we explore this possibility more explicitly by using education instead of occupation to generate results analogous to those in Figure 1a. That is, we present locally weighted least squares estimation of the probability a woman has the same level of education as her father and as her father-in-law by her birth cohort. If the results for education look like those for occupation, one might be worried that we are not picking up something about specific human capital transmission between fathers and daughters but instead something more akin to general human capital. Figure 2 looks nothing like Figure 1a. The change over time in the probability a woman has the same education level as her father (or father-in-law) is u-shaped with respect to her year of birth.

5.2 Sons, Fathers, and Fathers-in-Law

Up to now, we have focused exclusively on daughters. However, the trends across birth cohorts in the probabilities that sons enter the occupations of their fathers and their fathers-in-law is important for two reasons. First, as discussed in Section 3.4, our key identification assumption, that assortative mating over the dimension of a daughter's father's occupation has not fallen over time, can be analyzed by estimating the relationship over birth cohorts in the occupations of sons-in-law and fathers-in-law. Second, the main motivation for our focus on fathers and daughters (rather than sons) is the claim that something special changed in the relationship between fathers and daughters as a result of the increased entry of women into the labor market and into traditionally male occupations. To the extent that this is true, we should not see the same trends in the probabilities that sons and daughters work in their fathers' occupations.[28]

In Table 5 we present results of our search for any evidence of declines over birth cohorts in assortative mating by estimating the changing probability that a son works in his father-in-law's occupation. We estimate a series of regressions where we regress a binary variable

[28] In the analysis in this section we exclude from our sample men who are not working. We think it unlikely that most of these men are actively engaged in home production, but instead that they are temporarily out of the labor force, so that the fact that they are not in their father's or father-in-law's occupation at the time of the survey is transitory. This eliminates very few men in practice and including a separate out-of-labor force category for these men does not affect the results.

indicating whether the husband of a woman in our data works in the same occupation as her father (his father-in-law) on dummy variables for the data set from which the observation is drawn, the husband's age, and the husband's birth cohort. We model the husband's birth cohort alternatively as a linear term (in columns 1 and 2), and as a more flexible set of seven cohort dummy variables (in columns 3 and 4). The odd numbered columns report results from simple OLS regressions, and the even numbered columns report results from weighted regressions where the weights serve to control for changes in the marginal distributions of the occupations of fathers-in-law over time (e.g., the decline in farming).[29]

In both columns 1 and 2, the coefficient on the linear birth year term is very small (0.0008 and -0.0007, respectively) and statistically insignificant. In columns 3 and 4, where we instead include dummy variables for the birth cohort (approximately by decade) of the husband, leaving as the omitted group the first decade of our data, the point estimates suggest a somewhat nonlinear pattern, but the results are again small and not statistically significant. In all cases, then, while the probability that a husband works in his father-in-law's occupation is high (around 25 percent), there is no evidence in any specification of a decline across birth cohorts in this probability. This provides support for the key identifying assumption that assortative mating by a woman's fathers' occupation has not decreased over birth cohorts.

Finally, in order to contrast the differences between sons and daughters, Figure 3 presents results from locally weighted least squares that are analogous to those in Figure 1b. We contrast the rate of growth of the fraction of daughters in the same occupation as their father and their father-in-law to the respective rates of growth for sons. The top two "lines" illustrate growth rates in the probability that a woman is in her father or father-in-law's occupation over time, while the bottom "lines" show the same trends but for sons. There are several things to note in this graph. First, while the baseline probability that a son enters his father or father-in-law's occupation is much higher than that for women, the growth rate across birth cohorts for men are essentially flat. There is a slight rise toward the end of the period for fathers and sons, but this is an artifact of a nonlinear increase in the probability of fathers and sons working in the same occupation in the SIPP. This increase is not seen in the GSS over the same birth cohorts. Second, consistent with Table 5, the flat trend for sons and fathers-in-law supports the identifying assumption that assortative mating by occupation has not changed over time.

[29]The weighting algorithm is described in detail in Hellerstein and Imbens, 1999, and is conceptually parallel to adjusting the marginal distributions of contingency tables to be equal.

We also estimated regressions for sons that parallel those of Tables 3 and 4, and they confirm that there is no evidence of a rise in the probability that a son enters his father's occupation across birth cohorts, and no robust evidence that this probability has increased faster than the probability that a son enters his father-in-law's occupation.[30] Finally, and most importantly, the rate of growth in the probability a woman is in her father's occupation dwarfs any changes that men experienced over the same time period.[31]

6 Conclusion

Labor markets in the 20th century have been profoundly affected by increases in female labor force participation. One potential implication of increased female labor force participation is that it changes the incentives for fathers to invest in their daughters. In particular, it can increase the incentive to invest in human capital that is specific to a father's occupation, causing a rise in the probability that a woman enters her father's occupation.

Simply documenting that there has been an increase over time in the propensity of a woman to enter her father's occupation is not enough to determine whether there has been increased occupation-specific human capital transmission between fathers and daughters. A woman will be more likely to enter her father's occupation even absent such an increase, because she will be more likely to enter any number of traditionally male-dominated occupation, including her father's.

We demonstrate that under the assumption that assortative mating by fathers' occupation has not decreased over time, an assumption for which we find support in our data, we can compare the rates of change over time in the probabilities that a woman enters her father's occupation and her father-in-law's occupation to determine whether there has been an increase in the transmission of occupation-specific human capital from fathers to daughters.

We combine three data sets from the United States containing information collected between

[30] Results available upon request.

[31] There is one caveat to using men as a falsification test of our results. Until this point, our model and discussion have focused on one-child families. If human capital transmission within the family is a purely private good, and if fathers over time invest more in their daughters, they may invest less in their sons. This will itself affect the results in Figure 3, rendering it a flawed falsification test of our model and results for women. The OCG and SIPP do contain data on the number and sex mix of siblings which could potentially be used to examine whether the impact on boys of having sisters has changed over time, but of course family size and sex mix are endogenous.

