Centre for Economic Policy Research

BANCO DE ESPAÑA
Eurosistema

European Summer Symposium in International Macroeconomics (ESSIM) 2008

Hosted by
Banco de España

Tarragona, Spain; 20-25 May 2008

Culture as Learning: The Evolution of Female Labor Force Participation over a Century

Raquel Fernandez

We are grateful to the Banco de España for their financial and organizational support.

The views expressed in this paper are those of the author(s) and not those of the funding organization(s) or of CEPR, which takes no institutional policy positions.

Culture as Learning: The Evolution of Female Labor Force Participation over a Century*

Raquel Fernández
New York University, CEPR, NBER, IZA

Revised November 2007†

Abstract

Married women's labor force participation increased dramatically over the last century. Why this occurred has been the subject of much debate. This paper investigates the role of changes in culture arising from learning in generating this increase. To do so, it develops a dynamic model of culture in which individuals hold heterogeneous beliefs regarding the relative long-run payoffs for women who work in the market versus the home. These beliefs evolve rationally via an intergenerational learning process. Women are assumed to learn about the long-term payoffs of working by observing (noisy) private and public signals. This process generically generates an S-shaped figure for female labor force participation, which is what is found in the data. The S shape results from the dynamics of learning. I calibrate the model to several key statistics and show that it does a good job in replicating the quantitative evolution of female LFP in the US over the last 120 years. The model highlights a new dynamic role for changes in wages via their effect on intergenerational learning. The calibration shows that this role was quantitatively important in several decades.

JEL Nos.: J16, J21, Z1, D19. Keywords: female labor force participation; cultural transmission; preference formation; learning; S shape; social norms.

†An earlier version of the model and simulation in this paper were presented in my Marshall Lecture at the EEA, Vienna, August 2006. The slides for this presentation are available at http://homepages.nyu.edu/~rf2/Research/EEAslidesFinal.pdf (pp 48-52).

*I thank Liz Potamites for excellent research assistance, and Christophe Chamley, John Knowles, Oriana Bandiera, Gianluca Violante, Elisabeth Schulte, and George Tridimas for helpful comments. I also wish to thank seminar audiences at the LAEF "Households, Gender and Fertility" conference, the NY/Philadelphia feds. "Quantitative Macroeconomics" Workshop, the NBER Summer Institute, the Silvaplana Political Economy workshop, the "Family Behavior and the Aggregate Economy" SITE workshop, LACEA, and Columbia U. for many helpful remarks. Lastly, I thank the NSF for financial support and the Russell Sage Foundation for its hospitality.

1 Introduction

A fundamental change over the last century has been the vast increase in female labor force participation. In particular, married women's participation in the formal labor market increased dramatically–from around 2% in 1880 to over 70% in 2000–though the pace of change was markedly uneven. As shown in figure 1, married women's labor force participation increased very slowly from 1880 to 1920, grew a bit more rapidly between 1920 and 1950, then accelerated between 1950 and 1990, and has since stayed relatively constant.[1]

Many explanations have been given for this transformation. Depending on the particular time period under consideration, potential causal factors have included structural change in the economy (the rise of the clerical sector), technological change in the workplace and in the household, medical advances (including the introduction and dissemination of the oral contraceptive), decreases in discrimination, institutional changes in divorce law, and the greater availability of childcare.[2]

A popular alternative explanation (though not with economists) is that changes in culture or social norms have exerted great influence on the evolution of women's role in the market work.[3] And, from multiple sources of evidence, it certainly appears that opinions about the role of women in the workplace have changed radically over time. Figure 2, for example, shows the evolution of the percentage of the population that answered affirmatively to the question "Do you approve of a married woman earning money in business or industry if she has a husband capable of supporting her?"[4] In 1936 fewer than 20% of individuals sampled agreed with the statement; in 1998 fewer than 20% of individuals disagreed with it.[5]

Merely pointing to the fact that society has changed the way in which it regards women, however, is not particularly enlightening. It begs the question as to why culture changed and

[1] These LFP numbers were calculated by the author from the US Census for white, married women between the ages of 25-44, born in the US, not in agriculture, non-farm, non-institutional quarters.

[2] The classic source for an economic history of female labor force participation is Goldin (1990). For various explanations for this change see, among others, Goldin (1990), Galor and Weil (1996), Costa (2000), Goldin and Katz (2002), Jones, Manuelli, and McGrattan (2003), Greenwood, Seshadri, and Yorukoglu (2005), Gayle and Golan (2006), Albanesi and Olivetti (2006, 2007), and Knowles (2007).

[3] The reluctance of economists to believe in cultural explanations stems, in large part, from the absence of empirical evidence that convincingly isolates cultural influences from their economic and institutional environment. There has been recent progress in this area, however (see Fernández (2007a) and Guiso, Sapienza, and Zingales (2006) for partial reviews of this literature). For example, Fernández and Fogli (2005) show that the variation in the work behavior of second-generation American women can be explained, in part, by the level of female LFP in their parents' country of origin (see also Antecol (2000)). Moreover, Fernández (2007b) shows that the attitudes towards women's work in the parental country of origin has important explanatory value for second-generation American women's work behavior in the US (see also Burda, Hamermesh, and Weil (2007) who show that these attitudes help explain married men and women's work decisions inside and outside the home). These papers show that cultural differences across societies matter for women's work decisions, but they are silent on the evolution of culture. Fernández, Fogli, and Olivetti (2004) give an indication for one way that culture may evolve over time by showing that working mothers seem to transmit a different set of beliefs or preferences to their sons, which then makes it more attractive for the wives of these men to work (relative to the wives of men whose mothers did not work).

[4] The exact wording of this question varied a bit over time. See The Gallup Poll; public opinion, 1935-1971.

[5] For additional evidence that individual attitudes and work behavior are correlated see, for example, Levine (1993), Vella (1994) , Fortin (2005), and Farré-Olalla and Vella (2007).

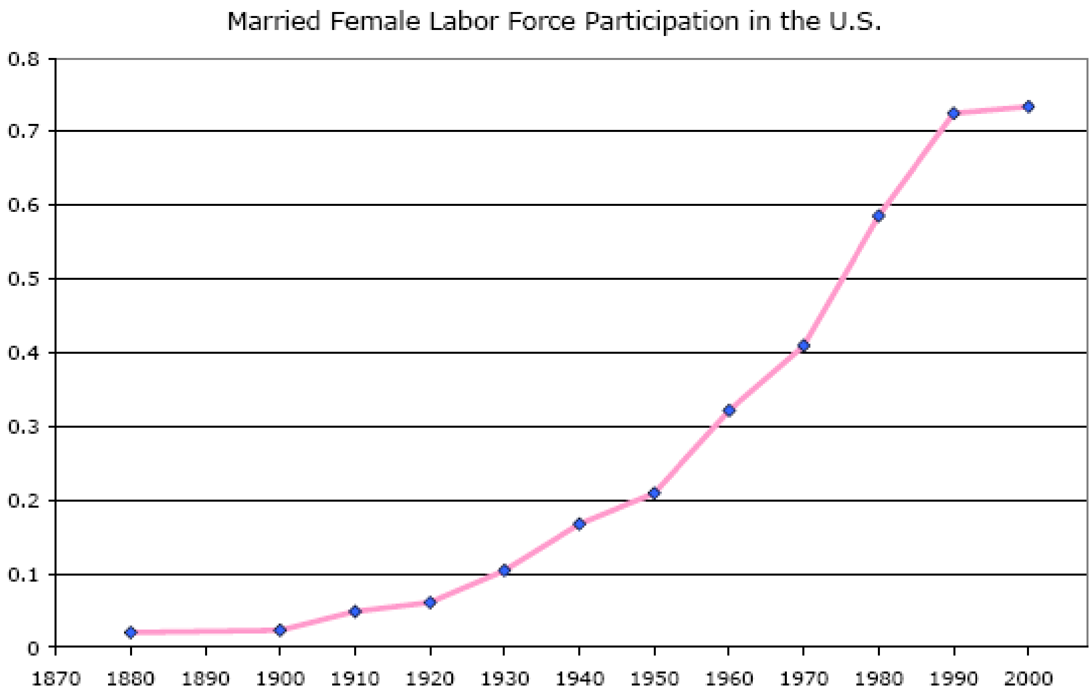

Figure 1: U.S. Census data 1880-2000. Percentage of white, married (spouse present) women born in the U.S., 25-44 years old (non-agricultural, non-group quarters), who report being in the labor force.

why these changes affected work behavior in such a gradual and uneven fashion. Indeed, one might be tempted, as surely some are, to dismiss the evolution of beliefs as mere changes in the superstructure of the economy that simply accompany and reflect the changes in material conditions brought about by technological change.[6] Viewed from this perspective, as technological advances altered women's work behavior, beliefs simply marched right along in step and changed with them. An alternative view of culture often provided in economic theory—that of a selection mechanism among multiple equilibria—likewise does not provide a very useful framework in which to think about questions of cultural change. Without a more developed theory of why culture changes, one is left with either sunspots causing a switch among equilibria or an evolutionary theory of gradual changes over time.[7]

Taking inspiration from the fact that women's labor force participation changed in a very uneven fashion over time in a form that resembles an "S-shape", this paper explores the idea that in some contexts it may be useful to think about cultural change as the evolution of beliefs that occurs over time as part of a rational, intergenerational *learning* process.[8]

[6] See, e.g., Guner and Greenwood (2006) who argue that the change in sexual mores reflect changes in the efficacy of contraception.

[7] For an interesting example of evolutionary theory applied to culture see Bowles (1998). Alternatively, social norms can be passed on from parents to children in an optimizing fashion as in Bisin and Verdier (2000) and Tabellini (2007).

[8] The idea that cultural change may be modelled as a learning process is already present in the seminal paper of Bikhchandani, Hirshleifer, and Welch (1992), though the focus there is on information cascades in which individuals stop learning.

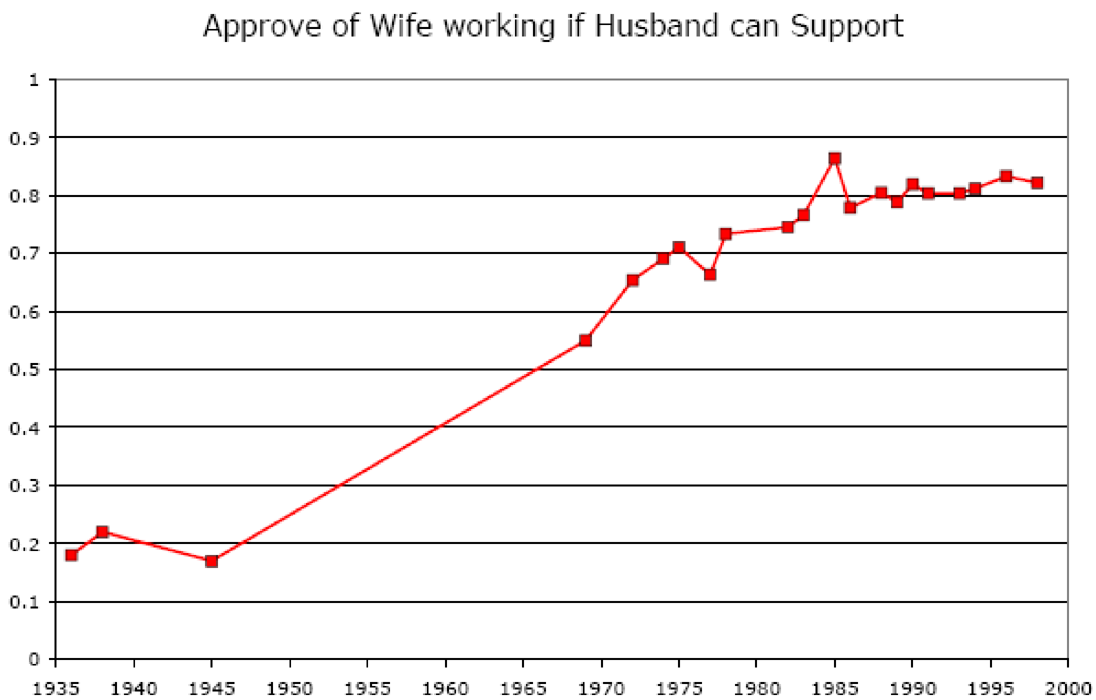

Figure 2: Sources: 1936-1938 and 1969 numbers are from the Gallup Poll (1972), 1945 is from Benjamin I. Page and Robert Y. Shapiro, The Rational Public, University of Chicago Press, 1992; pp. 101, 403-4. 1972 onwards are from the General Social Survey.

In particular, the S-shaped curve of female labor force dynamics is reminiscent of similarly shaped curves that are common in the process of technology adoption and may constitute an important clue that a similar mechanism of information diffusion is also at play in this context, though on a very different time scale.[9]

Where might learning play a role in the transformation of women's work? It is not an exaggeration to state that, throughout the last century, women's work has been a subject of great contention. As industrialization and urbanization progressed over time, so did specialization. Younger men and (unmarried) women were drawn into the paid workplace and away from sharing household chores, and the spheres of work and home became increasingly separate. This process left the wife in charge of the domestic realm and her husband in charge of supporting the family, and kicked off a debate on the effect of a wife working (outside the home) on her family and marriage as well as on her psyche and image (and on those of her husband's) that continues, in different guises, to this day.[10] For example, as noted by Goldin (1990), at the turn of the 20th century most working women were employed as domestic servants or in manufacturing. In this environment, a married woman's employment signalled that her husband was unable to provide adequately for his

[9]There is a large literature on learning and technology adoption. See, for example, Griliches (1957), Foster and Rosenzweig (1995), Conley and Udry (2003), Munshi (2004), Munshi and Myaux (2006), and Bandiera and Rasul (2006). See Chamley (2004) for a review of this literature.

[10]See Goldin (1990) for a very interesting account of this process of separation and specialization.

family and, consequently, most women exited the workplace upon marriage.[11] Over time, the debate shifted to the effect of a married woman working on family stability and to the general suitability of women for various types of work and careers. More recently, public anxiety regarding working women centers around the effect of a working mother on a child's intellectual achievements and emotional health [12] For example, a recent finding by Belsky et al. (2007) of a positive relationship between day care and subsequent behavioral problems became headline news all over the US. Thus, throughout the last century the expected payoff to a woman working has been the subject of an evolving debate.

In this paper I develop a simple model of women's work decisions in which beliefs about the (long-run) payoff to working evolve endogenously over time.[13,14] Using a framework broadly similar to Vives (1993) and Chamley (1999), I assume that women possess a private signal about how costly it is to work (e.g., how negative the outcome is for a woman's marriage, children, etc.) and that they also observe a noisy public signal indicatory of past beliefs concerning this value. This signal is a simple linear function of the proportion of women who worked in the previous generation and is equivalent to observing a noisy signal of the average utility of working women in the past. Women use this information to update their prior beliefs and then make a decision whether to work. In the following period, the next generation once again observes a noisy public signal generated by the decisions of women in the preceding generation, each woman obtains her individual private signal (or equivalently inherits that of her mother's), and makes her work decision. Thus, beliefs evolve endogenously via a process of intergenerational learning.

