EFABG
7/17/07
1:30 pm

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

Raquel Fernández
NYU, CEPR and NBER

Revised July 2007†

Women's labor force participation has increased dramatically over the last century. Why this has occurred has been the subject of much debate. This paper investigates the role of culture as learning in this change. 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. They then make a work decision. 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. When either small or large proportions of women work, learning is very slow and the changes in female labor force participation are also small. When the proportion of women working is close to 50%, rapid learning and rapid changes in female LFP take place. I calibrate the model to several key statistics and show that it does a very good job in replicating the quantitative evolution of female labor force participation 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.

†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, Christophe Chamley and Gianluca Violante for helpful comments, and the NSF for financial support.

# 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, changed preferences/skills transmitted from working mothers to their sons, 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 why these changes affected work behavior in such a gradual and uneven fashion. Indeed, one might be tempted, as surely some are, to dismiss these shifts in beliefs as mere changes in the superstructure of the economy that simply accompany and reflect the changes in

[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), Fernández, Fogli, and Olivetti (2004), Greenwood, Seshadri, and Yorukoglu (2005), and Albanesi and Olivetti (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 variation in past values of female LFP in their parents' country of origin. 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.

[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).

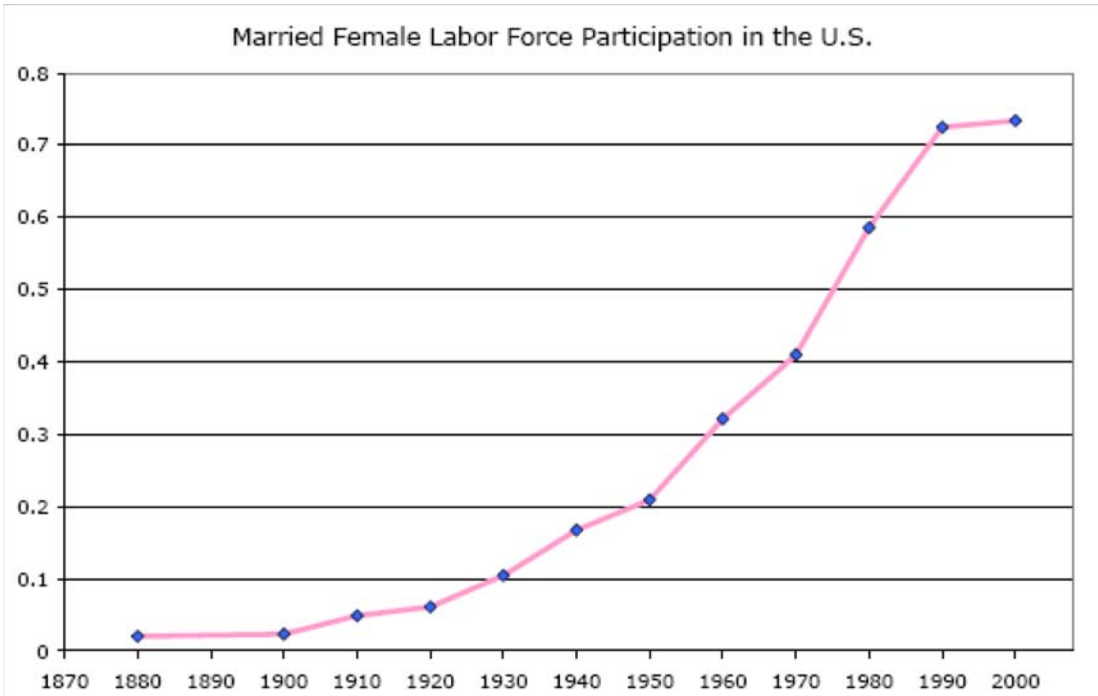


Figure 1: U.S. Census data 1880-2000. Percentage of white, married (spouse present) women born in the U.S. 25-44 years old (non-agricultural, non-group quarters) who report being in the labor force.