1973 and 2002 spanning birth cohorts born between 1909 and 1977. We show that over time the probability that a woman enters her father's occupation has increased significantly and substantially. For the full sample of married women, we estimate that with each successive year, the probability that a woman born in a particular year would enter her father's occupation increased by somewhere between 0.22 and 0.44 percentage points. The fraction of women entering their father-in-law's occupation increased anywhere from 0.16 to 0.38 percentage points. Across our many specification checks, the increase in the probability that a woman enters her father's occupation is always larger than the probability that a woman enters her father-in-law's occupation, a finding that we interpret as evidence of increased transmission of occupation-specific human capital between fathers and daughters. For the full sample of women, our results imply that the increase in the probability that a woman enters her father's occupation is around 20 percent higher than the increased probability that she enters another occupation in her choice set, an estimate that is likely a lower bound.

It is natural to speculate as to the form of transmission of occupation-specific human capital from fathers to daughters. Unfortunately, there is not much direct evidence that we know of that helps in this regard. For example, for the case of specific human capital investments, perhaps the most obvious form is investments in time. While there is some information from time-use diaries on how much time parents spend with children, the earliest reliable data come from time-use diaries from 1965, so we have no information on many of the birth cohorts in our sample. There is some evidence that fathers did not spend more time on primary childcare in the mid-1980's than they did in the mid-1960's, but that there has been an increase more recently. Of course, this is an increase that would not be represented in the birth cohorts of our sample. Disappointingly, there is no systematic evidence over time in how this time that fathers spend with children is allocated across daughters and sons, making it difficult to draw any inferences about investments of time in daughters specifically. (See the review in Raley and Bianchi, 2006, for more on what has been learned from time-use surveys on parental time with children.) Dahl and Moretti (2005) present evidence that the ways in which fathers are part of the lives of sons and daughters have changed over time (e.g., via divorce, custody, single parenthood), but these various ways all indicate that fathers have preferences for sons and do not suggest how this manifests itself in terms of changing investments in daughters over time.

We have focused on transmission from fathers to daughters in this paper because for many of the birth cohorts in our sample, the vast majority of their mothers were out of the labor force, and data on the occupations of mothers of adult daughters was not even collected. Therefore, any maternal investments (or transmission) that affected the occupation choice of daughters is difficult to formalize and is perhaps second order to those made by fathers when it comes to changing occupational choice across birth cohorts. However, as recent cohorts of women with high levels of labor force attachment in high(er)-paying occupations themselves become mothers, there should be changing incentives for these women to make investments of occupation-specific human capital in their own daughters (and sons). It will be quite interesting to examine for future cohorts how potentially "competing" transmission of the occupation-specific human capital of fathers and mothers affect the occupation choices of children, and in particular how they affect the occupation decisions of daughters relative to sons.

REFERENCES

Acemoglu, Daron, David Autor, and David Lyle, "Women, War and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid-Century," *Journal of Political Economy*, Vol. 112 (2004), 497-551.

Becker, Gary and Nigel Tomes, " An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, Vol. 87 (1979), 1153-1189.

______ and ______, "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, Vol. 4 (1986), S1-S39.

Chadwick, Laura and Gary Solon, "Intergenerational Income Mobility among Daughters," *American Economic Review*, Vol. 92 (2002), 335-344.

Dahl, Gordon and Enrico Moretti, "The Demand for Sons." *Review of Economic Studies*, forthcoming.

DiPrete, Thomas and David Grusky, "Recent Trends in the Process of Stratification" *Demography*, Vol. 27 (1990), pp. 617-637.

Ermisch, John, Marco Francesconi, and Thomas Siedler, "Intergenerational Social Mobility and Assortative Mating in Britain." *The Economic Journal*, Vol. 116 (2006), 659-679.

Featherman, David and Robert Hauser, *Opportunity and Change.* New York, NY: Academic Press, 1978.

Fernandez, Raquel, Alessandra Fogli, and Claudia Olivetti, "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *The Quarterly Journal of Economics*, Vol. 119 (2004), 1249-1299.

Ferrie, Joseph, "History Lessons: The End of American Exceptionalism? Mobility in the United States Since 1850." *The Journal of Economic Perspectives*, Vol. 19 (2005), 199-215.

Fertig, Angela, "Trends in Intergenerational Earnings Mobility in the U.S.," *Journal of Income Distribution*, Vol. 12 (2003), 108-130.

Galton, Francis, *Natural Inheritance*. London: Macmillan, 1889.

Goldin, Claudia, "The Role of World War II in the Rise of Women's Employment," *American Economic Review*, Vol. 81 (1991), 741-756.

Goldin, Claudia and Lawrence Katz, "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, Vol. 110, (2002), 730-770.

Goldstein, Joshua, and Catherine Kenney, 'Marriage Delayed or Marriage Forgone? New Cohort Forecasts of First Marriage for U.S. Women," *American Sociological Review*, Vol. 66, No. 4 (August, 2001), 506-519.

Grawe, Nathan and Casey Mulligan, "Economic Interpretations Of Intergenerational Correlations," *Journal of Economic Perspectives*, Vol. 16 (2002), 45-58.

Hellerstein, Judith and Guido Imbens, "Imposing Moment Restrictions from Auxiliary Data by Weighting." *The Review of Economics and Statistics*, Vol. 81 (1999), 1-14 .

Lam, David, and Robert Schoeni, "Family Ties and Labor Markets in the United States and Brazil," *Journal of Human Resources*, Vol. 29, no. 4, Fall 1994, pp. 1235-1258.

Lee, Chul-In and Gary Solon, "Trends in Intergenerational Income Mobility," *The Review of Economics and Statistics*, Vol. 91, No. 4 (2009), 766-772 .

Little, Roderick and Mei-Miau Wu, "Models for Contingency Tables With Known Margins When Target and Sampled Populations Differ." *Journal of the American Statistical Association*, Vol. 86 (1991), 87-95.

Mare, Robert, "Five Decades of Educational Assortative Mating." *American Sociological Review*, Vol. 56 (1991), 15-32.

Martin, Steven, "Reassessing Delayed and Foregone Marriage in the United States." *Russell Sage Foundation Working Papers: Series on Social Dimensions of Inequality*, (2004).