The model described above *generically* generates an S-shaped figure for female labor force participation. The S shape results from the dynamics of learning. When very few women participate in the labor market (as a result of initial priors that are pessimistic about the payoff from working), learning is very slow since the noisiness of the signal swamps the information content given by small differences in the proportion of women who would work in different states of the world. As the proportion of women who work increases and beliefs

[11] Over 80% of married women, not employed in 1939 but had worked at some point prior to marriage, exited the workplace at the precise time of marriage. These numbers are cited in Goldin (1990, p. 34) from the 1939 Retrospective Survey.

[12] See, for example, Bernal (2007), Keane and Bernal (2005) and Ruhm (2006) for reviews and recent findings of this literature.

[13] Whether preferences or beliefs changed is often impossible to distinguish and, in a reduced-form setup, it is also unnecessary. The assumption that changes in beliefs were driven by learning is important, however, as Bayesian updating thus constrains the path taken by beliefs. An additional advantage of this modelling choice is that is straightforward to think about social welfare, which is not the case if preferences themselves are affected (see Fernández (2007a) for a discussion of these issues).

[14] A recent paper by Fogli and Veldkamp (2007) independently develops a related idea. They study the labor force participation of women with children from 1940-2000 and assume that women learn about the ability cost to a child from having a working mother. Learning occurs through sampling the ability outcomes of a small number of other women. Whereas in my model actions change because people modify their beliefs about the cost of working, in their model beliefs change only because of a reduction of uncertainty about the cost. Also related is Munshi and Myaux (2006) who model the change in contraceptive practice in rural Bangladesh as learning about the the preferences of individuals in one's social network. They too use a sampling model but there is, in addition, a strategic aspect to individual choices since an agent's payoff depends on the contraceptive choices of the other individual sampled. Lastly, Mira (2005) examines the links between fertility and infant mortality in a model which mothers are learning about a family-specific component of infant mortality risk.

about work become more positive, the information content in the signal improves. Once a large enough proportion of women work though, once again, the informational content in the public signal falls since the differences in the proportion of women who would work under different states of the world is small and thus swamped by the noise.

The model also introduces a new role for changes in wages or technological change, which to my knowledge has not been noted in the learning literature. Unlike in traditional models, increases in women's wages or new technologies that make it easier for women to work outside the home, have not only a static effect of making work more attractive and thereby increasing female LFP, but they also have a dynamic effect since they affect the informativeness of the public signal and hence the degree of intergenerational updating of beliefs.[15] In particular, when the average woman is pessimistic about the payoff to women's work, increasing the attractiveness of work improves the informativeness of the public signal by moderating the private signal that she requires in order to be willing to work.

To evaluate the ability of such a model to explain the quantitative evolution of female LFP, I first calibrate a version of the model without any learning to a few key statistics for the year 2000. I show that such a model performs very badly and that it grossly overestimates the proportion of women who would have worked for basically every time period. I then introduce learning as discussed above, calibrate the model incorporating additional statistics, and show that introducing learning greatly improves the capacity of the model to replicate the historical path of female LFP.

The calibrated model indicates that the paths of both beliefs and earnings played important roles in the transformation of women's work. In the decades between 1880-1950 the growth in female LFP was small, and most of the change in LFP was the result of changes in wages. From 1950 to 1970, both the dynamic and static effects of wage changes played a role in increasing female LFP, and from 1970 to 1990 the dynamic effect on beliefs of changes in earnings is critical in accounting for the large increase in the proportion of working women over that time period.

The paper is organized as follows. Section 2 presents a simple model of a woman's work decision in which the dynamics is generated by changes in wages. The next section introduces beliefs and learning into the simple model and explains why the intergenerational evolution of beliefs naturally generates an S-shaped curve for LFP. Section 4 calibrates the model with and without learning and decomposes the changes in LFP into a beliefs component, a static wage component, and a dynamic wage-belief component. Section 5 discusses the roles of various assumptions and concludes.

2 A Simple Model of a Woman's Work Decision

I start with a very simple model of a woman's work decision that depends on the two main variables that are typically assumed to play a role, namely her consumption possibilities as a

[15] Of course, changes in wages may have dynamic effects by changing borrowing constraints, parental education, schooling choices, etc. The point that is being emphasized here is that they have an additional dynamic effect in the learning model as they will also change the informativeness of the public signal.

function of her decision and her disutility from working. As I am interested in the difference in the *long-run* payoffs from working versus not working, the disutility from working should be viewed as arising not only from labor-leisure preferences, but also from what might happen to a woman's identity, marriage, or her children as a result of her decision. In this first model, I assume that the difference in disutility is known and constant. What is critical though is that its expected value does not evolve endogenously over time; whether it is known for sure is otherwise irrelevant.

A woman makes her work decision to maximize:[16]

$$U_i(w_f, w_h, v_i) = \frac{c^{1-\gamma}}{1-\gamma} - \mathbf{1}v_i \tag{1}$$

where $\gamma \geq 0$ and $\mathbf{1}$ is an indicator function that takes the value one if she works and zero otherwise. A woman's consumption is the sum of her earnings, w_f, (which are positive only if she works) and her husband's earnings, w_h. Husbands are assumed to always work, i.e.,

$$c = w_h + \mathbf{1}w_f \tag{2}$$

The disutility of work, v_i, is assumed to consist of two parts,

$$v_i = \beta + l_i \tag{3}$$

where the first component β is common to all women and the second component is idiosyncratic and normally distributed, $l \sim N(0, \sigma_l^2)$.

Clearly, a woman will work iff

$$\frac{1}{1-\gamma}[(w_{ht} + w_{ft})^{1-\gamma} - w_{ht}^{1-\gamma}] - \beta \geq l_i \tag{4}$$

and thus, assuming that there is a continuum of agents of mass one in each period, the aggregate number and proportion of women who work at time t is given by

$$\omega_t = G\left(l_t^*; \sigma_l\right) \tag{5}$$

where $G(\cdot)$ is the cdf of the l distribution and l_t^* is the value of l such that (4) is a strict equality.

Note that in this simple model, the dynamics of female labor force participation is determined entirely by the dynamics of earnings. As earnings evolve, so does l^*. In particular, women's LFP is increasing in their own earnings, i.e., $\frac{\partial l^*}{\partial w_f} > 0$, whereas it is decreasing in their husbands' earnings, $\frac{\partial l^*}{\partial w_h} < 0$.

[16] We consider only the extensive margin, i.e., she either works or not.

3 The Simple Work Model with Learning

We next incorporate beliefs and learning in the simple model above. Women are assumed to be uncertain about the common value of the disutility of labor, β, e.g., they are unsure how bad working will be for their marriage, children, identity, etc. This is not something that can be learned by entering the labor market for a short period of time nor by experimentation, but rather reveals its effects over a lifetime.

For simplicity, we assume that β can take on only two values, high (H) and low (L), i.e., $\beta \in \{\beta_H, \beta_L\}$.[17] Note that β_L is the good state of nature in which working is not so costly, i.e., $\beta_H > \beta_L \geq 0$. An individual woman now makes her work decision to maximize her expected utility, i.e., equation (1) is modified to reflect uncertainty about the payoff to working:

$$\frac{c^{1-\gamma}}{1-\gamma} - \mathbf{1}(E_{it}v_i) \tag{6}$$

where E is the expectations operator and $E_{it}v_i = E_{it}(\beta) + l_i$.

The model incorporates two sources of learning. One is a private signal regarding the true value of β, β^*. The second is a public intergenerational signal of the decisions taken by women in the preceding generation. It is the latter social source of learning that is key. The exact mechanics are made more precise below.

Consider a woman in period t who has a prior belief about β^* as summarized in the log likelihood ratio (LLR) $\lambda_t = \ln \frac{Pr(\beta^*=\beta_L)}{Pr(\beta^*=\beta_H)}$. Prior to making her work decision, she receives a private signal s_{it} regarding β^*. This signal can be thought of as arising from many sources (e.g., the scientific literature that existed at that time regarding the effect of a woman working) and can be either newly generated each period or inherited from the woman's mother.[18] The private signal is given by:

$$s_{it} = \beta^* + \epsilon_{it} \tag{7}$$

where $\epsilon \sim N(0, \sigma_\epsilon^2)$ and its cumulative and probability distribution functions are denoted by $F(\cdot;\sigma_\epsilon)$ and $f(\cdot;\sigma_\epsilon)$, respectively.[19] The private signals are assumed to be iid across women.

After receiving (or inheriting) her private signal, s, each woman i updates her prior belief accordingly using Bayes' rule, resulting in a new LLR, $\lambda_{it}(s)$, given by

$$\begin{aligned} \lambda_{it}(s) &= \lambda_t + \ln\left(\frac{Pr(s|\beta^*=\beta_L)}{Pr(s|\beta^*=\beta_H)}\right) \\ &= \lambda_t - \left(\frac{\beta_H-\beta_L}{\sigma_\epsilon^2}\right)\left(s-\bar{\beta}\right) \end{aligned} \tag{8}$$

[17] Alternatively, one can think of individuals obtaining an ex-post realization β_i of a random variable with a mean equal to either β_H or β_L. Individuals would thus be learning about the true mean over time (hence even if one were able to observe an individual realization of β, it would convey little information about the benefits of working).

[18] In the calibration of the model we use the latter interpretation.

[19] The results do not depend on ε being normally distributed. Rather, as will be made clear further on, one requires a cdf that changes slowly, then rapidly, and lastly slowly again.

where $\bar{\beta} = (\beta_L + \beta_H)/2$.[20] Note that $\frac{\partial \lambda_{it}(s)}{\partial s} < 0$ since observing higher values of s increases the likelihood that the true value of β is β_H. Note also that the revision of λ is decreasing with the variance of the noise term, $\sigma^2_\in$, since it lowers the informativeness of the signal.

Assume that women have a common prior in period t, λ_t.[21] What proportion of women will choose to work that period? A woman will work in period t iff

$$\frac{1}{1-\gamma}[(w_{ht} + w_{ft})^{1-\gamma} - w_{ht}^{1-\gamma}] - E_{it}(\beta) \geq l_i \tag{9}$$

that is, the expected net benefit from working must exceed the idiosyncratic disutility of work. For notational ease, we henceforth denote the difference in consumption utility $\frac{1}{1-\gamma}[(w_{ht} + w_{ft})^{1-\gamma} - w_{ht}^{1-\gamma}]$ by $W(w_{ht}, w_{ft})$.

Note first that given $\{\beta_H, \beta_L\}$ and earnings (w_{ht}, w_{ft}), irrespective of their beliefs and thus of the signal they receive, women with very low l's ($l \leq \underline{l}(w_{ht}, w_{ft})$) will always work and women with very high l's ($l \geq \bar{l}(w_{ht}, w_{ft})$) will never work, where

$$\underline{l}(w_{ht}, w_{ft}) \equiv W(w_{ht}, w_{ft}) - \beta_H \tag{10}$$

$$\bar{l}(w_{ht}, w_{ft}) \equiv W(w_{ht}, w_{ft}) - \beta_L \tag{11}$$

Next, for each women of type l_j, $\underline{l} < l_j < \bar{l}$, one can solve for the critical value of the private signal $s_j^*(\lambda)$ such that, for any $s \leq s_j^*$, given her prior belief λ, she would be willing to work. Let $p = Pr(\beta^* = \beta_L)$ and let p_j^* be the critical probability such that a woman of type l_j is indifferent between working and not, i.e.,

$$p_j^* \beta_L + (1 - p_j^*)\beta_H = W(w_{ht}, w_{ft}) - l_j \tag{12}$$

Using (10), we obtain $p_j^*(w_{ht}, w_{ft}) = \frac{l_j - \underline{l}(w_{ht}, w_{ft})}{\beta_H - \beta_L}$ and hence,

$$\ln \frac{p_j^*}{1 - p_j^*} = \ln \frac{l_j - \underline{l}}{\bar{l} - l_j} \tag{13}$$

Thus, the critical value, s_j^*, of the private signal a woman of type l_j must receive in order to work, given a prior of λ_t, is given by

$$\lambda_t(s_j^*) = \lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma^2_\epsilon}\right)(s_j^* - \bar{\beta}) = \ln\left(\frac{l_j - \underline{l}}{\bar{l} - l_j}\right)$$

and hence

$$s_j^*(\lambda_t; w_{ht}, w_{ft}) = \bar{\beta} + \left(\frac{\sigma^2_\epsilon}{\beta_H - \beta_L}\right)\left(\lambda_t + \ln\left(\frac{\bar{l}(w_{ht}, w_{ft}) - l_j}{l_j - \underline{l}(w_{ht}, w_{ft})}\right)\right) \equiv s_j^*(\lambda_t) \tag{14}$$

We can conclude from the derivation above that the proportion of women of type l_j,

[20] To obtain (8) one uses the fact that $Pr(s|\beta)$ is equal to the probability of observing a signal s generated by a normal distribution $N(\beta, \sigma^2_\epsilon)$.

[21] The structure of the model will ensure that this is the case.

$\underline{l} < l_j < \bar{l}$, that will work in time t given a prior of λ_t and a true state of nature β^*, $\omega_{jt}(\beta;\lambda_t)$, is the proportion of this type that receives signals lower than $s_j^*(\lambda_t)$, i.e.,

$$\omega_{jt}(\beta^*;\lambda_t) = F(s_j^*(\lambda_t) - \beta^*;\sigma_\epsilon) \tag{15}$$

Thus, the total proportion of women that will work in period t is given by:

$$\omega_t(\beta^*;\lambda_t) = G(\underline{l}) + \int_{\underline{l}}^{\bar{l}} F(s_j^*(\lambda_t) - \beta^*;\sigma_\epsilon)g(l_j)dl \tag{16}$$

where $g(\cdot)$ is the pdf of the l distribution $G(\cdot)$. Note that, as in the prior model, $\frac{\partial\omega_t}{\partial w_f} > 0$ and $\frac{\partial\omega_t}{\partial w_h} < 0$.

3.1 Intergenerational Transmission

What information is passed on from generation t to generation $t+1$? I assume that each woman passes on to her child her prior, $\lambda_{it}(s)$. Equivalently, generation $t+1$ inherits the prior of generation t (its "culture"), λ_t, which each individual then updates with her private signal (which can be assumed to be either inherited from her mother or the result of a new random draw s). If solely this information was transmitted intergenerationally, then the learning model would behave in the same way as the earnings only model since this implies $\lambda_{it}(s) = \lambda_{it+1}(s)$; the only change in work behavior over time would result from changes in wages. There is, however, an additional source of information available to women in $t+1$ that was not available to women at time t – the proportion of women who worked in period t.