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 these 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]

This paper puts forward and develops 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. In particular, it takes seriously the fact that women's labor force participation changed in a very uneven fashion over time in a form that resembles an "S-shape". 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 (though on

[6] See, e.g., Guner and Greenwood (2006) who argue that the change in sexual mores reflect changes in the efficacy of contraception. This is no doubt a partial explanation but does not explain, for example, why attitudes towards homosexuality have changed.

[7] For an interesting example of the latter theory applied to culture see Bowles (1998).

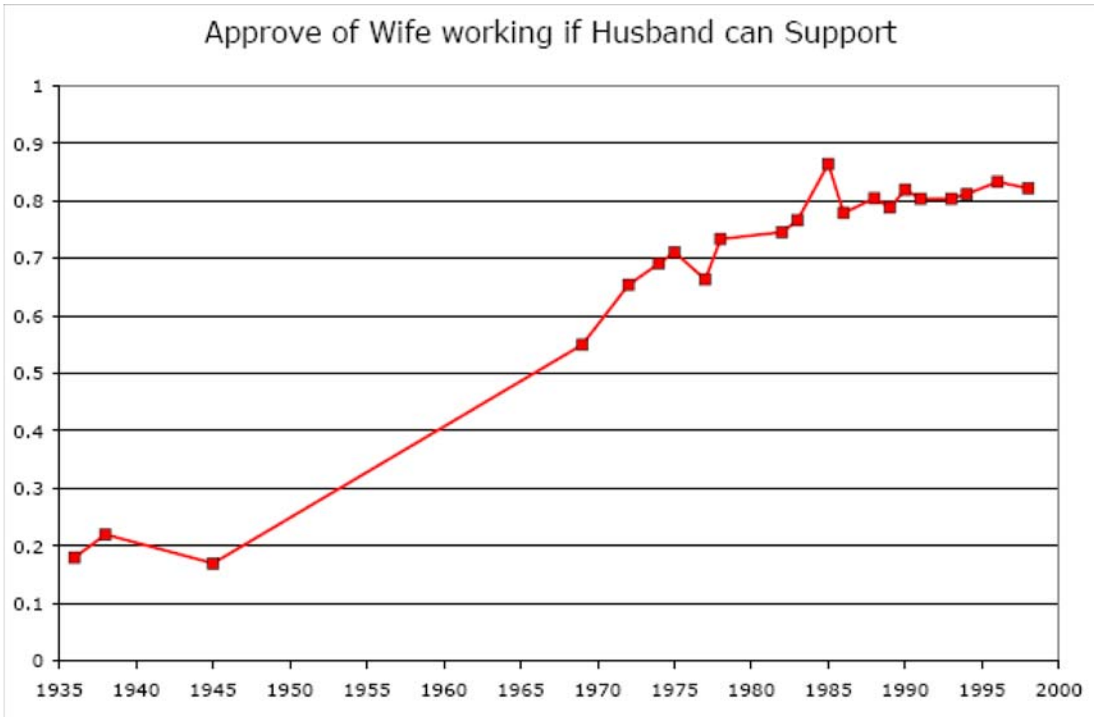


Figure 2: Sources: 1936-1938 and 1969 numbers are from the Gallup Poll (1972), 1945 is from Benjamin I. Page and Robert Y. Shapiro, The Rational Public, University of Chicago Press, 1992; pp. 101, 403-4. 1972 onwards are from the General Social Survey.

a very different time scale) is also at play in this context.[8],[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, the payoff to women's work has been the subject of great contention and fraught with uncertainty. Industrialization and urbanization contributed to the separation of the spheres of work and home, introducing far greater specialization, drawing in younger men and (unmarried) women into the paid workplace and away from sharing household chores, and kicking off a debate on the effect of working on a woman's marriage, on her psyche and image (and on those of her husband's) and, more recently, on her children's welfare that continues until this day.[10] Goldin (1990) notes that at the turn of the 20th century, most working women were employed as domesic

[8]For examples of learning contributing to technology adoption see, e.g., Griliches (1957), Foster and Rosenzweig (1995), Conley and Udry (2003), Munshi (2004, 2006), and Bandiera and Rasul (2006). 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.