Mazumder, Bhashkar, "Fortunate Sons: New Estimates of Intergenerational Mobility in the U.S. Using Social Security Earnings Data," *Review of Economics and Statistics*, Vol. 87 (2005) 235-255.

McFadden, Daniel, "Econometric Models of Probabilistic Choice," in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, 198-272, MIT Press: Cambridge, MA, 1981.

Mincer, Jacob, "Intercountry Comparisons of Labor Force Trends and of Related Developments: An Overview." *Journal of Labor Economics*, Vol. 3 (1985), S1-S32.

Mosteller, Frederick, "Association and Estimation in Contingency Tables." *Journal of the American Statistical Association*, Vol. 63 (1968), 1-28.

Mulligan, Casey, "Galton versus the Human Capital Approach to Inheritance." *The Journal of Political Economy*, Vol. 107 (1999), S184-S224.

Neal, Derek, "Industry-Specific Human Capital: Evidence From Displaced Workers." *Journal of Labor Economics*, Vol. 13 (1995), 653-77.

Raley, Sara and Suzanne Bianchi, "Sons, Daughters and Family Processes: Does Gender of Children Matter?" *Annual Review of Sociology*, Vol. 32 (2006), 401-421.

Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander, Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2008.

Solon, Gary, "Intergenerational Income Mobility in the United States," *American Economic Review*, Vol. 82 (1992), 393-408.

Stevenson, Betsey, and Justin Wolfers, "Marriage and Divorce: Changes and their Driving Forces." *Journal of Economic Perspectives*, Vol. 21 (2), Spring 2007, 27-52.

Weeden, Kim A., and David B. Grusky. "The Case for a New Class Map." *American Journal of Sociology*, Vol. 111 (2005), 141-212.

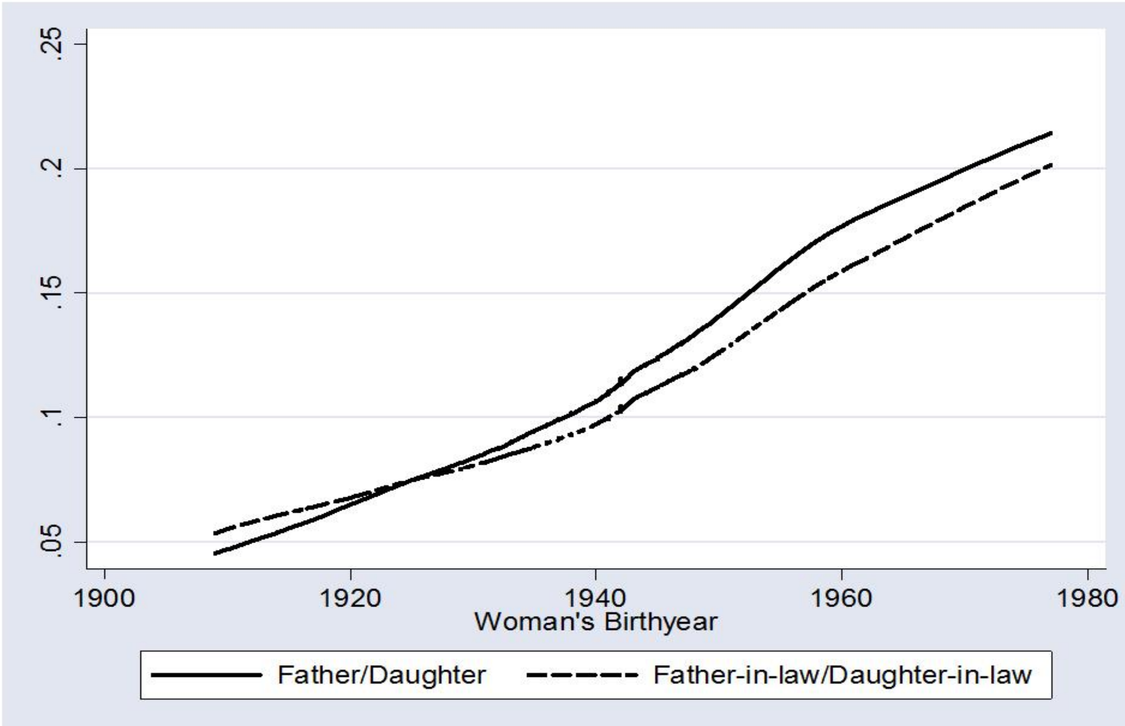


Figure 1A: The fraction of women in the same occupation as their father and their father-in-law