If generation $t+1$ were able to observe perfectly the aggregate proportion of women who worked in period t, ω_t, they would be able to back out the true state of nature, β^*, as a result of the law of large numbers (i.e., using equation (16)). While assuming that information about how many women worked in the past is totally unavailable seems extreme, the notion that this knowledge is completely informative seems equally implausible and is merely an artifact of the simplicity of the model. I employ therefore the conventional tactic in this literature and assume that women are able to observe a noisy function of the aggregate proportion of women worked.[22]

In particular, I assume that women observe a noisy signal of ω_t, y_t, where

$$y_t(\beta^*;\lambda_t) = \omega_t(\beta;\lambda_t) + \eta_t \tag{17}$$

and where $\eta_t \sim N(0,\sigma_\eta^2)$ with a pdf denoted by $h(\cdot;\sigma_\eta)$.[23] Thus, given a common inherited

[22] An alternative assumption, pursued in Fernández and Potamites (2007), is that agents know the work behavior of a small number of other women in their social circle (as in Banerjee and Fudenberg (2004)). This yields similar results. It has the advantage, for the calibration, of not requiring a specification of an aggregate shock but the disadvantage of being sensitive to assumptions about the size of a woman's social group. Amador and Weill (2006) also obtain an S shape in the behavior of aggregate investment by assuming that agents observe a noisy private signal of other's *actions* as well as a noisy public signal of aggregate behavior. They are interested in the welfare properties of the two sources of information.

[23] The assumption that η is distributed normally implies, as usual, that some observations of y_t will be

Figure 3: Timeline of Learning Model

t									**t + 1**
λ_t	$s_{it} = \beta^* + \epsilon_{it}$	$\rightarrow$	$\lambda_{it}(s)$	$\rightarrow$	ω_t	$\rightarrow$	$y_t = \omega_t + \eta_t$	$\rightarrow$	λ_{t+1}
Public Belief	Private Signal		Private Belief Update		Work Decision (Aggregate)		Observation		Public Updating of Belief
	Private Learning						***Social Learning***		

prior of λ_t, after observing last period's signal of aggregate female LFP, y_t, Bayes' law implies an updated common belief for generation $t+1$ of:

$$\begin{aligned}\lambda_{t+1}(\lambda_t, y_t) &= \lambda_t + \ln \frac{h(y_t|\beta^* = \beta_L)}{h(y_t|\beta^* = \beta_H)} \\ &= \lambda_t + \left(\frac{\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)}{\sigma_\eta^2}\right)\left(y_t - \frac{\omega_t(\beta_L;\lambda_t) + \omega_t(\beta_H;\lambda_t)}{2}\right) \quad (18)\end{aligned}$$

Note that (18) is the law of motion of aggregate beliefs (culture) for the economy.

One way to think about the above assumption is that it is a shorthand for agents knowing the proportion of women who worked but uncertain about the distribution of married men and women's incomes. Alternatively, one could assume that individuals perfectly observe LFP, but are uncertain about the distribution of an idiosyncratic utility factor affecting the disutility of work. The value of some parameter in the distribution of the idiosyncratic utility factor would change randomly every period (e.g. by depending on an unobservable aggregate factor in the economy).[24] This formulation would be mathematically equivalent, but would entail a considerable amount of additional notation.

Figure 3 summarizes the time line for the economy. Individuals start period t with a common (updated) prior, λ_t. Each woman updates the common prior with her (inherited or observed) private signal and makes her work decision, generating an aggregate ω_t and a noisy signal y_t. Generation $t+1$ observes y_t and uses it to update the old common prior (λ_t), generating λ_{t+1} – the "culture" of generation $t+1$.[25] The process continues as described in each period. It should be noted that instead of assuming women in $t+1$ inherit λ_t (or λ_{it}) which they update with the information contained in y_t, we can equivalently assume that women observe the entire history of y_τ, $\tau = 0, 1, 2..., t$. This would yield the same value of λ_{t+1} (or of λ_{it+1}).

negative (and some greater than one) and so should be taken as an approximation for analytical simplicity. Alternatively, one can assume that the distribution is a truncated normal and allow the truncation to change with the range, for example, but this just renders the analytical expressions and computations more cumbersome.

[24] See, for example, Chamley (1999).

[25] Thus, we can think of generation τ as having a shared culture given by λ_τ with the individual deviations around λ_τ (given by the normal distribution of $\lambda_{i\tau}(s)$) constituting the distribution of beliefs induced by different individual's dynastic histories (i.e., by their inheritance of different realizations of s).

3.2 Some Properties of the Learning Model

In additional to generating qualitatively similar comparative statics as in the model with no learning (i.e., $\frac{\partial \omega_t}{\partial w_{ft}} > 0$, $\frac{\partial \omega_t}{\partial w_{ht}} < 0$), the learning model has several important properties that will prove useful in generating LFP dynamics similar to those in figure 1.

Note first that beliefs in this model are unbounded. Hence, in the long run beliefs must converge to the truth.[26] Since female LFP has been increasing over time, this implies that it is likely that $\beta^* = \beta_L$ and we shall henceforth assume that this is the case.

A key characteristic of this model is that it naturally generates an S-shaped LFP curve. To see why, note that given $\beta^* = \beta_L$, we can rewrite (18) as

$$\lambda_{t+1} = \lambda_t + \left(\frac{\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)}{\sigma_\eta^2}\right)\left(\eta_t + \frac{\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)}{2}\right) \tag{19}$$

Hence, the change in the LLR is increasing in the difference between the aggregate proportion of women who work when $\beta^* = \beta_L$ relative to the proportion who work when $\beta^* = \beta_H$. A large change in the LLR will, ceteris paribus, imply a relatively large change in the proportion of proportion of women who change their work decisions; if beliefs hardly change, there will be few women who change their work decision over time (for given wages).

To understand when the aggregate work difference $\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)$ will be large or small, we can start by noting that for a given $l_j \in \left(\underline{l}, \bar{l}\right)$ type this difference is equal to:

$$F(s_j^*(\lambda_t) - \beta_L;\sigma_\epsilon) - F(s_j^*(\lambda_t) - \beta_H;\sigma_\epsilon) \tag{20}$$

Taking the derivative with respect to λ yields the f.o.c.

$$\left[f\left(s_j^* - \beta_L\right) - f\left(s_j^* - \beta_H\right)\right]\left(\frac{\sigma_\epsilon^2}{\beta_H - \beta_L}\right) = 0 \tag{21}$$

Recalling that $f\left(s_j^* - \beta\right) = \frac{1}{\sqrt{2\pi}\sigma_\epsilon}\exp - \left\{\left(\frac{(s_j^* - \beta)^2}{2\sigma_\epsilon^2}\right)\right\}$, (20) is minimized at $s_j^* = \pm\infty$ and it is at a maximum at $s_j^* = \overline{\beta}$.

Thus, if the critical signal $s_j^*(\lambda_t)$ is far from $\overline{\beta}$ in absolute value, (20) will be small. This implies that the difference in the value of the aggregate signal $y_t(\beta^*;\lambda)$ across the two states will be swamped by the variance of the aggregate noise term η_t. Thus, the amount of intergenerational updating will be small and hence the change in the proportion of women who work that period, ceteris paribus, will likewise be small.

This property of the normal distribution is illustrated in figure 4 which depicts the distribution of ϵ, $N\left(0, \sigma_\epsilon^2\right)$. As can be seen in the figure, when $s^* - \beta$ is far from zero, the difference in proportion of women who work in the two states is small, i.e., the difference between ω_j at $s^* - \beta_L$ and $s^* - \beta_H$, (i.e., the shaded area) is small, and thus not very informative, given the noise, about the true state of nature. The opposite is true at $s^{*\prime}$.

[26] See, e.g., Smith and Sorensen (2001). Chamley (2004) gives an excellent explanation of the conditions required for cascades to occur.

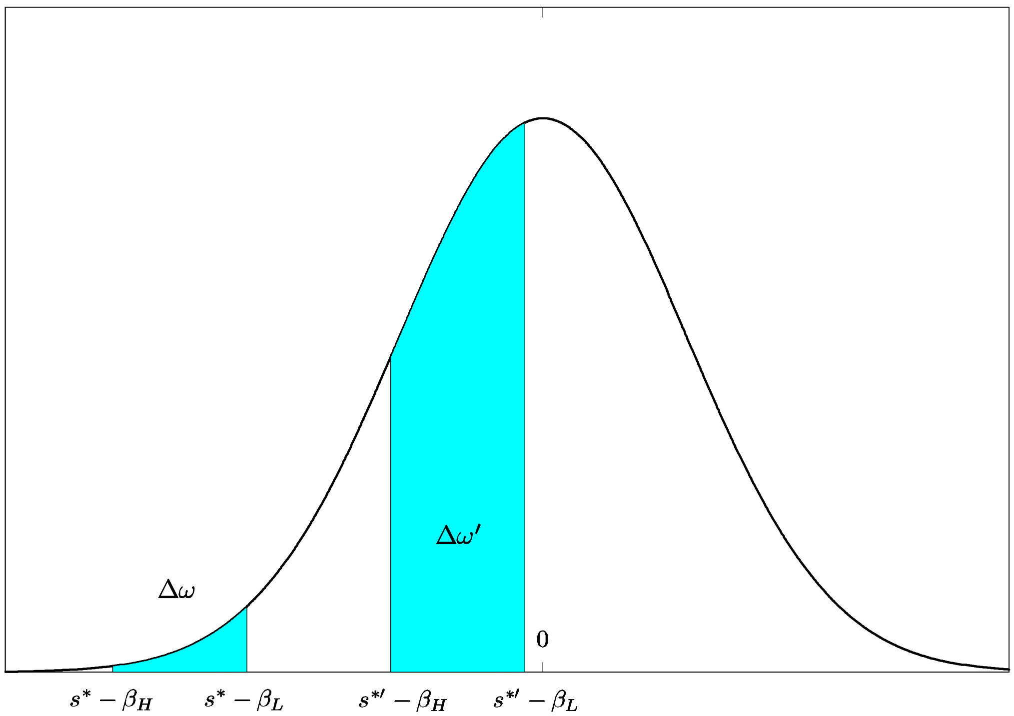

Figure 4: Normal PDF

Again, as shown in the figure for the same two values of β, when $s^{*\prime} - \beta$ is close to zero, the difference between ω_j' at the two states of nature is large.

Note that a similar conclusion holds once we aggregate over the l_j types. Taking the derivative of (16) we obtain

$$\frac{\partial \omega_t}{\partial \lambda_t} = \left(\frac{\sigma_\epsilon^2}{\beta_H - \beta_L}\right) \int_{\underline{l}}^{\bar{l}} \left[f\left(s_j^*(\lambda_t) - \beta_L\right) - f\left(s_j^*(\lambda_t) - \beta_H\right)\right] g(l_j) dl_j \tag{22}$$

Thus, if the critical signal $s_j^*(\lambda_t)$ is, for the average individual in $\left(\underline{l}, \bar{l}\right)$, far from $\overline{\beta}$, (22) will be small in absolute value, intergenerational updating will be small, and the evolution of LFP over time will be slow.[27] The opposite is true when the critical signal is close to $\overline{\beta}$ for the average individual.

It follows from the logic above that if parameter values are such that few women would choose to work if they assigned a low probability to $\beta^* = \beta_L$ (λ_t is low) whereas many women would choose to work if they assigned a high probability to this state (λ_t is high), then the amount of intergenerational learning that occurs when female LFP is either very low or very high will be relatively small as the average woman requires a very low realization of s to convince her to work in the first case, and a very high realization of s to convince her

[27] The assumption of heterogeneous types complicates matters since one must also be concerned about the size of $g(l)$. Thus, in order for the change in ω to be large, we need s_j^* to be close to $\overline{\beta}$ for types with a large frequency not only in $\left(\underline{l}, \bar{l}\right)$ but overall.

not to work in the second case. In both of these cases, the aggregate noise term dominates in (18) and hence the period to period change in female LFP will be likewise small. So, in these cases learning occurs, but it takes time. When, instead, the difference in the proportion of women who choose to work across states is large, i.e., when s_j^* is close to $\overline{\beta} \equiv \frac{\beta_H+\beta_L}{2}$ for l_j close to 0 (see footnote 27), then observing the aggregate signal tends to be informative, intergenerational learning is rapid, and the period to period change in female LFP will be large. Putting these statements together, it is easy to see that in this model the evolution of beliefs on their own (i.e., independently of earnings dynamics) will tend to generate an S-shaped curve, with a slow evolution of female LFP at the beginning, followed by rapid increases over time, and then tapering off again to small increases in female LFP until there is no more learning. At that point, any further changes in female LFP result solely from changes in earnings.[28]

3.3 Wages, Technology, and Learning

The learning model generates a novel role for changes in wages or for technological change that facilitates women's market work (e.g., the washing machine in Greenwood et al (2005) or the introduction of infant formula as in Albanesi and Olivetti (2007)). An increase in female wages, for example, will have the traditional static effect of increasing female LFP. In this model, however, it will an additional, dynamic effect; it will also affect the amount of intergenerational updating that takes place, i.e., $\lambda_{t+1} - \lambda_t$. This occurs not because it increases the proportion of women who work, but rather because it increases s_j^*.

If, for example, the average individual requires a very low value of the signal in order to work, the increase in s^* induced by an increase in women's wages will render y_t more informative for the next generation. As explained in the preceding section, an increase in s^* for the average individual increases the difference across states in the proportion of women who work (when λ is low) and hence increases the informativeness of the aggregate signal for the next generation. Thus, increases in female earnings or technological progress or policies that make it more attractive for women to work have a positive dynamic externality when the average woman requires a very low value of s in order to work, and have a negative dynamic externality under the opposite circumstances (i.e., when it would take a very large value of s for the average woman not to work). This gives a very different lens through which to evaluate the effects of changes in earnings, technology, and policy and one of the objectives of the next section will be to ask whether this effect is quantitatively important in explaining the historical evolution of female LFP.

[28] As should be clear from the intuition provided above, a normal distribution of the noise term ϵ is not critical. Rather the distribution needs to be able to give rise to a cdf that is increasing very slowly at the beginning, rapidly towards the middle, and then slowly once again towards the end.

4 Empirical Analysis

In this section I examine the ability of the simple learning model to replicate the dynamic path of female labor force participation over the last 120 years. We start with the model with no learning which I calibrate to match three key statistics of female LFP in the year 2000. This gives one a benchmark with which to measure how much the incorporation of endogenously evolving beliefs adds to the ability of the model to replicate the data. Next, I calibrate the learning model to match four additional statistics and show that the fully calibrated model does a good job of predicting the historical LFP series. The section concludes by examining the quantitative roles of beliefs relative to wages in the evolution of female LFP and distinguishing between the static and dynamic contribution of the changes in earnings to this process.