[9]A recent paper by Fogli and Veldkamp (2007) independently develops a similar idea. They study the labor force participation of women with children from 1940-2000 and assume that agents are learning the true cost to a child's ability from having a working mother by sampling the outcomes of a small number of other women. Munshi and Myaux (2006) model the change in contraceptive practice in rural Bangladesh as learning about the the preferences of individuals in one's social network. As in Fogli and Velkamp, they use a matching model (only one individual is sampled) and there is, in addition, a strategic aspect to individual choices. 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.

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

servants or in manufacturing. Thus, a married woman's employment served as a signal that her husband was unable to provide adequately for his family and most women exited the workplace upon marriage.[11] In more recent times, the locus of the debate has shifted to the effect of a working mother on a child's intellectual achievements and emotional health. Although the research evidence on this topic is far from conclusive, it nonetheless commands a great deal of public attention.[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, despite showing only a small quantitative effect.

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] Using a framework broadly similar to Vives (1993) and Chamley (1999), I assume that women receive a private signal about how costly it is to work (e.g., how negative the outcome is for one's marriage, children, etc.) and that they also observe a noisy public signal about 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 payof 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 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 in technology, which to my

[11] According to figures from the 1939 Retrospective Survey, in 1939, of all married women between the ages of 40-49 not working currently but who had worked at some point prior to marriage, more than 80% exited the workplace at the precise time of marriage (Source: Goldin (1990) p. 34).

[12] See, for example, Bernal (2007), Keane and Bernal (2005), Hill, Waldfogel, Brooks-Gun and Han (2005), and Ruhm (2006) for reviews and recent findings of this literature. The level of attention devoted to evidence in this area is tremendous. As an interesting indication of culture, note that the effect of having a working father is rarely investigated.

[13] Whether preferences or beliefs changed is often impossible to distinguish (and may often be an invalid question as the two may influence one another). In a reduced-form setup, it is not necessary to distinguish between the two. For a model of learning, however, the distinction is important as beliefs will evolve in a Bayesian fashion. Furthermore, it is straightforward to think about welfare properties in a model in which changes in beliefs that do not otherwise affect preferences. If preferences themselves change, however, this requires a non-standard welfare treatment (see Fernández (2007b) for a discussion of these issues).

knowledge has not been noted in the learning literature. Unlike in traditional models, new technologies that make it easier for women to work outside the home or increases in female wages, 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.[14] 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 model indicates that both the dynamic paths of beliefs and earnings played an important role in the transformation of women's work. From the decades between 1880-1950 the growth in female LFP was small, and most of the change in LFP comes from the gradual evolution of beliefs, independently of any changes in wages. Both the dynamic and static effects of wage changes played a role in increasing female LFP from 1950 to 1970, and from 1970-1990 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.

## 2 A Simple Model of a Woman's Work Decision

We start with a very simple model of a woman's work decision. We include the main variables that are typically assumed to play a role in this decision, namely her consumption possibilities as a function of her decision and her disutility from working. As we are interested in the difference in the *long-run* payoffs from working versus not working, we view the disutility from working as stemming not only from labor-leisure preferences, but also from what might happen to her identity, marriage, or her children as a result of decision. In this first model, we 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:

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

where $\gamma \geq 0$ and $\mathbf{1}$ is an indicator function that takes the value one if she works and zero

[14] 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.

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 $\beta$ 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 $l_i$ 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, $\frac{\partial l^*}{\partial w_h} < 0$.

# 3 The Simple Work Model with Learning

We next incorporate beliefs and learning in the simple model above. In particular, women are assumed to be uncertain about the common value of the disutility of labor, $\beta$, 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 or by experimentation but rather reveals its effects over a lifetime.

For simplicity, we assume that $\beta$ can take on only two values, high (H) and low (L), i.e., $\beta \in \{\beta_H, \beta_L\}$.[15] Note that $\beta_L$ is the "good" state of nature in which working is not so costly, i.e., $\beta_H > \beta_L > 0$. An individual woman now makes her work decision to maximize

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

her expected utility, i.e., (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.