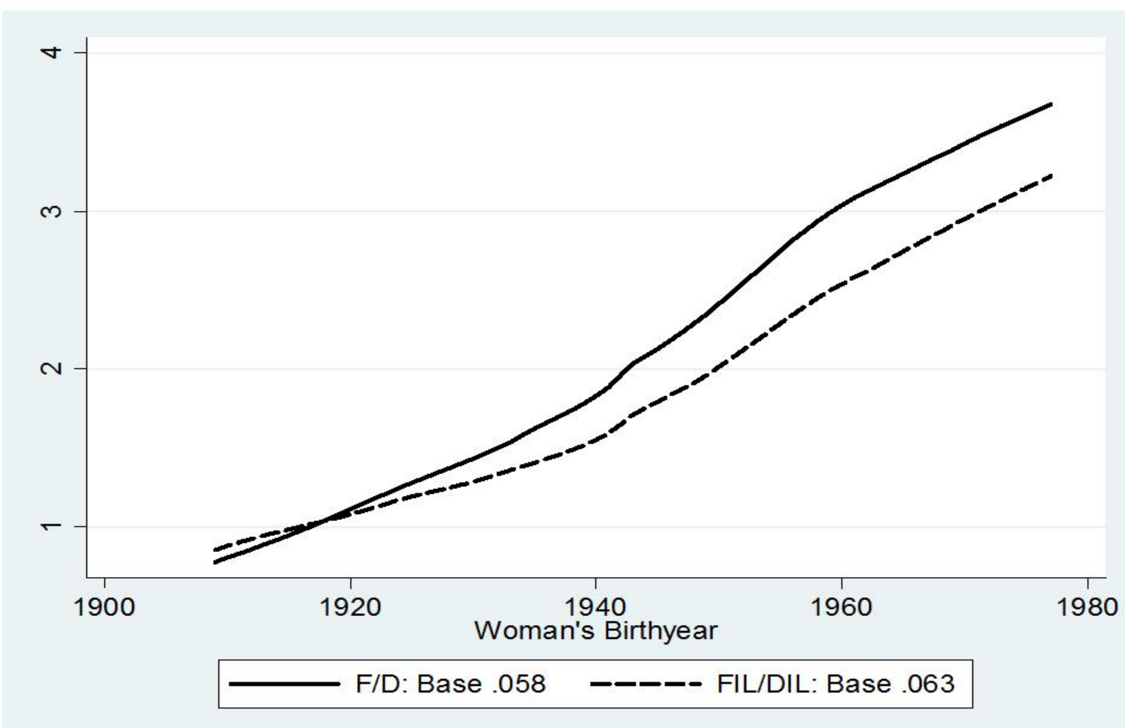


Figure 1B: The rate of growth of the fraction of women in the same occupation as their father and their father-in-law

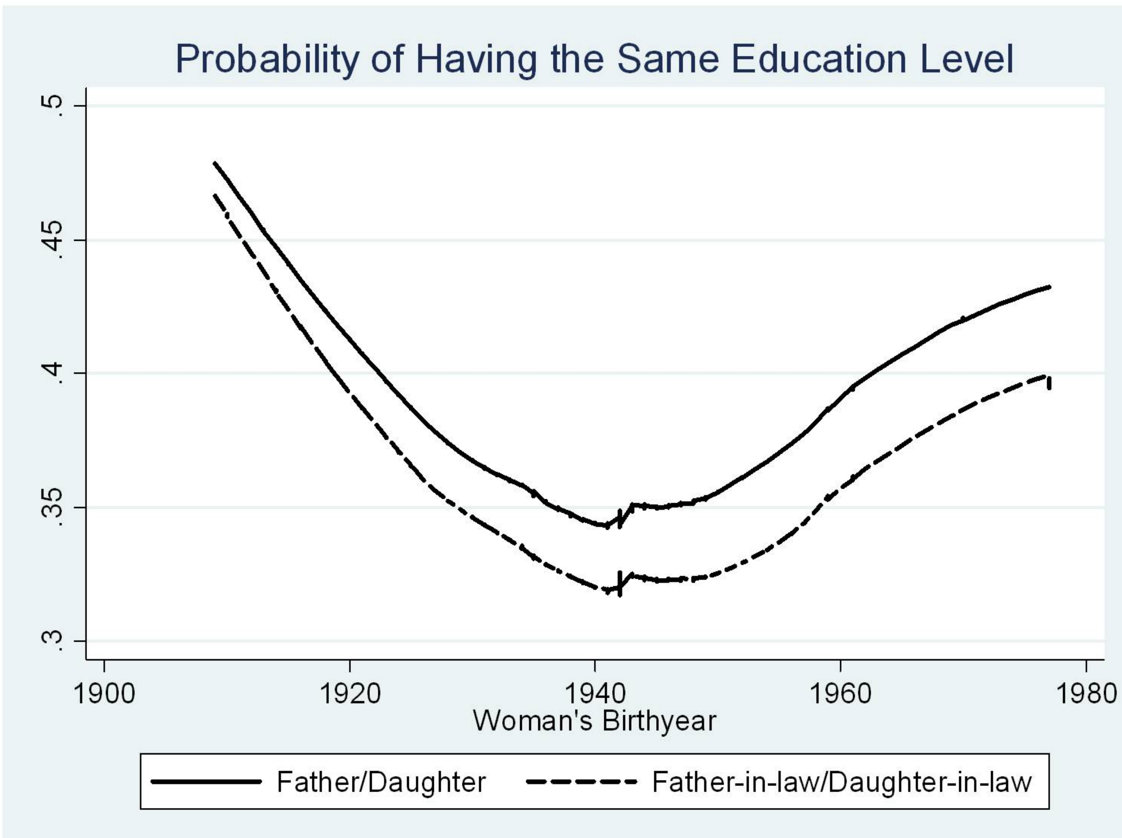


Figure 2: The fraction of women in the same education level (measured in 4 broad categories) as their father and their father-in-law