It should be noted from the outset that the empirical analysis is not a "test" of the model. In particular, the paper does not attempt to quantify the contributions of other potentially important factors discussed in the introduction to explain the data, except insofar as these are reflected in earnings changes (e.g., as would be the case for many forms of technological change or changes in wage discrimination). On the other hand, it should be clear that some of these alternative drivers of change, while considered exogenous and "belief free" in much of the literature, also reflect changed beliefs about the desirability of employing women and thus nesting these explanations is far from trivial.[29] To given an example, the pace of technological change in the household is likely to have been influenced by the perceived potential demand for these implements, which in turn is influenced by whether women are working outside the home.[30] The literature tends to ignore the effect of beliefs on the demand for household technological innovation. The contribution of this section is thus to evaluate the potential ability of a simple learning model to replicate the dynamics of female LFP and to examine the quantitative role of wages and beliefs in that process, abstracting from other, possibly complementary, channels.

4.1 Calibration Strategy

In both variants of the model, married women decide whether to engage in market work. Taken their husbands' earnings as given, they are faced with increasing their consumption with their own earnings if they choose to work or foregoing the consumption increase and not bearing the disutility of being a working woman. Thus, calibrating the models requires parameter values for the chosen analytical forms and an earnings or wage series for men and women. Since the model does not incorporate an intensive work margin, it is not clear how one should measure the opportunity cost of women's work. Given the paucity of data prior to 1940, I decided to use the (median) earnings of full time (white) men and women for which some data was available as of 1890. This choice exaggerates the earnings of working

[29] See Gayle and Golan (2006) for the estimation of a dynamic model in which firms (statistically) discriminate against women and beliefs evolve endogenously over time.

[30] See Adshade (2007) for a model in which the expansion of skill levels induces organizational and technological change which in turn increases female labor force participation.

women in general, as many work less than full time. As will be clear further on, however, the main conclusions are robust to reasonable alternatives.

For earnings data prior to 1940, I rely on numbers provided in Goldin (1990) who uses a variety of sources (Economic Report of the president (1986), Current Population Reports, P-60 series, and the U.S. Census among others) to calculate earnings for men and women.[31] As Goldin does not provide data for earnings in 1880 and 1910, these are constructed using a cubic approximation with the data from 1890 -1930 (inclusive).

As of 1940, I use the 1% IPUMS samples of the U.S. Census for yearly earnings (incwage) and calculate the median earnings of white 25-44 years old men and women who were working full time (35 or more hours a week) and year round (40 or more weeks a year) and were in non-farm occupations and not in group quarters.[32] As is commonly done, observations that report weekly earnings less than a cutoff a cutoff are excluded. The latter is calculated as half the nominal minimum wage times 35 hours a week and nominal weekly wages are calculated by dividing total wage and salary income last year by weeks worked last year.[33]

Figure 5 shows the evolution of female and male median earnings as calculated above over the 120 year period 1880-2000 (with earnings expressed in 1967 dollars). In order to compare procedures, the figure plots both the numbers obtained from the calculations above as of 1940 (they are shown in (red) dots) as well as Goldin's numbers (which continue to 1980 and are shown in (blue) x's). The only significant difference is with male earnings in 1950 which are higher for Goldin.[34]

To calibrate the models and to compare the predictions to the data requires female LFP numbers from 1980-2000. I use the numbers shown in figure 1 calculated from the US Census, which are for married white women (with spouse present), born in the US, between the ages of 25 and 44, who report being in the labor force (non-farm occupations and non-group quarters).

Both models are calibrated to match female LFP in the year 2000 as well as the own and cross wage elasticity of female LFP in that same year. For the learning model, I also match the cross-wage elasticity in 1990, female LFP in 1990, the relative probability of a woman working in 1980 (conditional on whether her mother worked), and female LFP in 1980. See table 1 for a list of the targets.

[31] See Goldin (1990) pages 64-65 and 129 for greater detail about the earnings construction for various years. I use the data for white men and women. I restrict the sample to white women as black women have had a different LFP trajectory with much higher participation rates earlier on.

[32] The sample is limited to full-time year-round workers because hourly wages are not reported. Even with this restriction, the usual issues remain (see Appendix). Furthermore, the sample could have been restricted to include only married men and women, but I chose not to do this in order to be consistent with the data from the earlier time period.

[33] See, for example, Katz and Autor (1999). This procedure is somewhat more problematic for the decades 1940-1960, when the federal minimum wage did not apply to all workers (prior to the 1961 amendment, it only affected those involved in interstate commerce). Nonetheless, I use the same cutoff rule as in Goldin and Margo (1992) as a way to eliminate unreasonably low wages. Note that by calculating median earnings, I do not have to concern myself with top-coding in the Census.

[34] Goldin's 1950 number is from the Current Population Reports, series P-60 number 41 (January 1962). It is for all men over 14 which may explain the discrepancy since our census figure leaves out men older than 44 who would, on average, have higher earnings.

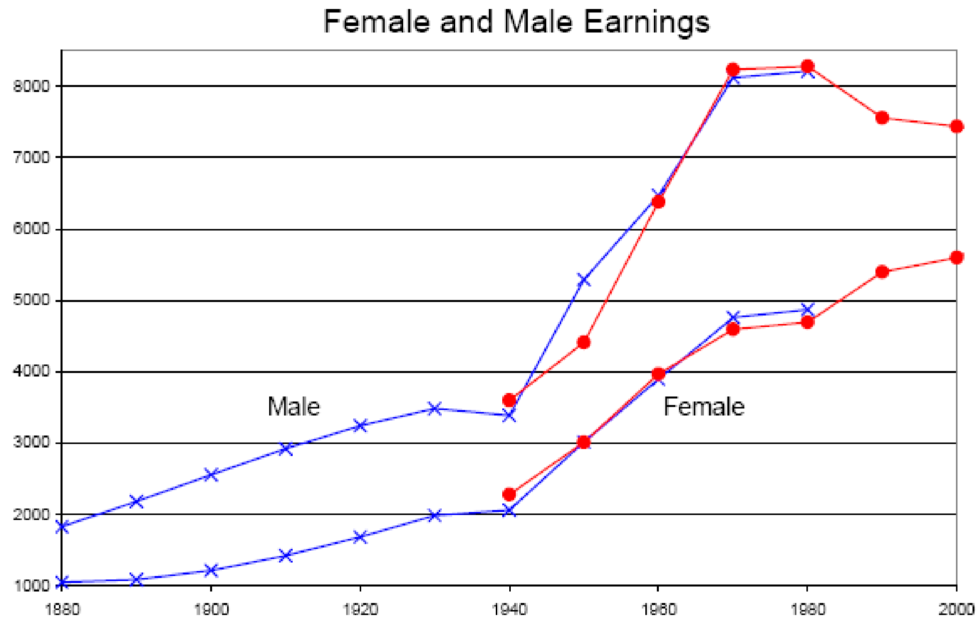

Figure 5: Crosses (blue) represent the yearly median earnings data from Goldin (1990), Table 5.1. Dots represent our calculations using U.S. Census data (red). They are the median earnings of white men and women between the ages of 25-44 in non-farm occupations and not living in group quarters. All earnings are expressed in 1967 \$. See text for more detail.

For the elasticity estimates I use those in Blau and Kahn (2006). The authors use the March CPS 1989-1991 and 1999-2001 to estimate married women's own-wage and husband's-wage elasticities along the extensive margin.[35] I use the results obtained from the basic probit specification, which does not control for education, as this way the elasticity measure obtained does not control for a measure of permanent income. This is preferable since I am more interested in an elasticity with respect to some measure of lifetime earnings. I also chose the specification without children as a control variable as it is endogenous. For the year 2000, Blau and Kahn estimate an own-wage elasticity of 0.30 and the cross-elasticity (husband's wage) of -0.13. The cross elasticity in 1990 is -0.14.[36]

To calculate the probability that a woman worked in 1980 conditional on her mother's work behavior, I use the General Social Survey (GSS) from 1977, 1978, 1980, 1982, and 1983.[37] The sample includes all white married women between the ages 25-45 who were

[35] They impute wages for non-working wives using a sample of women who worked less than 20 weeks per year, controlling for age, education, race and region, and a metropolitan area indicator (page 42). They run a probit on work (positive hours) including log hourly wages (own and husband's), non-wage income, along with the variables used to impute wages, both including and excluding education. The sample is restricted to married women 25-54 years old (with spouses in the same age range).

[36] Using the elasticities estimated from a specificaiton with education controls does not affect the results as the elasticities are very similar (0.28 and -0.12 for 2000 and -.15 in 1990).

[37] We used the ratio of the conditional probabilities rather than a conditional probability on its own since the latter is not consistent with the proportion of women who worked the previous generation. This is due to the fact that women in the GSS are more likely to report that their mother worked (given our lenient

born in the U.S.[38] The GSS asked a variety of questions regarding the work behavior of the respondent's mother. I used the response to the question "Did your mother ever work for pay for as long as a year, after she was married?" (MAWORK) to indicate whether a woman's mother worked. For each sample year, I calculated the ratio of the probability of a woman working (i.e., she reported being in the labor force) given that her mother worked relative to the probability of her working given that her mother didn't work (henceforth referred to as the work risk ratio). I averaged this ratio across the years in the sample to obtain an average risk ratio of 1.13, i.e., women whose mother worked are 13% more likely to work in 1980 than women whose mother didn't work. In the calibration each period is a decade and, for the purpose of computing the work risk ratio, daughters will make their work decisions two periods after their mothers (i.e., a separation of 20 years).

4.2 Calibrating the Model Without Learning

We start out by calibrating the model without learning (which will also be referred to as the "earnings only" model). In that model, only changes in earnings (male and female) can explain why labor supply changed over time. The unknown parameters are γ, β, and σ_l. As there is no direct evidence as these parameters, we calibrate them so that they are able to reproduce female LFP, a woman's own-wage elasticity, and her cross-wage (husband's wage) elasticity, all in the year 2000. These are useful statistics for the model to match as the ratio of the elasticities gives information about the curvature of the utility function and an elasticity and LFP value combined give information both about the magnitude of the common disutility of working, β, and about how dispersed the l types must be in order to generate a response to a change in wages. The simplicity of the model allows one to solve for the parameter values analytically.

Note that the wage elasticity ε (own, f, or cross, h) is given by:

$$\varepsilon_k = g\left(l^*\right) \frac{\partial l^*}{\partial w_k} \frac{w_k}{\omega} \tag{23}$$

$k = f, h$. Taking the ratio of the two elasticities and manipulating the expression yields a closed-form expression for γ, from which one can obtain a parameter value by using the earnings and elasticity numbers in 2000, i.e.,

$$\gamma = \frac{\log\left(1 - \frac{w_f}{w_h}\frac{\varepsilon_h}{\varepsilon_f}\right)}{\log\left(1 + \frac{w_f}{w_h}\right)} = 0.503 \tag{24}$$

Next one can use one of the elasticity expressions and the requirement that $G\left(l^*; \sigma_l\right) = \omega$ in 2000 to solve for β and σ_l. Note that since G is a normal distribution, one can write:

$$l^* = \sigma_l \Phi^{-1}\left(\omega\right)$$

work requirement) than what would be consistent with the Census numbers.

[38] Women who were students or retired were not included.

where Φ^{-1} is the inverse of a standard normal distribution $N(0,1)$. After some manipulation of (23), one obtains:

$$\sigma_l = \frac{A}{\exp\left(\frac{\Phi^{-1}(\omega)^2}{2}\right)} = 2.29 \qquad (25)$$

where $A = \frac{w_f (w_f + w_h)^{-\gamma}}{\sqrt{2\pi}\varepsilon_f \omega}$. One can then solve for β directly from the definition of l^*, yielding $\beta = 0.321$. To interpret this value, note that this is 4.7% of the consumption utility from working in 2000 or 22.4% of the difference in the consumption utility between working and not working in that year. In 1880, however this number represents 10.4% of the consumption utility from working or 88.1% of the difference in the consumption utility between working and not working.

As can be seen in figure 6, the calibrated earnings only model does a terrible job of matching the female LFP data (the data is shown in small circles and the (blue) line is the model's predicted LFP). It grossly overestimates the amount of female LFP that should exist in all decades other than 1990 and 2000.[39]

This basic inability of the earnings only model to match the historical data is robust to a wide range of values for the elasticities (I explored with values ranging from twice to half of those in Blau and Kahn). It is also robust to alternative specifications of the share of consumption that a woman obtains from her husband's earnings. In particular, one can modify the model so that the wife obtains only a share $0 < \alpha \leq 1$ of her husband's earnings as joint consumption. Figure 7 shows the results obtained from recalibrating the model using values of α that vary from 0.1 to 1. As is clear from the figure, this does little to remedy the basic problem. Furthermore, introducing any sensible time variation in this share would also not help matters as it would require women to have obtained a much larger share of husband's earnings in the past than in the present in order to explain why they worked so much less then. Since women's earnings relative to men's are higher now than in the past, most reasonable bargaining models would predict the opposite, i.e., a greater ability to obtain a higher share of male earnings now than in the past.[40]

The results are also robust to the exact choice of earnings series. One might argue that, over time, the average hours worked by women has changed and this intensive margin is not incorporated into the model. In order to more fully account for this margin, rather than use the median earnings of full-time women I constructed a series of the median annual earnings for all working women from 1940 to the year 2000. The sample consisted of 25-44 year old women who were born in the U.S., not living in group quarters, and working in a non-farm occupation. The adjustment to earnings was sizeable, ranging from 18% to 30% lower depending on the decade. This resulted in different parameter values ($\gamma = 0.49$, $\beta = .25$, $\sigma_l = 2.01$) but the predicted path of LFP generated was similar to the one obtained with the original series and hence still did an abysmal job of predicting the historical LFP

[39] It must, by construction, perfectly predict the 2000 level of LFP.

[40] Note that, in any case, to obtain the very low LFP numbers in 1880 would require women to fully share husband's earnings in that decade and to obtain a share of only 0.0001 of husband's earnings in the year 2000.

path.

4.3 Calibrating the Learning Model

We now turn to calibrating the learning model. As LFP has been increasing throughout and, from the results of the previous section we know that changes in wages alone are unlikely to explain this phenomenon, I assume that the true state of nature is given by $\beta^* = \beta_L$. In this case, learning over time about the true cost of working would, ceteris paribus, increase female LFP.

There is an additional complication in calibrating this model that was not present in the earnings only model – the presence of an aggregate observation shock in each period (i.e., individuals observe a noisy *public* signal of aggregate female LFP). This implies that the path taken by the economy depends on the realization of this shock. Each realization η_t generates a corresponding different public belief λ_{t+1} in the following period, and consequently a different proportion of women who choose to work after receiving their private signals. Note that we cannot simply evaluate the model at the mean of the expected η shocks (i.e., at zero) since, although λ_{t+1} is linear in η, the work outcomes ω_{t+1} are not.