The model incorporates two sources of learning. One is an individual source whereby a woman receives a noisy private signal regarding the true value of $\beta$, $\beta^*$. The second is an intergenerational source whereby all women in generation $t$ observe a noisy public 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 $\beta^*$ 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 $\beta^*$. 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 mother's work on children or families) and can be either newly generated each period or inherited from the woman's mother.[16] 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. 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

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

where $\bar{\beta} = (\beta_L + \beta_H)/2$.[17] Note that $\frac{\partial \lambda_{it}(s)}{\partial s} < 0$ since higher signals increase the likelihood that the true value of $\beta$ is $\beta_H$. Note also that the revision of $\lambda$ is decreasing with the variance of the noise term, $\sigma_\varepsilon^2$, since it lowers the informativeness of the signal.

Assume that women have a common prior in period $t$, $\lambda_t$.[18] 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 $\frac{1}{1-\gamma}[(w_{ht}+w_{ft})^{1-\gamma}-w_{ht}^{1-\gamma}]$ by $W(w_{ht}, w_{ft})$.

[16] To calibrate the model to the conditional probability that a woman works given that her mother worked, we use the latter interpretation.

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

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

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 $l_j$, $\underline{l} < l_j < \bar{l}$, we 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 $\lambda$, 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

$$p_j^*(w_{ht}, w_{ft}) = \frac{l_j - \underline{l}(w_{ht}, w_{ft})}{\beta_H - \beta_L} \tag{13}$$

and hence, $\frac{p_j^*}{1-p_j^*} = \frac{l_j - \underline{l}}{\beta_H - \beta_L + \underline{l} - l_j} = \frac{l_j - \underline{l}}{\bar{l} - l_j}$.

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 $\lambda_t$, is given by

$$\lambda_t(s_j^*) = \lambda_t + \left(\frac{\beta_L - \beta_H}{\sigma_\epsilon^2}\right)(s_j^* - \bar{\beta}) = \ln\left(\frac{l_j - \underline{l}}{\bar{l} - l_j}\right) \tag{14}$$

and hence

$$s_j^*(\lambda_t; w_{ht}, w_{ft}) = \bar{\beta} + \left(\frac{\sigma_\epsilon^2}{\beta_H - \beta_L}\right)\left(\lambda_t + \ln\left(\frac{\bar{l}(w_{ht}, w_{ft}) - l_j}{l_j - \underline{l}(w_{ht}, w_{ft})}\right)\right) \equiv s_j^*(\lambda_t) \tag{15}$$

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

$$\omega_{jt}(\beta; \lambda_t) = F(s_j^*(\lambda_t) - \beta; \sigma_\epsilon) \tag{16}$$

Thus, the total proportion of women that will work in period $t$ if the true state of nature is $\beta$, is given by:

$$\omega_t(\beta; \lambda_t) = G(\underline{l}) + \int_{\underline{l}}^{\bar{l}} F(s_j^*(\lambda_t) - \beta; \sigma_\epsilon) g(l_j) dl_j \tag{17}$$

where $g(\cdot)$ is the pdf of the $l$ distribution $G(\cdot)$.

### 3.1 Intergenerational Transmission

What information is passed on from generation $t$ to generation $t+1$? We assume that each woman passes on to her child her prior, $\lambda_{it}(s)$. Alternatively, generation $t+1$ inherits the

Figure 3: Timeline of Learning Model

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

***Private Learning*** ***Social Learning***

prior of generation $t$ (its "culture"), $\lambda_t$, which each individual then updates with her private signal (assumed to be either inherited from her mother or the result of a new random draw $s$). If this is all the information that was transmitted intergenerationally, then the learning model would behave in the same way as the earnings only model; 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 the aggregate proportion of women who worked in period $t$, $\omega_t$, they would be able to back out the true state of nature, $\beta^*$, as a result of the law of large numbers. While assuming that that this information is totally unavailable seems extreme, the notion that this knowledge is completely informative seems equally implausible. We employ instead the conventional tactic in this literature and assume that women are able to observe a noisy function of the aggregate proportion of women worked.[19] One way to think about this 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 which is a very reasonable assumption. Alternatively, one could model individuals as observing LFP perfectly, but being uncertain about the distribution of an idiosyncratic utility factor whose distribution could change randomly every period (e.g. by depending on an unobservable aggregate factor in the economy). The route chosen below saves on a considerable amount of additional notation.