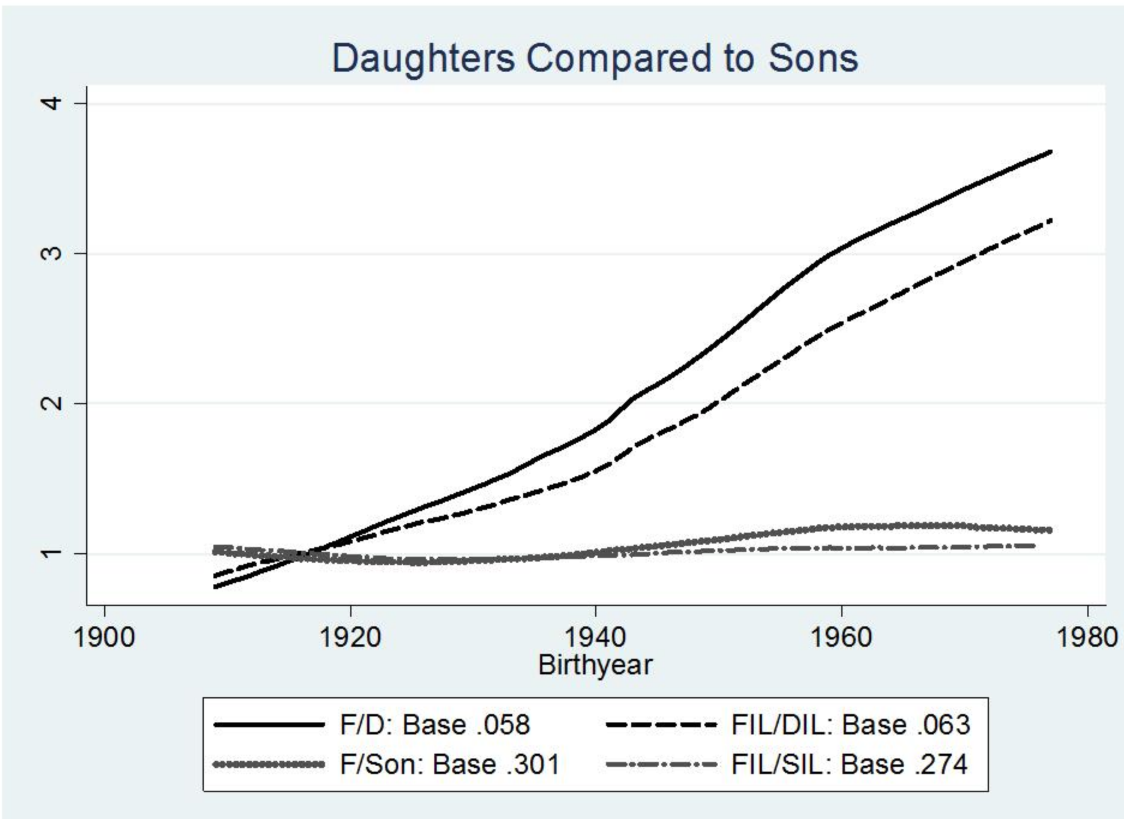


Figure 3: A comparison of the rates of growth of the fraction of women versus men in their father's occupation and father-in-law's occupation

Table 1: Summary Statistics for Women

		ALL	OCG	GSS	SIPP
Women	(1) Managerial and Professional Specialty	0.155	0.103	0.215	0.178
	(2) Technical, Sales, and Admin. Support	0.228	0.174	0.254	0.283
	(3) Service	0.078	0.068	0.082	0.088
	(4) Farming, Forestry, and Fishing	0.009	0.010	0.006	0.010
	(5) Precision Production, Craft, and Repair	0.011	0.008	0.012	0.015
	(6) Operators, Fabricators, and Laborers	0.058	0.067	0.050	0.052
	(7) Not in Labor Force	0.462	0.570	0.381	0.373
Fathers	(1) Managerial and Professional Specialty	0.191	0.180	0.223	0.190
	(2) Technical, Sales, and Admin. Support	0.119	0.096	0.131	0.147
	(3) Service	0.050	0.055	0.040	0.048
	(4) Farming, Forestry, and Fishing	0.195	0.232	0.156	0.157
	(5) Precision Production, Craft, and Repair	0.230	0.224	0.240	0.235
	(6) Operators, Fabricators, and Laborers	0.216	0.213	0.210	0.224
Father-in-Laws	(1) Managerial and Professional Specialty	0.184	0.172	0.225	0.185
	(2) Technical, Sales, and Admin. Support	0.111	0.090	0.117	0.145
	(3) Service	0.048	0.050	0.040	0.048
	(4) Farming, Forestry, and Fishing	0.222	0.263	0.174	0.174
	(5) Precision Production, Craft, and Repair	0.223	0.214	0.242	0.228
	(6) Operators, Fabricators, and Laborers	0.213	0.212	0.202	0.221
Fraction of Women in Father's Occupation		0.107	0.079	0.138	0.134
Fraction of Women in Father-in-Law's Occupation		0.099	0.077	0.120	0.128
Woman's Age		42.100	41.800	42.300	42.300
		(10.9)	(10.7)	(11.0)	(11.2)
Woman's Birth Year		1939.1	1931.2	1946.2	1944.4
		(13.6)	(10.7)	(13.3)	(11.2)
Sample Size of Women		40,360	17,617	11,006	11,737

Table 2: Baseline Results for Probability of Daughters in Same Occupation as Father

Dependent variable: In same occupation as father or father-in-law

	Pooled	OCG	GSS	SIPP
	(1)	(2)	(3)	(4)
Birthyear Daughter	0.268	0.088	0.328	0.305
	(0.013)	(0.019)	(0.031)	(0.028)
	[4.591]	[1.504]	[5.629]	[5.235]
Birthyear Daughter-in-law	0.213	0.065	0.270	0.250
	(0.013)	(0.019)	(0.032)	(0.031)
	[3.399]	[1.031]	[4.310]	[3.990]
Const.	0.005	0.051	-.014	-.0008
	(0.005)	(0.006)	(0.014)	(0.012)
DIL Eqn Dummy	0.013	0.005	0.009	0.015
	(0.006)	(0.007)	(0.020)	(0.016)
Relative Diff F/D vs FIL/DIL	0.055	0.023	0.058	0.055
	(0.017)	(0.023)	(0.045)	(0.037)
	[1.192]	[0.473]	[1.319]	[1.245]
Obs.	63,076	32,700	11,006	19,370

Notes: Standard errors are in parentheses. Rates of increase (relative to baseline 1909-1919 birth cohorts' fraction of women in the same occupation as their father or father-in-law, 0.058 for daughters and 0.063 for daughters-in-law) are in brackets. Results are from linear probability models. Coefficients on birth year, age, and the relative difference in slopes are in percentage point terms. Standard errors are robust and account for correlation across observations that arise from a daughter and daughter-in-law representing the same woman.