I deal with the aggregate shock in the following way. For each period $t+1$, given LFP in the previous period ω_t, I calculate the proportion of women who would work, ω_{t+1}, for each possible realization of the shock, η_t, i.e., for each induced belief $\lambda_{t+1}(\eta_t)$. Integrating over the shocks, I find the expected value of LFP for that period, $E_t\omega_{t+1}(\lambda_{t+1}(\eta_t))$, and then back out the particular public belief (or shock) that would lead to exactly that same proportion of women working, i.e., I solve for $\lambda^*_{t+1}(\eta^*_{t-1})$ such that:[41]

$$E_t\omega_{t+1}(\lambda_{t+1}(\eta_t)) = \omega_{t+1}\left(\lambda^*_{t+1}(\eta^*_t)\right) \tag{26}$$

Performing this exercise in each period determines the path of beliefs.[42]

Continuing with the calibration, after some algebra and noting that $\frac{\partial \bar{l}}{\partial w_k} = \frac{\partial l}{\partial w_k}$, $k = f, h$, one can show that the ratio of the elasticities in this model can be written as

$$\frac{\varepsilon_{w_f}}{\varepsilon_{w_h}} = \frac{\frac{\partial l}{\partial w_f}}{\frac{\partial l}{\partial w_h}} \frac{w_f}{w_h}$$

Noting further that $\frac{\partial l}{\partial w_k} = \frac{\partial l^*}{\partial w_k}$, this implies that by performing the same manipulations as in the previous section one obtains (24), and thus the same value of γ as in the earnings

[41] For the computation, I take a large number of draws of entire histories for η (500 histories) in order to calculate the expected value of ω. See the Appendix for details.

[42] An alternative derivation can be obtained by modeling the economy as populated by a large number (or continuum) of communities k, each of which observes $y_{t,k} = \omega_t + \eta_{t,k}$ where η is an iid draw from the normal distribution $N\left(0, \sigma^2_\eta\right)$. Given a common prior, λ_t (and the same distribution of individual signals as before), the proportion of individuals that work in period $t+1$ is obtained by integrating over the $\eta_{t,k}$. Thus, as before the aggregate labor force is given by equation (26), i.e., $\int_{\eta_k} \omega_{t+1}\left(\lambda_{t+1,k}\left(\eta_{t,k}\right)\right) = \omega_{t+1}\left(\lambda^*_{t+1}(\eta^*_t)\right)$. To maintain the common prior assumption,one would need to assume that in each period communities inherit the common "average" prior of the previous generation consistent with the aggregate work decision, i.e., generation $t+1$ would inherit the average cultural belief $\lambda^*_{t+1}(\eta^*_t)$.

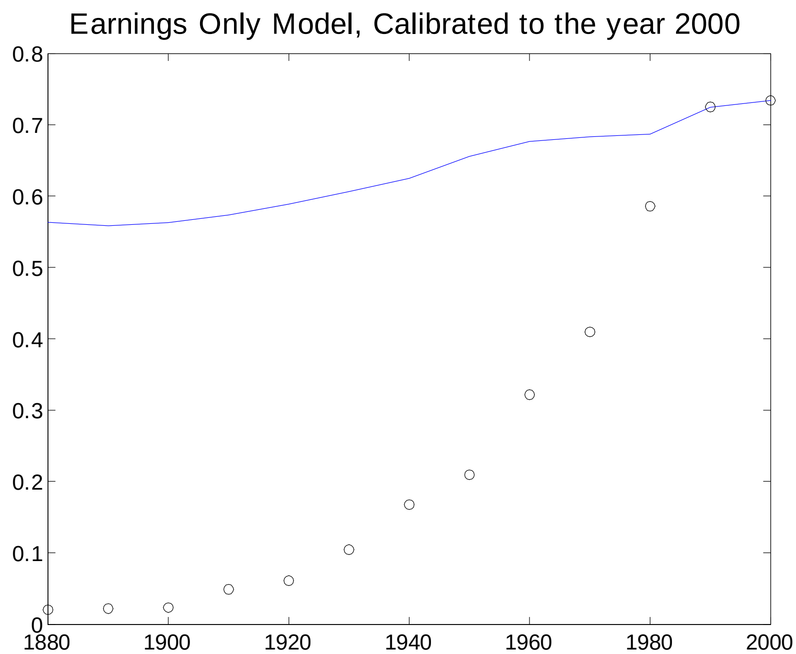

Figure 6: Parameters: $\gamma = 0.503$, $\beta = 0.321$, and $\sigma_L = 2.293$

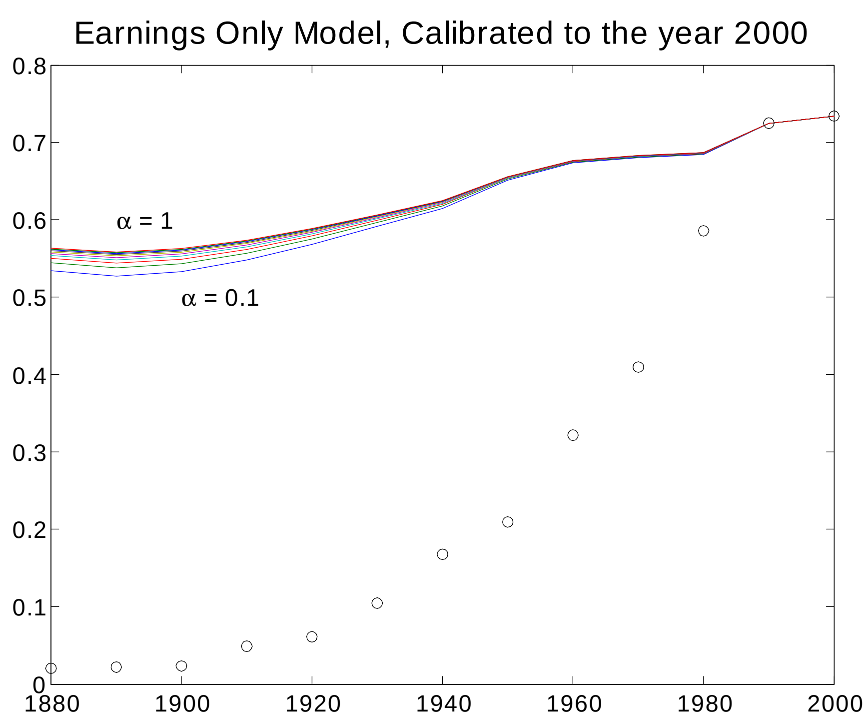

Figure 7: α is the fraction of husband's earnings that enters a wife's utility via consumption.

only model, i.e., $\gamma = 0.503$.

Before turning to the remaining calibration targets, it may be useful to first examine the maximum potential of this model by calibrating it solely to the same set of statistics from 2000 as the earnings only model. As the earnings only model is in this way nested within the learning model, it is not possible for the latter to do a worse job. How much better it can do, however, is not clear ex ante. As shown below, it greatly improves the ability of the model to match the data.

The results of this partial calibration exercise are shown in figure 8; table 1 reports the parameter values under the column "partially calibrated". The (blue) solid line in figure 8 shows the evolution of the expected value of female LFP and the (red) dashed line(labeled P) shows the evolution of public beliefs, i.e., the belief, $p_t\left(\eta_{t-1}\right)$, that the true state is β_L in period t (derived from λ_t). As can be seen from figure 8, what we henceforth denote the "partially calibrated model" does an excellent job of replicating the LFP time series.

Table 1

Calibration Targets		Earnings Model	Partially Calibrated	Learning Model
Own-Wage Elasticity (2000)	0.30	**0.30**	**0.30**	**0.29**
Cross-Wage Elasticity (2000)	-0.13	**-0.13**	**-0.13**	**-0.13**
Female LFP (2000)	0.734	**0.734**	**0.736**	**0.744**
Female LFP (1990)	0.725	0.725	0.696	**0.716**
Cross-Wage Elasticity (1990)	-0.14	-0.13	-0.14	**-0.14**
Female LFP (1980)	0.586	0.687	0.601	**0.585**
Work Risk Ratio (1980)	1.13	1	1.27	**1.13**
Parameters				
γ		0.503	0.503	0.503
σ_L		2.293	2.067	2.085
β		0.321		
β_H			7.481	4.935
β_L			.0004	.001
$P_0(\beta = \beta_L)$			0.110	0.057
σ_ϵ			5.408	5.288
σ_η			0.157	0.055

All elasticities are from Blau & Kahn (2006). The work risk ratio uses data from GSS (see text). The values in bold (first panel) are the model's predicted values for its calibration targets.

We now return to the full calibration exercise in order to impose more discipline on the free parameters of the model. Parameter values are chosen such that, in addition to being

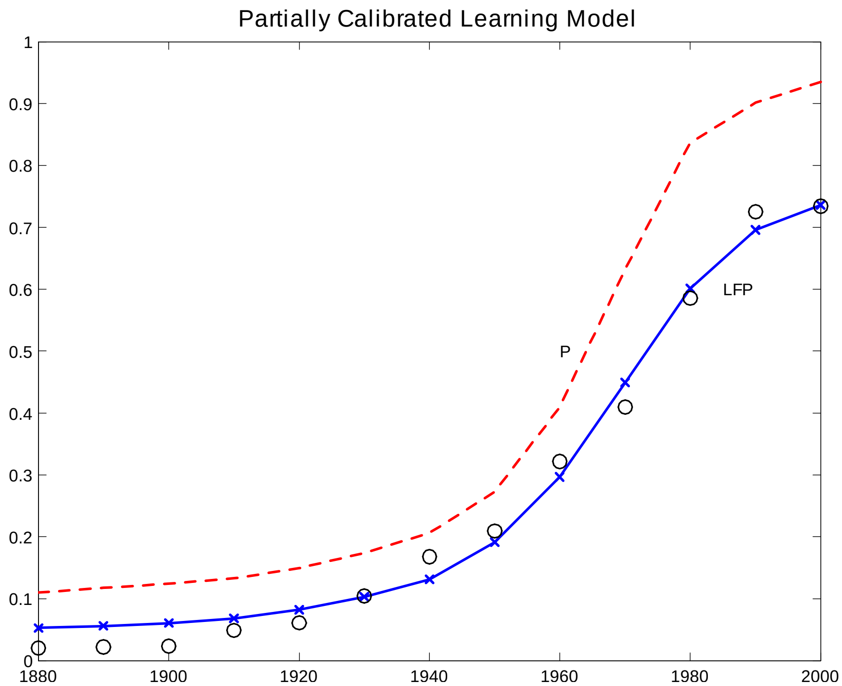

Figure 8: x indicates the predicted LFP path (blue). The dashed (red) line (p) is the belief path. Sum of squared errors (distance of predicted LFP from actual LFP) is 0.009.

able to generate the statistics discussed above, the model is able to match the cross-wage elasticity in 1990, female LFP in 1990, the work risk ratio in 1980, and female LFP in 1980. The values of these statistics are shown in table 1. As in the earnings only model, the additional elasticities and values of female LFP give one information both about how how bad women believe, on average it is to work and how dispersed women should be in their willingness to work at those wages. Unlike before, however, this dispersion is given not only by that of the distribution of the l types, σ_l, but also by the dispersion of private information, σ_ϵ. Furthermore, as the expected value of β is evolving over time with the beliefs λ, the values of LFP from 1980-2000 yields information as well on how rapidly λ needs to evolve and hence on how noisy the signal η should be (i.e., on σ_η).

In order to calculate a daughter's conditional probability of working (as a function of her mother's work behavior), we need to specify an inherited characteristic; otherwise, the conditional probability of working is the same as the non-conditional probability, which is not true in the data. In the learning model, either the private information (the signal) or the l_j type could be inherited. We assume that the signal is perfectly inherited whereas the l_j type is a random draw from the normal distribution $G(\cdot)$ that is *iid* across generations.[43]

Thus, given a signal s we can define l_s as the l_j type that is just indifferent between working and not at that signal value (i.e., $s^*_{l_s} = s$). Hence, the probability that a woman

[43] Thus, this model yields a postive correlation between a mother and her daughter's work "attitudes" ($E_{it}\beta + l_i$ and $E_{i',t+1}\beta + l_{i'}$ where i indexes the mother and i' the daughter). See Farré-Olalla and Vella (2007) for recent evidence on the correlation of mother's and daughter's attitudes towards work.

with signal s works is $G(l_s)$, i.e., it is the probability that her l type is smaller than l_s. Rearranging the expression for s_j^* in (14), we obtain

$$l_{st} = \frac{\underline{l}_t + \bar{l}_t \exp\left(\lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma_\epsilon^2}\right)(s - \bar{\beta})\right)}{1 + \exp\left(\lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma_\epsilon^2}\right)(s - \bar{\beta})\right)} \tag{27}$$

And, using Bayes rule and $\beta^* = \beta_L$, we can calculate the probability that a daughter works given that her mother worked as:

$$\begin{aligned}
\Pr(DW_t|MW_{t-2}) &= \frac{\Pr(DW_t \text{ and } MW_{t-2})}{P(MW_{t-2})} \\
&= \frac{\int_{-\infty}^{\infty} \Pr(DW_t \text{ and } MW_{t-2}|s) f(s - \beta_L) ds}{\omega_{t-2}(\beta_L)} \\
&= \frac{\int_{-\infty}^{\infty} G(l_{st}) G(l_{s,t-2}) f(s - \beta_L) ds}{\omega_{t-2}(\beta_L)}
\end{aligned} \tag{28}$$

where DW and MW stand for daughter works and mother worked, respectively. We use the predicted LFP from two periods earlier to calculate the probability that mothers worked (hence the $t-2$ in expressions such as $G(l_{s,t-2})$). Note that in (28), the probability that both mother and daughter worked, $\Pr(DW_t$ and $MW_{t-2}|s)$, is multiplied by $f(s - \beta_L)$ as this is the proportion of daughters (or mothers) who have a private signal s in any time period.

A similar calculation to the one above yields

$$\Pr(DW_t|MNW_{t-2}) = \frac{\int_{-\infty}^{\infty} G(l_{st})(1 - G(l_{s,t-2})) f(s - \beta_L) ds}{1 - \omega_{t-2}(\beta_L)} \tag{29}$$

where MNW denotes a mother who did not work. The work risk ratio is thus given by

$$R_t = \frac{\Pr(DW_t|MW_{t-2})}{\Pr(DW_t|MNW_{t-2})} \tag{30}$$

The results of the fully calibrated model are shown in figure 9; table 1 reports the parameter values and calibration targets. As in figure 8, the (blue) solid line shows the evolution of the expected value of female LFP and the (red) dashed line shows the evolution of the probability that the true state is β_L. See table 1 for a comparison of the calibration targets and the model's predicted values.

The calibrated model does a good job of replicating the historical path of female LFP.[44] It under-predicts LFP from 1940 to 1970, however, and slightly over predicts it from 1880 to 1900. Individuals start out in 1880 with pessimistic beliefs about how costly it is to work. They assign around a 6% probability to the event $\beta^* = \beta_L$ (i.e., λ_0 implies $p_0 = 0.06$). Beliefs evolves very slowly over the first seventy years or so (remaining no higher than 10%

[44] The sum of squared errors (between actual and model predicted LFP) is 0.052.