In particular, we assume that women observe a noisy signal, $y_t$, of $\omega_t$, where

$$y_t\left(\beta;\lambda_t\right) = \omega_t\left(\beta;\lambda_t\right) + \eta_t \tag{18}$$

and where $\eta_t \sim N(0, \sigma_\eta^2)$ with a pdf denoted by $h(\cdot;\sigma_\eta)$. Thus, given a common inherited prior of $\lambda_t$, after observing last period's signal of aggregate female LFP, $y_t$, Bayes' law

[19] 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 alternative assumption of information obtained from a limited (discrete) sample yields similar results. It has the advantage, for the calibration, of not requiring a specification of an aggregate shock. Amador and Weill (2006) use a matching model with private and public observation of actions that also yields an S shaped curve.

implies an updated common belief for generation $t+1$ of:

$$\begin{aligned}\lambda_{t+1}(\lambda_t, y_t) &= \lambda_t + \ln \frac{h(y_t|\beta^* = \beta_L)}{h(y_t|\beta^* = \beta_H)} \\ &= \lambda_t + \left(\frac{\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)}{\sigma_\eta^2}\right)\left(y_t - \frac{\omega_t(\beta_L;\lambda_t) + \omega_t(\beta_H;\lambda_t)}{2}\right) \quad (19)\end{aligned}$$

Note that (19) is the law of motion of aggregate beliefs (culture) for the economy.

Figure 3 summarizes the time line for the economy. Individuals start period $t$ with a common (updated) prior, $\lambda_t$. Each woman then updates the common prior with her (inherited or observed) private signal and makes her work decision, generating an aggregate $\omega_t$ and a noisy signal $y_t$. Generation $t+1$ observes $y_t$ and uses it to update the old common prior ($\lambda_t$), generating $\lambda_{t+1}$ – the "culture" of generation $t+1$.[20] The process continues as described in each period.

It should be noted that instead of assuming women in $t+1$ inherit $\lambda_t$ which they update with the information contained in $y_t$, we can assume that women observe the entire history of $y_\tau$, $\tau = 0, 1, 2..., t$.

## 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., $\frac{\partial \omega_t(\beta;\lambda_t)}{\partial w_{ft}} > 0$, $\frac{\partial \omega_t(\beta;\lambda_t)}{\partial w_{ht}} < 0$), the learning model has several important properties that will be prove useful when we try to match the data in figure 1.

Note first that beliefs in this model are unbounded. Hence, in the long run beliefs must converge to the truth.[21] 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 will naturally generate an S-shaped curve. To see why, note that given $\beta^* = \beta_L$, we can rewrite (19) as

$$\lambda_{t+1} = \lambda_t + \left(\frac{\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)}{\sigma_\eta^2}\right)\left(\eta_t + \frac{\omega_t(\beta_L;\lambda_t) - \omega_t(\beta_H;\lambda_t)}{2}\right) \quad (20)$$

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 $\beta^* = \beta_H$.

To understand when the 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 $l_j \in (\underline{l}, \bar{l})$ type this difference is equal to:

$$F(s_j^*(\lambda_t) - \beta_L; \sigma_\epsilon) - F(s_j^*(\lambda_t) - \beta_H; \sigma_\epsilon) \quad (21)$$

[20] Thus, we can think of generation $\tau$ as having a shared culture given by $\lambda_\tau$ with the individual deviations around $\lambda_\tau$ (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 $s$).