Table 3: Full Specification for Women

Dependent variable: In same occupation as father or father-in-law

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Birthyear Daughter	0.218	0.441	0.434	0.435	0.279	0.297	0.242
	(0.014)	(0.042)	(0.055)	(0.040)	(0.038)	(0.040)	(0.038)
	[3.743]	[7.571]	[7.440]	[7.464]	[4.793]	[5.091]	[4.147]
Birthyear Daughter-in-law	0.160	0.371	0.379	0.378	0.234	0.242	0.196
	(0.014)	(0.042)	(0.056)	(0.040)	(0.038)	(0.040)	(0.038)
	[2.554]	[5.924]	[6.053]	[6.037]	[3.737]	[3.860]	[3.127]
Const.	0.034	-.173	-.164	-.166		-.108	
	(0.006)	(0.037)	(0.048)	(0.035)		(0.034)	
SIPP	0.004	0.009	0.004	0.009	0.007	0.010	0.008
	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
OCG	-.023	0.011	0.007	0.011	0.003	0.007	0.002
	(0.004)	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)
DIL Eqn Dummy	0.015	0.028	0.009	0.015		0.014	
	(0.006)	(0.025)	(0.069)	(0.006)		(0.006)	
DIL Dummy*SIPP			0.009				
			(0.008)				
DIL Dummy*OCG			0.008				
			(0.013)				
Daughter's Age		0.247	0.239				
		(0.044)	(0.057)				
Daughter-in-law's Age		0.226	0.235				
		(0.045)	(0.058)				
Constrained D/DIL's Age				0.237	0.190	0.155	0.158
				(0.041)	(0.038)	(0.040)	(0.038)
F/FIL Occ Cntrls	No	No	No	No	Yes	No	Yes
Own Ed Cntrls	No	No	No	No	No	Yes	Yes
Relative Diff F/D vs FIL/DIL	0.058	0.070	0.055	0.057	0.045	0.055	0.046
	(0.017)	(0.031)	(0.079)	(0.017)	(0.016)	(0.017)	(0.016)
	[1.188]	[1.647]	[1.386]	[1.428]	[1.056]	[1.231]	[1.020]
Obs.	63,076	63,076	63,076	63,076	63,076	63,032	63,032

Notes: See Table 2.

Table 4: Robustness Checks

Dependent variable: In same occupation as father or father-in-law

	Prime Age (35-55)		In Labor Force		13 Occupations		15 Industries	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birthyear Daughter	0.455	0.256	0.333	0.079	0.191	0.122	0.224	0.150
	(0.056)	(0.052)	(0.065)	(0.057)	(0.029)	(0.028)	(0.034)	(0.032)
	[6.075]	[3.414]	[2.007]	[0.479]	[5.922]	[3.794]	[4.841]	[3.241]
Birthyear Daughter-in-law	0.399	0.218	0.255	0.030	0.150	0.092	0.182	0.110
	(0.057)	(0.052)	(0.065)	(0.057)	(0.028)	(0.027)	(0.034)	(0.032)
	[5.282]	[2.893]	[1.446]	[0.172]	[4.189]	[2.580]	[3.688]	[2.237]
Const.	-.170		0.015		-.067		-.082	-.087
	(0.049)		(0.057)		(0.025)		(0.029)	(0.028)
SIPP	0.004	0.004	0.004	0.002	0.005	0.005	0.003	0.008
	(0.006)	(0.005)	(0.006)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)
OCG	0.012	-.0007	0.022	0.020	0.004	-.0008	0.013	0.008
	(0.009)	(0.009)	(0.012)	(0.010)	(0.005)	(0.004)	(0.005)	(0.005)
DIL Eqn Dummy	0.013		0.016		0.011		0.016	0.012
	(0.011)		(0.013)		(0.004)		(0.005)	(0.008)
Constrained D/DIL's Age	0.248	0.166	0.105	0.074	0.106	0.087	0.135	0.103
	(0.062)	(0.057)	(0.066)	(0.059)	(0.029)	(0.028)	(0.034)	(0.033)
F/FIL Occ Cntrls	No	Yes	No	Yes	No	Yes	No	Yes
Relative Diff F/D vs FIL/DIL	0.056	0.037	0.078	0.049	0.041	0.030	0.042	0.039
	(0.032)	(0.030)	(0.031)	(0.029)	(0.012)	(0.012)	(0.015)	(0.014)
	[0.793]	[0.521]	[0.562]	[0.307]	[1.733]	[1.214]	[1.153]	[1.004]
Obs.	34,544	34,544	33,242	33,242	63,076	63,076	58,039	58,039

Notes: Standard errors are in parentheses. Rates of increase are in brackets. In columns (1) and (2) the relevant baseline fractions of women in the same occupation as their father or father-in-law, using the 1918-1928 birth cohort, were 0.075 for both the daughter and daughter-in-law. For columns (3) and (4) the baselines were 0.166 for the father-daughter pair and 0.176 for the father-in-law-daughter-in-law. The baselines for columns (5) and (6) were 0.032 and 0.036 for the father-daughter and father-in-law-daughter-in-law respectively. The final two columns used the relevant baseline of the fraction of women in the sample in the same industry as their father or father-in-law, 0.046 and 0.049 respectively. Results are from linear probability models. Coefficients on birth year, age, and the relative difference in slopes are in percentage point terms. Standard errors are robust and account for correlation across observations that arise from a daughter and daughter-in-law representing the same woman.

Table 5: Assortative Mating: Sons-in-Law and Fathers-in-Law

Dependent variable: Man in same occupation as father-in-law.

	Unadjusted	Adjusted	Unadjusted	Adjusted
	(1)	(2)	(3)	(4)
Constant	0.2456	0.3118	0.2217	0.2772
	(0.0657)	(0.0654)	(0.0493)	(0.0497)
OCG	0.0056	0.0013	0.0146	0.0106
	(0.0146)	(0.0148)	(0.0121)	(0.0123)
SIPP	-0.0176	-0.0143	-0.0163	-0.0125
	(0.0081)	(0.0082)	(0.0082)	(0.0082)
SIL Age	0.0001	-0.0004	0.0007	0.0001
	(0.0008)	(0.0008)	(0.0007)	(0.0007)
Birthyear	0.0008	-0.0007	X	X
	(0.0008)	(0.0008)		
Birth Cohort 2	X	X	-0.0036	-0.0087
			(0.0119)	(0.0120)
Birth Cohort 3	X	X	0.0113	-0.0123
			(0.0161)	(0.0162)
Birth Cohort 4	X	X	0.0309	-0.0150
			(0.0214)	(0.0215)
Birth Cohort 5	X	X	0.0475	-0.0089
			(0.0273)	(0.0275)
Birth Cohort 6	X	X	0.0562	-0.0102
			(0.0325)	(0.0325)
Obs.	28,086	28,086	28,086	28,086

Notes: Standard errors are in parentheses. Odd numbered columns show results from OLS. Even numbered columns show results from weighted regressions, where the weights adjust the distribution of occupations of fathers-in-law in each cohort to be equal to the mean across all cohorts.