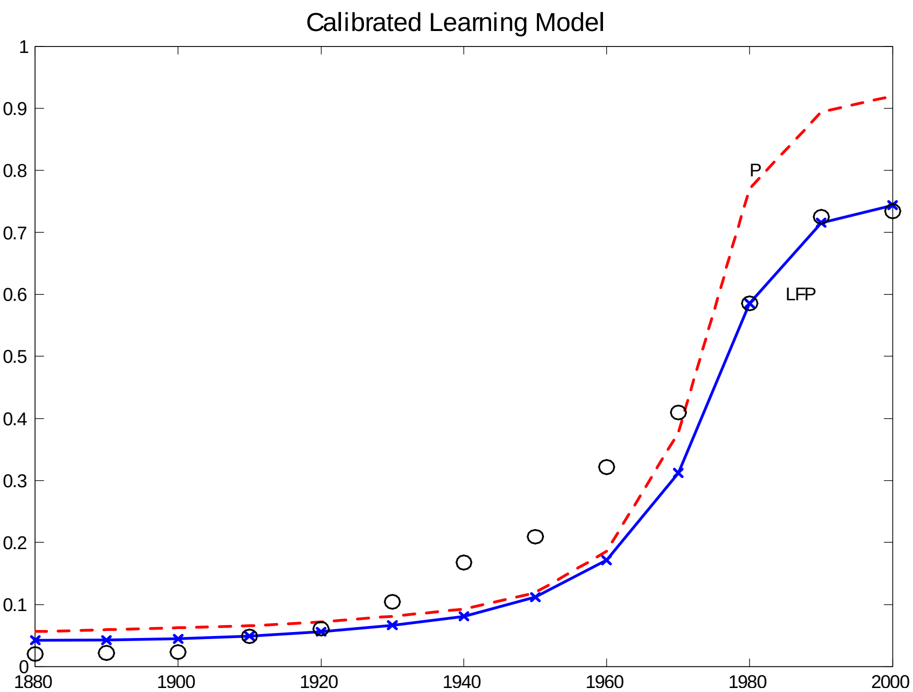

Figure 9: The dashed red line (p) is the belief path. The sum of squared errors (distance of predicted LFP from actual LFP) is 0.052.

for this period). Then, as of 1960, the change in beliefs accelerate, jumping from assigning a probability of 18.6% to β_L in 1960, to 37.7% in 1970, to 77.0% in 1980. By 2000, the public belief assigns a probability of 92.0% to $\beta^* = \beta_L$.

Individual beliefs are very dispersed as the private signal is very noisy. Figure 10 shows the path of beliefs once again, but this time for the individual with the median or mean LLR, $\lambda_{it}(s)$, as well as for the individuals with signals two standard deviations below and above this mean.[45]

The fact that the model's predictions are too low in the period 1940-1970 may indicate that another factor, such as technological change in the household, was also responsible for the higher levels of LFP during this period. Note that a characteristic of the learning model is that any technological change that occurred in the 1930s and 1940s (e.g., the clothes washer and other housework savings devices discussed in Greenwood et al (2005)) would have had repercussions in later decades through the dynamic impact of technological change on learning discussed earlier. Alternatively, it may be that world war II made women more willing to work and that this in turn increased the pace of intergenerational learning.

It is of interest to examine the pattern of own and cross wage elasticities predicted by the model. As has been long noted, a model that generates a constant wage elasticity is at odds with the data (see, e.g., Goldin (1990). Both the earnings only model and the learning

[45] Using (8), note that the median individual has a LLR given by $\lambda_t + \frac{(\beta_L - \beta_H)^2}{2\sigma_\epsilon^2}$.

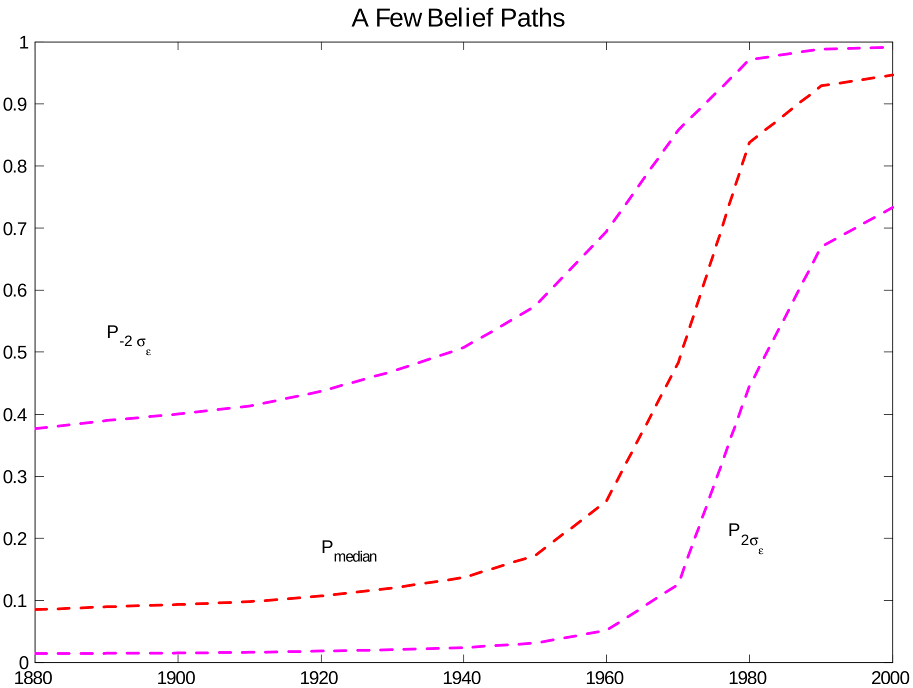

Figure 10: This shows $\Pr(\beta^* = \beta_L)$ for agents with $s = \beta^*$ and $s = \beta^* \pm 2\sigma_\varepsilon$.

model predict changing elasticities over time. In the earnings model this is due only to the heterogeneity of individuals with respect to their preferences (i.e., their l_j) whereas in the learning model, there is also heterogeneity in beliefs.

The learning model's elasticities predictions are shown in figure 11. Recall that the model is calibrated to match both elasticities in 2000 and the cross elasticity in 1990. As can be seen from the picture, over time both elasticities are first increasing (in absolute value) and then decreasing.[46] This pattern is similar to the historical one reported in Goldin (1990) with respect to women's own wage elasticity. One can speculate that it reflects, in the early decades, the unwillingness of women to work unless required to by a husband's low income. Over time, however, women become less pessimistic about the disutility of working and thus exhibit more sensitivity to their own (and husband's) wages until, further on in the process, by the 1960s, there is a much more widespread belief that it is not bad for a woman to work (recall that we find that indeed β_L is very close to zero) and there is a large drop with respect to the sensitivity to both her own and her husband's wages.[47]

A comparison of the earnings only model with the learning model is instructive. Why do they obtain such different LFP paths? As noted previously, the calibration implies that both models must have the same value of γ. Furthermore, the difference in the standard

[46] Note that Blau and Kahn (2006) also estimate decreasing absolute values for these elasticities over 1980-2000 period.

[47] See table 5.2 and the discussion in chapter 5 in Goldin (1990) . The correspondence between the model predictions and the data for the pattern of cross-wage elasticities is less clear as the studies reported in the table start in 1900 and show only a trend of becoming smaller in absolute value.

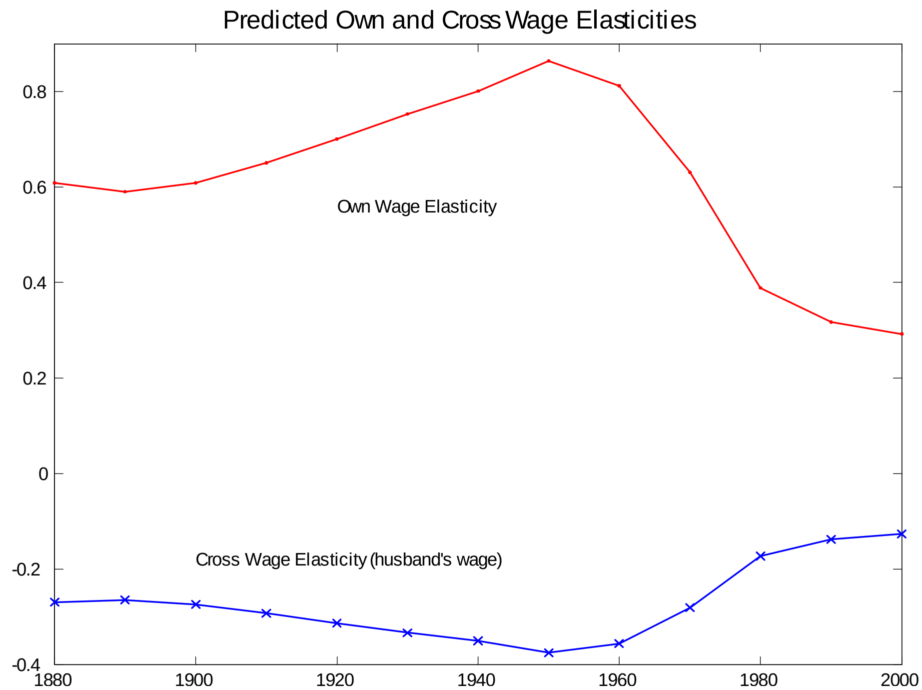

Figure 11: Parameter values from calibrated model. See the Appendix for a description of how the elasticities were calculated.

deviation of the normal distribution of types is relatively small: 2.29 versus 2.09. Lastly, the expected value of β in 2000 (a constant, of course, in the earnings only model) is also not very different across models: 0.32 as opposed to 0.40 in the learning model.[48] Thus, it is the endogenous evolution of the expected value of β in the learning model that is responsible for the difference in LFP behavior observed over time across the two models. Whereas by construction this remains constant in the earnings only model, in the learning model the expected value of β is close to 4.65 in 1880 and then evolves over time to 0.40 in 2000. This allows LFP to respond in dramatically different ways over time.

Another interesting difference between the two models is with respect to their predicted elasticities' paths. As just discussed, the learning model predicts large changes in elasticities over time. This is mostly generated by changes in beliefs. The earnings only model, on the other hand, while it does predict non-constant elasticity paths, generates much smaller changes over time as can be seen in figure 12.

It may be also be instructive to examine where the calibrated model does worse than the partially calibrated one. As can be seen from figures 8 and 9, the main decades in which the partially calibrated model does significantly better are 1950-1970. The requirement that in the fully calibrated model the parameters be able to match the work risk ratio appears to be mostly responsible for this. In the partially calibrated model, this ratio is quite a bit

[48] Note that the calibration does not require both models to have the same values of σ_l and β (for 2000) since the learning model has an additional source of heterogeneity (intra-generational heterogeneity in beliefs induced by private signals) which affects the elasticity.

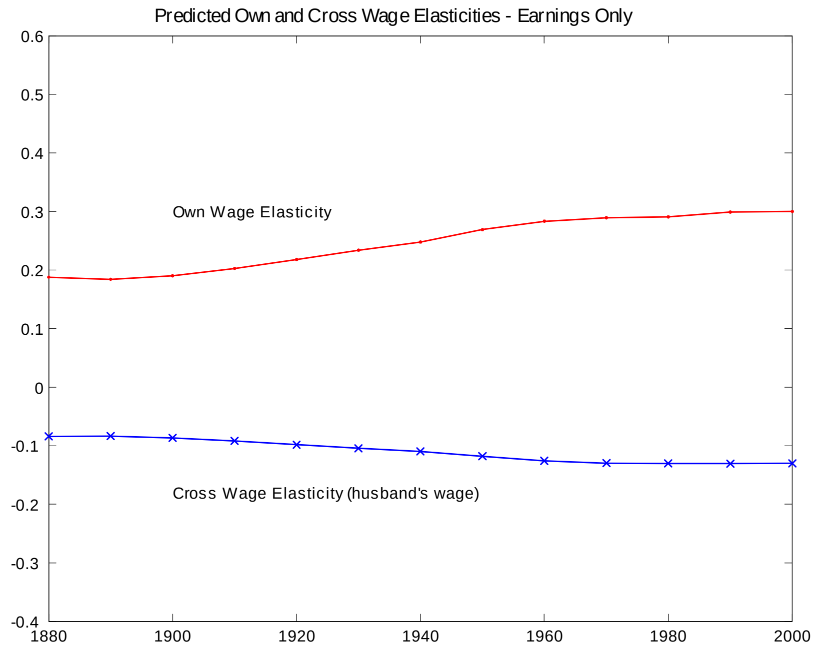

Figure 12: Uses the parameter values of the calibrated earnings only model.

higher than the target for the calibrated model (1.26 rather than 1.13).[49]

As a last exercise, we can use the calibrated learning model to generate a prediction for future female LFP and the elasticities. Using median earnings for men and women in 2005 as our guess for 2010 earnings ($7518 and $5959, respectively, in 1967 dollars and calculated as described earlier), our model predicts that 76.8% of women would work in 2010 with an own-wage elasticity of 0.29 and a cross-wage elasticity of -0.12.

From the discussion in this section, one can conclude that overall the simple learning model does a good job in predicting the historical path of LFP. We next turn to a quantitative assessment of the role of beliefs as well as the traditional static and non-traditional dynamic roles of changes in wages in generating the model's predicted LFP path.

4.4 The Roles of Wages and Beliefs

To investigate the roles of changes in earnings and in beliefs, we can start by not allowing public beliefs to evolve (i.e., the public signal is shut down). First, we can freeze beliefs at the 1880 level (i.e., at a prior of approximately 6% that $\beta^* = \beta_L$) and ask how labor force participation would have evolved in the absence of any updating of beliefs using the public signal. Thus, women have private information but there is no intergenerational evolution of beliefs. As show by the bottom line (with the caption "LFP if no public updating") in figure 13, female LFP would barely exceeded 10% by the year 2000.

[49] See table 2 for a comparison of the predictions of the calibration targets for the three models (earnings only, partially calibrated learning, and fully calibrated learning).

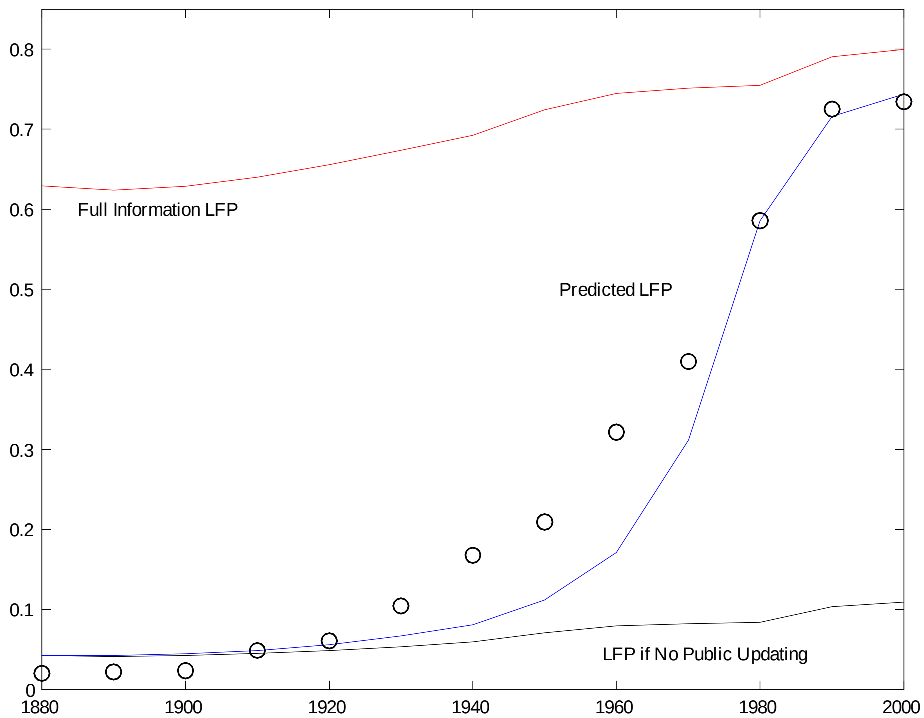

Figure 13: Uses the solution parameters from calibrated model but without public learning.