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

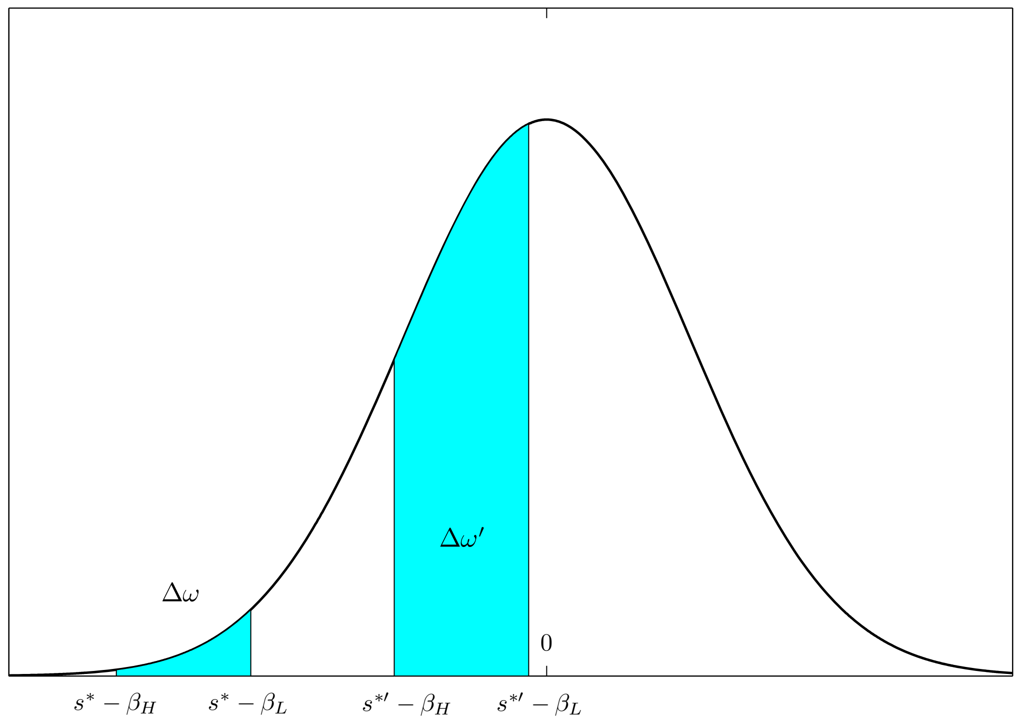


Figure 4: Normal PDF

Taking the derivative with respect to $s_j^*$ yields the f.o.c.

$$f\left(s_j^* - \beta_L\right) - f\left(s_j^* - \beta_H\right) = 0 \tag{22}$$

Recalling that $f\left(s_j^* - \beta\right) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp - \left\{\left(\frac{(s_j^* - \beta)^2}{2\sigma_\varepsilon^2}\right)\right\}$, (21) is minimized when $s_j^* = \pm\infty$ and it is at a maximum at $s_j^* = \overline{\beta}$.

Thus, if the critical signal $s_j^*(\lambda_t)$ is far from the $\beta'$s in absolute value, (21) 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 aggregate noise term $\eta_t$. 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. 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 $\omega_j$ at $s_j^* - \beta_L$ and $s_j^* - \beta_H$, (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 $s^{*\prime}$. Again, as shown in the figure for the same two values of $\beta$, when $s^{*\prime} - \beta$ is close to zero, the difference between $\omega_j'$ 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 (17) we obtain

$$\frac{\partial \omega_t}{\partial s^*} = \int_{\underline{l}}^{\overline{l}} \left[ f\left(s_j^*(\lambda_t) - \beta_L\right) - f\left(s_j^*(\lambda_t) - \beta_L\right)\right] g(l_j) dl_j \tag{23}$$

Thus, if the critical signal $s_j^*(\lambda_t)$ is, for the average individual, far from $\beta$, (23) will be small in absolute value, intergenerational updating will be small, and the evolution over time will be slow.[22]