A Data Appendix

A.1 Core Sample Description

In our sample, we exclude respondents who are younger than 25 years old or who report being married to someone under age 25. While we exclude women who are older than 64 years in the regressions (and make similar restrictions for men in the male sample regressions), we do include women who are married to men older than 64. Restricting to men and women older than 25 helps control for data quality, since occasionally children are incorrectly coded as spouses. The results we report throughout the paper are insensitive to restricting to *couples* that are between 25-64 years old, but relaxing the upper bound restriction allows for a larger sample. One effect of this asymmetric treatment is that the male sample is slightly smaller than the female sample, since men tend to marry women who are younger. All of our results are for married, white individuals who report being either the head of household or the spouse of the head of household.

A.2 Labor Force Participation and Occupations

Appendix Table A.1 describes how we define who participates in the labor force and who is dropped from the sample across each data set. As described in the text, we consider women who have decided not to work as a separate occupation "out of the labor force" (OLF). This category includes women who are "keeping house," as the OCG and GSS categorize them. The OLF category should not include women who are unemployed, looking for work, in school, or doing something else that is distinct from choosing to remain out of the labor force. Note that we never include OLF for fathers or fathers-in-law, since we do not have an employment status code for fathers and therefore can not distinguish between item non-response and a non-working father.

The SIPP is distinct from the OCG and GSS in how employment status is coded. In the SIPP, work status is asked separate of school enrollment or other activities. A respondent could be coded as having a job and being enrolled in school, while the GSS and OCG only report one employment status per individual. Since someone enrolled in school, either part-time or full-time, is likely to not be in their final occupation even if a valid occupation code is given, we restrict the SIPP sample to include only individuals who are not currently enrolled in school.

In order to get some sense of how comparable the data are across surveys and how they reflect general trends, we examine female labor force participation by birth cohort in each data set. We provide a graph in Appendix Figure 1 that shows the fraction of women who were employed in each year for each survey. Because we treat "out of the labor force" as an occupation in itself (one that daughters do not, by our definition, ever share with their fathers), it is important to examine female labor force participation rates in the context of occupation transmission between fathers and daughters.

We do not expect our data sets to provide identical female labor force participation rates for each birth cohort because of age effects. We therefore also graph female labor force participation rates by birth cohort for the 1970-2000 Decennial Census Public Use Micro Samples (PUMS), four nationally representative data sets drawn from years similar to our three data sets. Our samples consist of married, white women between the ages of 25 and 64 who report that they are not in school and are either working or out of the labor force (we exclude "unemployed" women and women in school).[32] We also restrict our attention to women who are either the head of household or the spouse of the head of household.

It is useful to begin by comparing data from the PUMS samples. For the birth cohorts that overlap between the samples, it is clear that overall female labor force participation rates increased over time. Looking at birth cohorts separately, one can see that for earlier birth cohorts there is exit out of the labor force as women age toward retirement while for later birth cohorts female labor force participation clearly increased over the decades. For all four data sets, a dip in female labor force participation exists for women in their 30's, presumably as a result of child-rearing. The changing labor force participation rates of women through their lifetimes foreshadows the importance of controlling for age in our analysis of intergenerational occupation between fathers and daughters.

Data from the GSS surveys of 1975-2002 provide the longest time period over which to examine labor force participation by birth cohort. The GSS spans the data from the SIPP and OCG, and nearly spans our Census years as well. As the graph in Appendix Figure 1 indicates, the GSS labor force participation rates do cut through those of the other data sets and rise from well below 20 percent for the birth cohorts early in the 20th century to well above 60 percent for women born in the 1960s and thereafter.

The OCG contains information on the labor force participation in 1973 of women born between 1909 and 1948. Average labor force participation of women in the OCG lies between the 1970 and 1980 PUMS graphs, as it should. Similarly, the SIPP profile of female labor force participation, derived from data collected between 1986 and 1988, is between the two PUMS profiles from 1980 and 1990, and is closer to 1990, as would be expected. In total, female labor force participation in our data reflects that seen in PUMS data, and across our three data sets the trends in female labor force participation over time by birth cohort are consistent with age effects of retirement and child-rearing.