Alternatively, one can ask what female LFP would have been if, throughout the entire time period, agents had known the true value of β, i.e., $\beta^* = \beta_L$. This scenario is shown for the parameters of the calibrated model by the top (red) line (with the caption "full information LFP"). It predicts a very different trajectory than before, with LFP starting close to 63% in 1880 and slowly evolving to 80% by 2000. Thus, as can be seen from contemplating either of the two extremes regarding constant public beliefs, the actual dynamics of beliefs induced by learning is essential to producing the predicted path of female LFP also reproduced in figure 13. The model with dynamics induced solely by changes in male and female earnings along with unchanged beliefs grossly under or over estimates female labor supply over the entire time period.[50]

Next, we can distinguish between the static and dynamic effects of changes in earnings on female LFP by performing the following instructive decomposition. First, as before, we can keep earnings constant at their initial 1880 levels and let beliefs change endogenously. The LFP path obtained in this fashion, denoted $LFP(p_{1880}, w_{1880})$ in figure 14, results only from the changes in beliefs that would have occurred had earnings stayed constant at their 1880 levels. It is thus a measure of the quantitative importance of the evolution of beliefs for female LFP dynamics in which changes in earnings play no part. This LFP path is given by the bottom (magenta) line in figure 14. Hence, the difference between the level of LFP in 1880 (given by the dotted horizontal line) and $LFP(p_{1880}, w_{1880})$ measures the

[50] This is simply a repetition, with slightly different parameter values, of the finding that earnings only model does a very bad job of replicating the LFP trajectory.

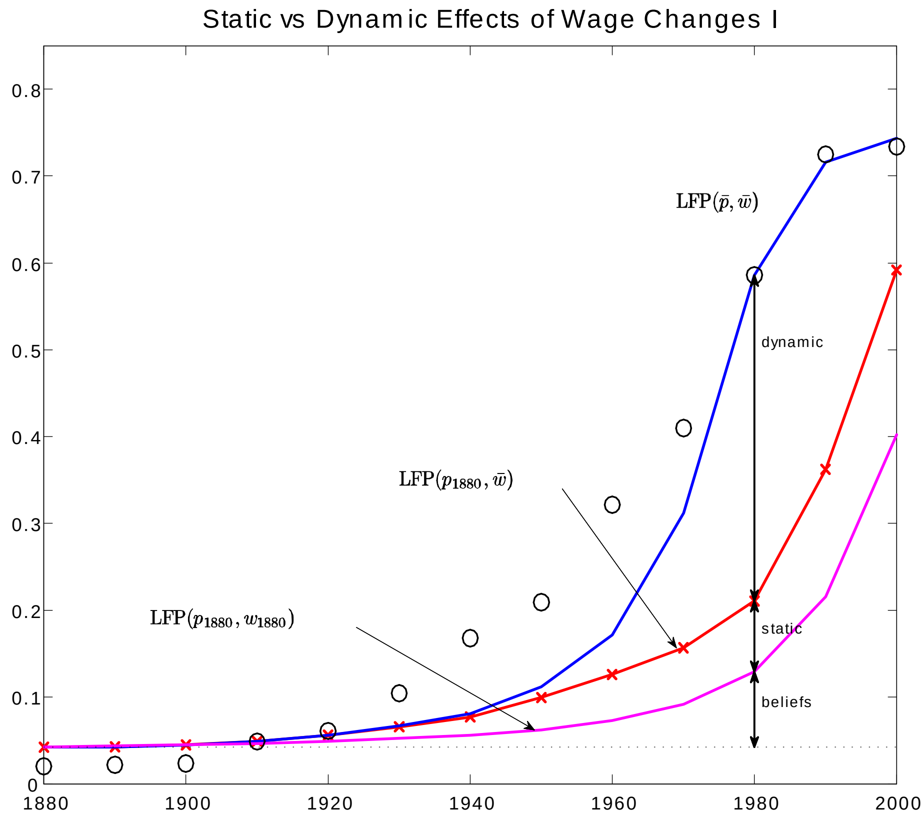

Figure 14: Decomposition of LFP. See the text for notation.

contribution of beliefs to the historical evolution of female LFP.

Combining the belief path obtained from the exercise above, p_{1880}, with the actual historical earnings path, $\overline{w}$, allows one to disentangle the dynamic from the static effect of wages. In this exercise, changes in earnings have the traditional direct effect of changing the attractiveness of working vs not working, but they do not have the dynamic effect on intergenerational beliefs since, by construction, these beliefs were derived from a constant (1880) wage path. We denote the LFP obtained this way by $LFP\,(p_{1880}, \overline{w})$ and it is shown with (red) x's in the figure. The difference between $LFP(p_{1880}, w_{1880})$ and $LFP\,(p_{1880}, \overline{w})$ measures the static contribution of wages to the evolution of LFP (as beliefs change over time in the same way for both curves whereas earnings change only in $LFP\,(p_{1880}, \overline{w})$).

Lastly, we allow wages to also influence intergenerational learning and thus beliefs and denote the LFP path obtained this way $LFP\,(\overline{p}, \overline{w})$). Note that this LFP path is the one predicted by the model and depicted previously in figure 9. It is the top (blue) curve shown in figure 14. The difference between $LFP\,(\overline{p}, \overline{w})$ and $LFP\,(p_{1880}, \overline{w})$ measures the dynamic contribution of wages to changing LFP obtained by changing beliefs (i.e., both series have the same historical earnings series, $\overline{w}$, but $LFP\,(\overline{p}, \overline{w})$ allows beliefs to respond to these changes and thus affect LFP whereas $LFP\,(p_{1880}, \overline{w})$ keeps the belief path that would have occurred had wages remained at their 1880 level).

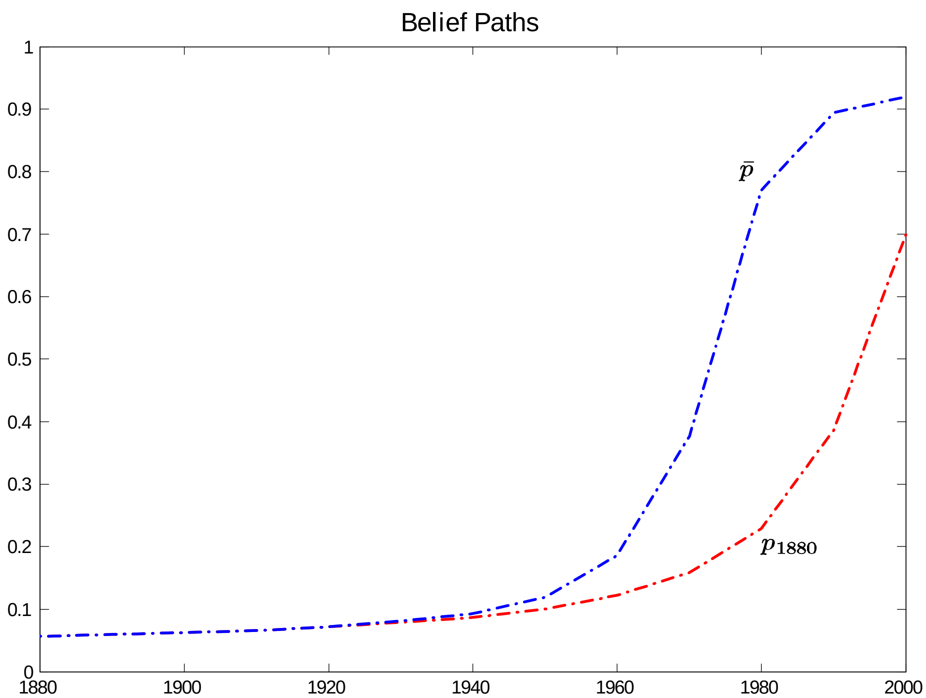

Figure 15: $P(\beta^* = \beta_L)$ for historical earnings series and for earnings constant at the 1880 levels.

As can be seen in figure 14, for the first several decades the static effect of wages is mostly responsible for the (small) increase in LFP. Over time, both the dynamic effect of wages on beliefs and the evolution of beliefs independently of wage changes become increasingly important, with the dynamic effect of wages accounting for over 50% of the change in LFP between 1970 to 1990, which are decades of large LFP increases.

To understand why the dynamic effect of wages is more important in some decades than others, it is useful to compare the two belief paths, $\overline{p}$ and p_{1880}, depicted in figure 15. Note that the difference in the probability assigned to $\beta^* = \beta_L$ is especially large in 1980 and 1990; these probabilities would have been 22.9 and 38.7 if earnings had not changed rather than 77.0% and 89.5% respectively. By 2000, however, the difference in probability assigned by the two belief paths diminishes considerably, which explains the decreased importance of the dynamic effect of earnings on beliefs.

The decomposition of LFP is not unique. One could alternatively eliminate the $LFP\,(p_{1880}, \overline{w})$ curve and replace it with the LFP path that would result if the beliefs followed the ones obtained from the historical earnings series, $\overline{p}$, but wages were kept constant at their 1880 levels. This curve is shown in figure 16 as $LFP\,(\overline{p}, w_{1880})$. The effect on LFP of beliefs with unchanged earnings ($LFP\,(p_{1880}, w_{1880})$) remains as before, but the dynamic effect of wages is now given by the difference between $LFP\,(\overline{p}, w_{1880})$ and $LFP\,(p_{1880}, w_{1880})$. These paths are obtained from the same constant 1980 earnings, w_{1880}, but in the first trajectory beliefs evolve as they would with the historical earnings profile, whereas in p_{1880} beliefs follow the path they would have taken had wages not changed over time. The static

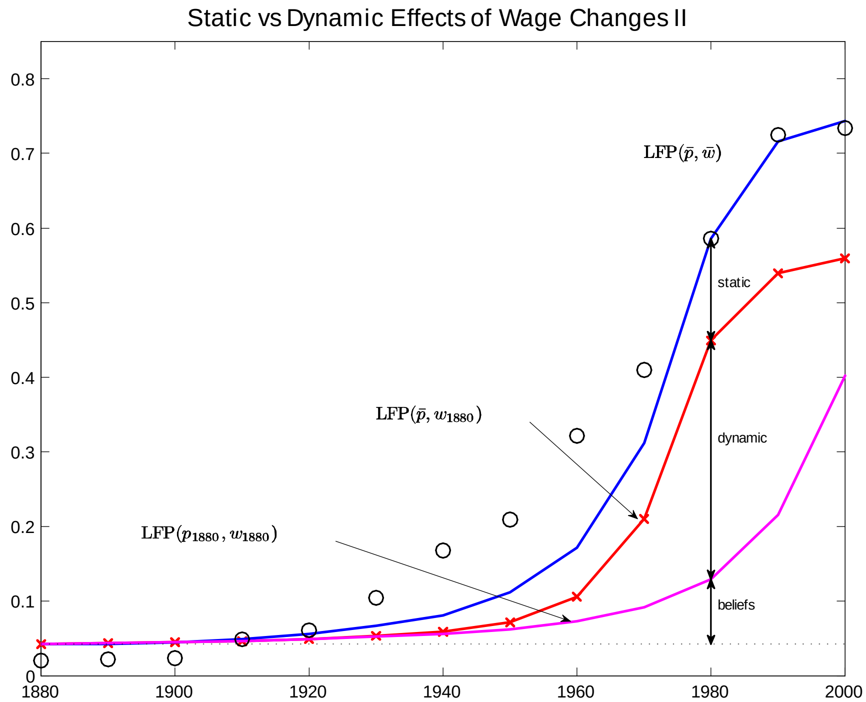

Figure 16: Alternative decomposition of LFP.

effect of earnings is now measured as the difference between $LFP\,(\overline{p}, w_{1880})$ and $LFP\,(\overline{p}, \overline{w})$, as beliefs evolve the same way for both series whereas earnings follow different paths.

With this alternative decomposition we obtain the same basic pattern as the one described above, with both the static and dynamic effect of wages becoming increasingly important over time, and with the dynamic effect accounting for between 40% to 60% of LFP in the decades 1970-1990. Thus, the way in which we decompose the wage effect into static and dynamic matters, but the basic conclusion remains the same as above.

We conclude from our decomposition of LFP that in some decades the dynamics of learning as induced by higher earnings was critical to the increases in female LFP. Overall, at different time periods, all three factors played important roles in the changes in female LFP.

5 Discussion and Conclusion

This paper models the dynamics of married women's labor force participation as reflecting a process of cultural change brought about by intergenerational learning. In this process, married women compare the benefits of increased consumption from labor earnings with the expected utility cost of working. This cost is unknown and women's beliefs about it evolve endogenously over time in a Bayesian fashion through the observation of noisy signals of the labor supply choices of women in the past and through the inheritance, through their mothers, of private information. I show that a simple model with these features, calibrated to key statistics from the later part of the 20th century, is capable of generating a time trend of female labor force participation that is similar to the historical one in the US over the last 120 years.

This model naturally generates the S-shaped curve of female LFP found in the data, shown in figure 1. This shape results from the dynamics of learning. When very few women participate in the labor market (as a result of initial priors that are very negative about the payoff from working), learning is very slow since the noisiness of the signal swamps the information content given by differences in the proportion of women who would work in different states of the world. As the proportion of women who work increases and beliefs about work become more positive, the information in the signal improves. Once a large enough proportion of women work though, once again, the informational content in the public signal falls since the difference in the proportion of women who would work under different states of the world is swamped by the variance in the noise.

To evaluate the ability of such a model to explain the quantitative evolution of female LFP, I first calibrate a version of the model without any evolution of beliefs to a few key statistics for the year 2000, namely married women's LFP, and the own and cross-wage elasticities of LFP. In this model, only changes in earnings over time can explain changes in female LFP. I show that such a model performs very badly and that it grossly overestimates the proportion of women who would have worked in virtually every decade since 1880. Introducing learning in this simple model and calibrating the model to additional statistics greatly improves its capacity to predict the historical path of female LFP.