Under what circumstances will the critical signal take an extreme value and lead to slow intergenerational learning? As can be seen from expression (15), for a given $l_j$ type, this occurs when $\lambda_t$ is either very small or very large. To understand why this is the case, note that when an individual assigns a very low (high) probability to $\beta^* = \beta_L$, it takes a very low (high) realization of the private signal to update her belief sufficiently to convince her to work (not work). The difference across states of nature, however, in the probability of obtaining such extreme values of $s$ is very small. Once again, this is the result of the shape of the cdf of a normal distribution which is very flat for extreme (positive or negative) values of $s$. As before (see footnote 22), our assumption of a normal distribution of types implies that what matters is when $l_j$ types close to 0 require extreme signal values, and this will be true, ceteris paribus, when $\lambda_t$ is very small or very large (and $l_j = 0 \in (\underline{l}, \overline{l})$).

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 $\beta^* = \beta_L$ whereas many women would choose to work if they assigned a high probability to this state, then the amount of intergenerational learning that will happen when female LFP is either very low or very high will be small as the aggregate noise term dominates in (19) and hence the period to period increase in female LFP will be likewise small. At these extremes, 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 $l_j = 0$, 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.

## 3.3 Wages, Technology, and Learning

The learning model generates a novel role for changes in wages or for technological change that makes it easier for women to work outside the household (e.g., the washing machine in Greenwood et al (2005) or the introduction of infant formula as in Albanesi and Olivetti

[22] 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 $\omega$ to be large, we need $s_j^*$ to be close to $\beta$ for types with a large frequency, i.e., individuals close to $l_j = 0$.

(2007)).[23] An increase in female wages, for example, has as before its static effect of increasing female LFP, i.e., $\frac{\partial \omega_t}{\partial w_{ft}} > 0$. In this model, it will futhermore have a dynamic effect; it will also affect the amount of intergenerational updating that takes place, i.e., $\lambda_{t+1}$. 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 induced increase in $s^*$ 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 $\lambda$ is low) and hence increases the informativeness of the aggregate signal for the next generation. Thus, increases in female earnings and changes in technology or in 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 it has 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 might be quantitatively important.

# 4 Empirical Analysis

In this section we 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 we calibrate to three key statistics of female LFP in the year 2000. This gives us a benchmark by which to measure how much the incorporation of beliefs that evolve endogenously over time is able to add to the ability of the model to replicate the data. We next calibrate the learning model to four additional statistics and show that the fully calibrated model does a good job of fitting the historical LFP series. We conclude by examing the roles of beliefs relative to wages in the evolution of female LFP and distinguish between the static and dynamic contribution of the changes in earnings.

## 4.1 The Data

Our model requires historical data on earnings for men and women as well as female LFP statistics. It is difficult to obtain earnings data prior to 1940. For this time period, we rely on data 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. We use the data for white men and women.[24] As Goldin does not provide data for earnings in 1880 and 1910, we construct these using a cubic approximation with the data from 1890 -1930 (inclusive).

[23] While this role exists in

[24] See Goldin (1990) pages 64-65 and 129 for greater detail about the earnings construction for various years. We look at white women as black women have had a different LFP trajectory, with much higher participation rates earlier on.