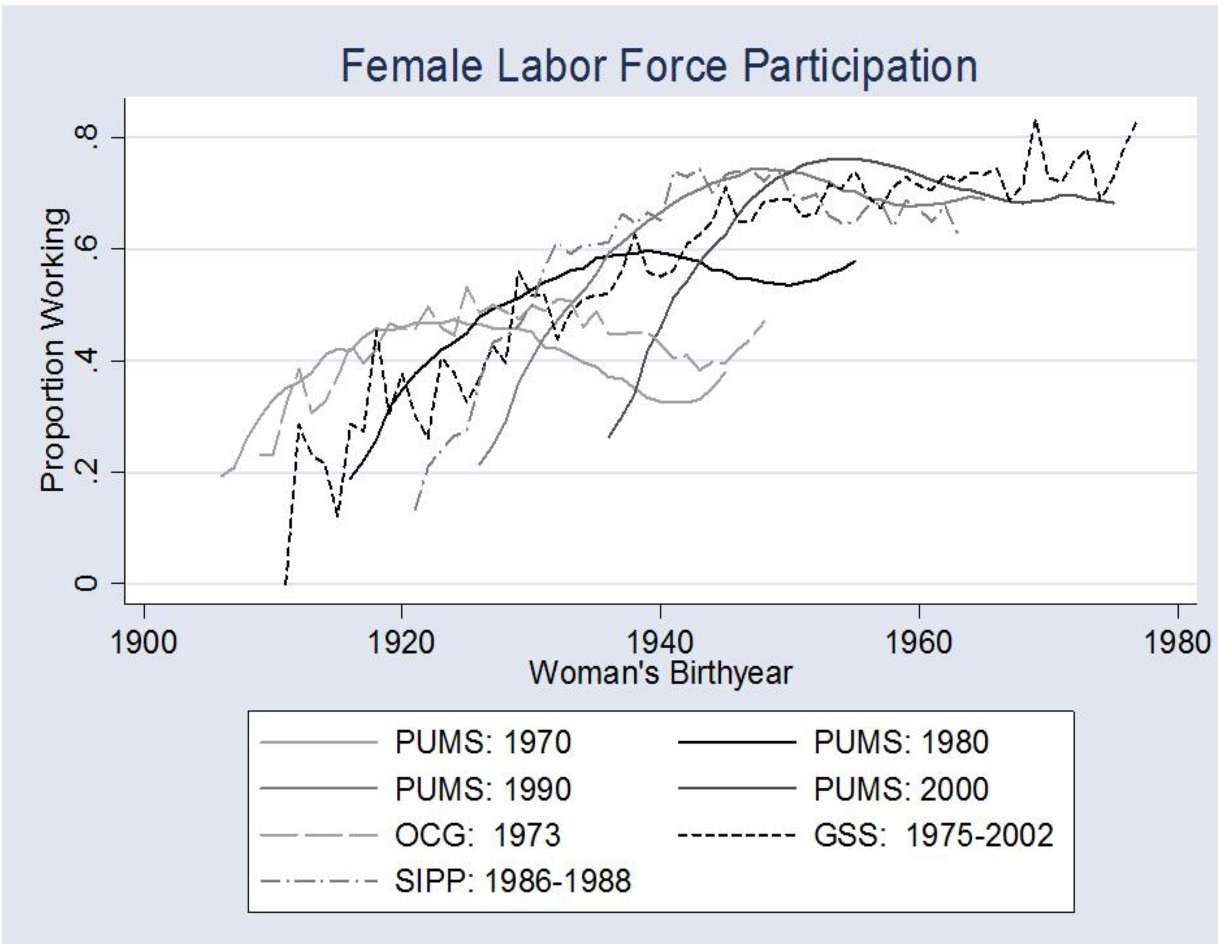
A.3 Occupation Coding and Concordances

Because our three surveys contain different occupation classifications, we had to create a consistent definition of "occupation." Occupation codings change across decennial censuses. Though the occupation codings tend to be similar across most census years, they are not identical. The 1973 OCG reports 1970 (and 1960) Census Occupation Codes, while the SIPP reports 1980 Occupation Codes. The GSS, on the other hand, uses 1970 codes for survey years 1975-1990 and 1980 codes for survey years 1988-2002. This provides us with 3 survey years (1988, 1989, and 1990) for which both 1970 and 1980 occupation codes are included. To get a consistent definition of occupation we created a concordance from the 1970 to 1980 Census Occupation Codes using the GSS.

To create the concordance we take the 1980 occupation code that is most frequently matched to each 1970 occupation code, choosing the smallest code by default in a tie. Once we have this mapping from 1970 to 1980, we merge the 1980 occupation codes onto the early years of the GSS with only 1970 occupation codes and onto the OCG. Tests of the sensitivity to using categorizations of the 1970 and of the 1960 codes provided consistent results.

Appendix Table A.2 lists the occupation groupings used in our analysis.

[32] Note that the PUMS definition of labor force participation is closest to that of the SIPP.



Appendix Figure 1: Female Labor Force Participation by Birth Cohort

Appendix Table A1: Description of Labor Force Definitions

	OCG	GSS	SIPP	PUMS
Question wording	"What was _ doing most of LAST WEEK?"	"Last week were you working full time, part time, going to school, keeping house, or what?"	Work status: Month 4; School: During any of the past 4 months	Employment Status Recode (empstatg): Previous week; School attendance: During the past 2 months
Possible Responses	(1) Working (2) With a job, but not at work (3) Looking (4) Housework (5) School (6) Unable to work (7) Other	(1) Working full time (2) Working part time (3) With a job, but not at work b/c of temp illness, vacation, or strike (4) Unemployed (5) Retired (6) In School (7) Keeping House (8) Other (9) No answer	With a job the entire month: (1) worked all weeks (2) missed one or more weeks, no time on layoff (3) missed one or more weeks, spent time on layoff With a job one or more weeks: (4) no time spent looking or on layoff (5) spent one or more weeks looking or on layoff No job during month: (6) spent entire month looking or on layoff (7) spent one or more weeks looking or on layoff (8) no time spent looking or on layoff	Employment Status Codes: (1) Employed (2) Not Employed (3) Not in Labor Force
Women	Working: (1) or (2) Out of Labor Force: (4) or (7) Dropped: (3), (5), (6)	Working: (1), (2), or (3) Out of Labor Force: (5), (7), or (8) Dropped: (4), (6), or (9)	If not enrolled in school: Working: (1) - (5) Out of Labor Force: (8) Dropped: (6) or (7)	If not enrolled in school: Working: (1) Out of Labor Force: (3) Dropped: (2)
Men	Working: (1) or (2) Dropped: (3) - (7)	Working: (1), (2), or (3) Dropped (4)-(9)	If not enrolled in school: Working: (1) - (5) Dropped: (6) - (8)	N/A

Appendix Table A2: 1980 Census Occupation Code Groupings

1980 Census Occupation Codes	
Six Occupation Categories	**13 Occupation Categories**
(1) Managerial and Professional Specialty	(1) Executive, Administrative, and Managerial Occupations and Management Related Occupations (2) Professional Specialty Occupations
(2) Technical, Sales, and Administrative Support	(3) Technologists, Technicians and Related Support Occupations" (4) Sales Occupations (5) Administrative Support Occupations, Including Clerical
(3) Service	(6) Service Occupations, Private Household Occupations (7) Protective Service Occupations (8) Service Occupations, Except Protective and Household
(4) Farming, Forestry, and Fishing	(9) Farming, Forestry, and Fishing Occupations
(5) Precision Production, Craft, and Repair	(10) Precision Production, Craft, and Repair Occupations
(6) Operators, Fabricators, and Laborers	(11) Machine Operators, Assemblers and Inspectors (12) Transportation and Material Moving Occupations (13) Handlers, Equipment cleaners, Helpers, and Laborers"