The model also indicates a novel role for increases in women's wages (or for technological change), beyond the traditional direct effect of making it more attractive for women to work outside the home. In particular, when beliefs are relatively pessimistic, increases in women's wages make the private information (signal) required by the average woman in order to work less extreme, and thus render the public signal more informative. Thus, factors that make working more attractive when women are, on average, pessimistic, have an additional dynamic impact though the increased intergenerational updating of beliefs. Analysis of the calibrated model indicates that the dynamic effect of wages on beliefs played a quantitatively important role in changing female LFP, particularly over the period 1970-1990.

The model makes some heroic simplifying assumptions, including an unchanged true (psychic) cost of working over 120 years. It would not be difficult to incorporate changes in the cost structure, but without direct empirical evidence it seemed better to leave it constant and not introduce additional parameters. The model also ignored costs that are endogenous in nature. In particular, by modeling changes in culture arising solely as a process of learning about exogenous costs, it neglected the endogenous, socially imposed, costs stemming from social (cultural) reactions to married women in the work force. Questions of identity (as emphasized in the economics literature by Akerlof and Kranton (2000)), and society's reactions to and portrayals of working women, most likely also played an important role in determining the path of female LFP, as might have changes in vested economic interests. Other assumptions in the model, such as the normal distributions of the noise terms, could easily be replaced with others (e.g., single-peaked distributions and relatively thin tails on both sides of the modal frequency) that would preserve the same qualitative features, particularly the S-shaped curve.

The calibrated model finds that at the outset women were very pessimistic about the true cost of working. This lack of neutrality may indicate that particular social forces were at play in determining culture. Common economic interests for certain groups in industrial societies at that time (e.g., men?), may help explain why most countries shared the view that women working outside the home was harmful. Endogenizing this initial prior, however, is outside the model presented here and would require, in my opinion, a political economy framework to explain why certain opinions become dominant.[51] In future work, therefore, in addition to exploring the informational role of different social networks, it would also be of interest to incorporate the contribution that social rewards and punishments may play in changing behavior over time and to find a way to quantify their importance relative to learning.[52] Some interesting initial work in this area has been done by Munshi and Myaux (2006) who incorporate strategic interactions in the context of a learning model with multiple equilibria in which individuals are deciding whether to adopt modern contraception.[53]

In future research, it would be interesting to explore also the potential inefficiencies that arise because individuals do not take into account the effect of their actions on learning and to examine the role that policy could play. At the empirical level, it is important to depart from focussing exclusively on aggregate features of the data over a very long time horizon. In particular, sharper hypotheses about cultural change over a shorter time period would allow a greater use of microdata and permit one to learn more about the process of cultural diffusion.[54] Lastly, if one could reliably identify variation in policies or technologies across otherwise similar economic space, this could allow us to empirically quantify the dynamic effect of these on beliefs. Examining variation across states in the importance of WWII shocks may permit some progress in this direction.

[51] As the economy changed, so may have the interests of firms (capitalists) and perhaps men in general with respect to having women in the work force. For economic theories of changes in women's conditions (e.g. voting) see, for example, Doepke and Tertilt (2007) and Edlund and Pande (2002).

[52] The interaction of social networks and endogenous punishments is the topic explored in Fernández and Potamites (2007).

[53] In their model, the payoff in a period to an individual using birth control depends on her type (whether she is a "reformer" or not) and the contraceptive choice of the woman she interacts with in that period (this is a model with random matching). Thus, there is a strategic aspect to a woman's choice as her payoff depends upon the choice of the woman she meets. The authors show that if society starts in an equilibrium with no modern contraceptive use, whether it can transit to an equilibrium with contraceptive use will depend upon the proportion of individuals who are reformers, a constant fraction of which are assumed to use (for exogenous reasons) modern contraception every period. Reformers preferences are such that they obtain a higher payoff from using modern contraception.

[54] Munshi and Myaux test their hypothesis, for example, using microdata from a 10 year interval in Bangladeshi villages. Bandiera and Rasul (2006) and Conley and Udry (2003) use self-reported data on social contacts to construct networks to test their models of learning about new technologies. Mira (2005) structurally estimates his model using Malaysian panel data.

6 Appendix

6.1 Data

To construct the earnings sample from 1940 onwards we used the 1% IPUMS samples of the U.S. Census. We limited the sample to full-time year-round workers because hourly wages are not reported. Even with this restriction, there are some issues as has been noted by all who use this data. In particular, individuals report earnings from the previous year, weeks worked last year, and hours worked last week. We included earnings from those individuals who worked 35 or more hours last week and 40 or more weeks last year. From 1980 onwards, individuals are asked to report the "usual hours worked in a week last year." Hence for these years we require that people answer 35 or more hours to that question and we drop the restriction on hours worked last week. In 1960 and 1970, the weeks and hours worked information was reported in intervals. We take the midpoint of each interval for those years.

Sample weights (PERWT) were used as required in 1940, 1990, 2000. In 1950 sample line weights were used since earnings and weeks worked are sample line questions. The 1960-1980 samples are designed to be nationally representative without weights.

For the LFP numbers we used the 1% IPUMS samples for 1880, 1900-1920, 1940-1950, 1980-2000, and the 0.5% sample in 1930 and the 1970 1% Form 2 metro sample. For 1890, we use the midpoint between 1880 and 1900.[55] We restricted our sample to married white women (with spouse present), born in the US, between the ages of 25 and 44 who report being in the labor force (non-farm occupations and non-group quarters).

6.2 Calibration of the learning model

In order to estimate $\lambda_0, \sigma_\epsilon, \sigma_\eta, \beta_H, \beta_L$, and σ_l we minimized the sum of the squared errors between the predicted and actual values of our calibration targets (see table 1). All statistics were weighted equally.

The simplex algorithm was used to search for an optimal set of parameters. Multiple starting values throughout the parameter space were tried (specifically over 2,000 different starting values with λ_0 ranging between [-10, -.01], σ_ϵ in [0.1, 5], σ_η in [0.01, 2], σ_l between [0.5, 4], β_L in [.01, 1], and β_H to be between [1, 10] units greater than β_L.

A period is 10 years. 500 different public shocks were generated for each period (these draws were held constant throughout the minimization process). For each shock, there is a corresponding public belief that subjects begin the next period with. For each belief, a different percentage of women will choose to work after they receive their private signals.

300 discrete types were assumed between $\underline{l}(w_h, w_f)$ and $\bar{l}(w_h, w_f)$ in each year to approximate the integral in equation 16. Then we average over the η shocks to determine the expected number of women working. We then back out the belief that would lead to exactly that many women working. This determines the path of beliefs.

[55] The individual census data is missing for this year.

The elasticities were calculated computationally by assuming either a 1% increase in female earnings or male earnings and calculating the corresponding changes in LFP predicted by the model in those histories in which the (original) predicted LFP was close to the true LFP value (specifically those histories in which the predicted LFP was within ± .05 of the true LFP that year). These elasticities were calculated individually for all histories meeting this criterion and were then averaged.

In order to approximate the integrals that are needed to compute $\Pr(DW_t|MW_{t-2})$ and $\Pr(DW_t|MNW_{t-2})$, 400 discrete signals from $\beta_L - 4\sigma_\epsilon$ to $\beta_L + 4\sigma_\epsilon$ were used.

Lastly, in the partial calibration of the learning model to the same three statistics as in the earnings only model, we estimated $\lambda_0, \sigma_\epsilon, \sigma_\eta, \beta_H, \beta_L$, and σ_l by minimizing the sum of the squared errors between predicted and actual LFP (12 observations).

References

[1] Akerlof, G. and Kranton, R. E. (2000), "Economics and identity," *Quarterly Journal of Economics* 115, 715–33.

[2] Albanesi, Stefania and Claudia Olivetti (2006), "Home Production, Market Production and the Gender Wage Gap: Incentives and Expectations," Working Paper.

[3] Albanesi, Stefania and Claudia Olivetti (2007), "Gender Roles and Technological Progress," working Paper.

[4] Amador, Manuel and Pierre-Oliver Weill (2006), "Learning from Private and Public Observations of Others' Actions," UCLA working paper.

[5] Antecol, H. (2000), "An Examination of Cross-Country Differences in the Gender Gap in Labor Force Participation Rates," *Labour Economics* 7, 409-426.

[6] Adshade, Marina (2007), "Female Labor Force Participation in an Era of Organizational and Technological Change," Queen's Economics Department wp 1130.

[7] Bandiera, Oriana and Imran Rasul (2006), "Social Networks and Technology Adoption in Northern Mozambique," *Economic Journal*, Forthcoming

[8] Belsky, Jay, Deborah Lowe Vandell, Margaret Burchinal, K. Alison Clarke-Stewart, Kathleen McCartney, and Margaret Tresch Owen (2007), "Are There Long-Term Effects of Early Child Care?" *Child Development* 78 (2), 681-701.

[9] Bernal, Raquel and Michael Keane (2006), "Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers," Working Paper.

[10] Bernal, Raquel (2007), "The Effect of Maternal Employment and Child Care on Children's Cognitive Development," Northwestern Working Paper.

[11] Bikhchandani, S., D. Hirshleifer, and I. Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992-1026.

[12] Bisin, A., and T. Verdier (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics* 115(3), 955-988.

[13] Blau, Francine D. and Lawrence M. Kahn (2006), "Changes in the Labor Supply Behavior of Married Women: 1980-2000," IZA Discussion Paper.

[14] Bowles, Sam (1998), "Endogenous preferences: the cultural consequences of markets and other economic institutions," *Journal of Economic Literature* 36, 75–111.

[15] Burda, Michael, Daniel Hamermesh, and Philippe Weil (2007), "Total Work, Gender and Social Norms," NBER wp 13000.

[16] Chamley, Christophe (1999), “Coordinating Regime Switches,” *Quarterly Journal of Economics* 114, 869-905.

[17] Chamley, Christophe (2004), *Rational Herds*, Cambridge University Press.

[18] Conley, T. and Chris Udry (2003), “Learning About a New Technology: Pineapple in Ghana,” mimeo Yale University.

[19] Costa, Dora (2000), “From Mill Town to Board Room: The Rise of Women’s Paid Labor,” *The Journal of Economic Perspectives* 14(4), 101-122.

[20] Doepke, Matthias and Michele Tertilt (2007), “Women’s Liberation: What Was in It for Men?,” working paper.

[21] Edlund, Lena and Rohini Pande (2002), “Why have Women Become Left-Wing: the Political Gender Gap and the Decline in Marriage,” *Quarterly Journal of Economics*, 117, 917-961.

[22] Farré-Olalla, Lidia and Francis Vella (2007), “The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labor Force Participation,”. IZA Discussion Paper No. 2802.

[23] Fernández, Raquel (2007a), “Culture and Economics," in the *New Palgrave Dictionary of Economics*, 2nd edition, edited by Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan (Basingstoke and New York), forthcoming.

[24] Fernández, Raquel (2007b), “Women, Work, and Culture,” *Journal of the European Economic Association*, forthcoming.

[25] Fernández, Raquel and Alessandra Fogli (2005), “Culture: An Empirical Investigation of Beliefs, Work, and Fertility,” NBER Working Paper No. 11268.

[26] Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti (2004), “Mothers and Sons: Preference Formation and Female Labor Force Dynamics,” *Quarterly Journal of Economics* 119(4), 1249-1299.

[27] Fernández, Raquel and Elizabeth Potamites (2007) “Learning, Social Networks, and Changes in Female Labor Force Participation,” work in progress, NYU.

[28] Fogli, Alessandra and Laura Veldkamp (2007), “Nature or Nurture? Learning and Female Labor Force Participation,” Working Paper.

[29] Fortin, Nicole (2005), "Gender Role Attitudes and the Labour Market Outcomes of Women Across OECD Countries," *Oxford Review of Economic Policy*, 21, 416-438.

[30] Foster, Andrew and Mark Rosenzweig (1995), “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture,” *Journal of Political Economy* 103, 1176-1209.

[31] Gallup, George H. (1972), *The Gallup Poll; public opinion, 1935-1971* New York: Random House.

[32] Galor, Oded and David Weil (1996), "The Gender Gap, Fertility, and Growth," *American Economic Review* 86(3), 374-87.

[33] Gayle, George-Levi and Limor Golan (2006), "Estimating a Dynamic Adverse Selection Model: Labor Force Experience and the Changing Gender Earnings Gap," Carnegie Mellon Working Paper.

[34] Goldin, Claudia (1990), *Understanding the Gender Gap: An Economic History of American Women*, New York: Oxford University Press.

[35] Goldin, Claudia and Lawrence Katz (2002), "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy* 110(4), 730-770.

[36] Goldin, Claudia and Robert A. Margo (1992),"The Great Compression: The Wage Structure in the United States at Mid-century," *The Quarterly Journal of Economics*, 107(1), 1-34.

[37] Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu (2005), "Engines of Liberation," *Review of Economic Studies* 72(1) 109-133.

[38] Griliches, Zvi (1957), "Hybrid Corn: An Exploration in the Economics of Technological Change", *Econometrica* 25, 501-22.

[39] Guiso, Luigi, Paola Sapienza and Luigi Zingales (2006), "Does Culture Affect Economic Outcomes," *Journal of Economic Perspectives* 20(2), 23-48.

[40] Guner, Nezih and Jeremy Greenwood (2006), "Social Change," working paper.

[41] Jones, Larry, Rodolfo Manuelli, and Ellen McGrattan (2003), "Why are Married Women Working so much?" Working Paper.

[42] Katz, Lawrence and David Autor (1999), "Changes in the wage structure and earnings inequality," in O. Ashenfelter and D. Card (ed.), *Handbook of Labor Economics*, edition 1, vol. 3, chap. 26, 1463-1555, Elsevier.

[43] Knowles, John (2007), "High-Powered Jobs: Can Contraception Technology Explain Trends in Women's Occupational Choice?" Working Paper.

[44] Levine, David (1993), "The Effect of Non-Traditional Attitudes on Married Women's Labor Supply," *Journal of Economic Psychology* 14, 665-679.

[45] Mira, Pedro (2005), "Uncertain infant mortality, learning and life-cycle fertility," working paper.

[46] Munshi, Kaivan (2004), “Social Learning in a Heterogeneous Population: Social Learning in the Indian Green Revolution,” *Journal of Development Economics,* 73, 185-213

[47] Munshi, Kaivan and Myaux, J. (2006), “Social norms and the fertility transition,” *Journal of Development Economics,* 80, 1– 38

[48] Ruhm, Christopher J. (2004), “Parental Employment and Child Cognitive Development,” *Journal of Human Resources* 39(1), 155-192.

[49] Smith, Lones and P. Sorensen (2001), “Pathological Outcomes of Observational Learning,” *Econometrica*, 68, 371-398.

[50] Tabellini, Guido, (2007), “The Scope of Cooperation: norms and incentives,” working paper, IGIER.

[51] Vella, Francis (1994), “Gender Roles and Human Capital Investment: The Relationship between Traditional Attitudes and Female Labour Force Performance,” *Economica* 61, 191-211.

[52] Vives, Xavier (1993), “How Fast do Rational Agents Learn,” *The Review of Economic Studies*, 60 (2), 329-347.