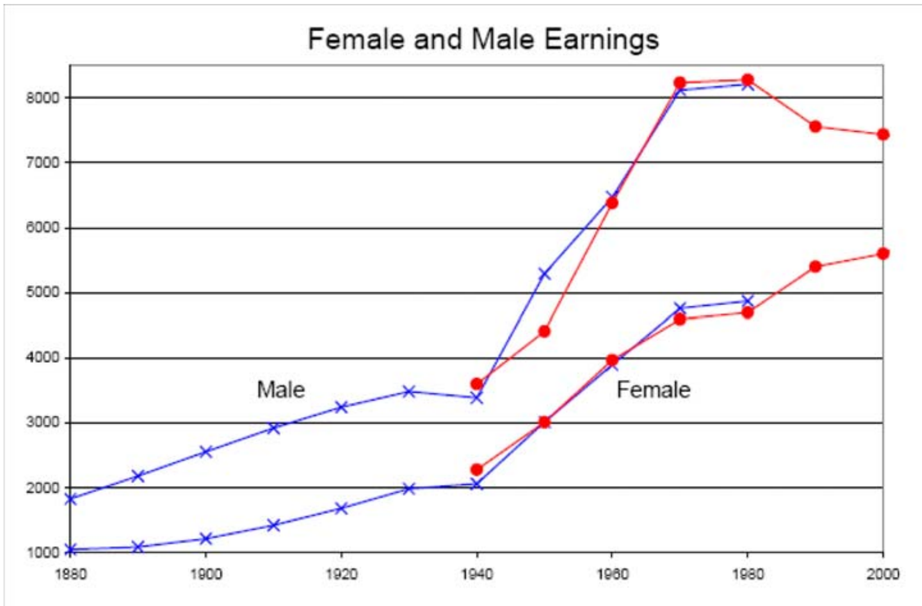


Figure 5: Crosses (blue) represent the yearly median earnings data from Goldin (1990), Table 5.1. Dots represent our calculations using U.S. Census data (red). They are the median earnings of white men and women between the ages of 25-44 in non-farm occupations and not living in group quarters. All earnings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

As of 1940 we use the 1% IPUMS samples of the U.S. Census for yearly earnings (incwage) and calculate the median earnings of white 25-44 year 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.[25] As is commonly done, we exclude observations that report weekly earnings less than half the minimum wage. We use half the nominal minimum wage times 35 hours a week as our cutoff for weekly wages and calculate nominal weekly wages by dividing total wage and salary income last year by weeks worked last year.[26]

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 our data with Goldin's we also plot her figures (which continue to 1980). These numbers are show in blue. The only significant difference is with male earnings in 1950 which are higher for Goldin.[27]

[25] We limited our sample to full-time year-round workers because hourly wages are not reported. Even with this restriction, the usual issues remain (see Appendix).

[26] 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, as in Goldin and Margo (1992), we use the same cutoff rule as a way to eliminate unreasonably low wages. Note that since we are calculating median earnings, we do not have to concern ourselves with top-coding in the Census.

[27] 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.

The LFP numbers 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).

We calibrate both models 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, we also match the cross-wage elasticity in 1990, female LFP in 1990, the probability that a woman worked in 1980 conditional on her mother working, and female LFP in 1980. See table 1 for a list of the targets.

## Table 1

| Parameter | Earnings Model | Partially Calibrated | Learning Model | | Calibration Targets |
|---|---|---|---|---|---|
| $\gamma$ | 0.503 | 0.503 | 0.503 | 0.30 | Own-Wage Elasticity (2000) |
| $\sigma_L$ | 2.293 | 2.067 | 2.052 | -0.13 | Cross-Wage Elasticity (2000) |
| $\beta$ | 0.321 | | | 0.734 | Female LFP (2000) |
| $\beta_H$ | | 7.481 | 3.701 | | |
| $\beta_L$ | | .0004 | 0.000 | 0.725 | Female LFP (1990) |
| $P_0(\beta = \beta_L)$ | | 0.110 | 0.048 | -0.14 | Cross-Wage Elasticity (1990) |
| $\sigma_\epsilon$ | | 5.408 | 5.746 | 0.586 | Female LFP (1980) |
| $\sigma_\eta$ | | 0.157 | 0.026 | 0.595 | $Pr(DW\|MW)$ in 1980 |

All elasticities are from Blau & Kahn (2006). $Pr(DW|MW)$ is calculated from GSS. The earnings model and the partially calibrated model are calibrated to first 3 targets. The learning model is calibrated to all the targets.

For the elasticity estimates we use those reported in Blau and Kahn (2006). The authors use the March CPS 1989-1991 and 1999-2001 and compute married women's own-wage and husband's-wage elasticities along the extensive margin restricting their sample to married women of age 25-54 (with spouses in the same age range).[28] We 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 we are more interested in an elasticity with respect to some measure of lifetime earnings. The specification we chose also did not control for children which we consider an endogenous variable. Blau and Kahn estimate an own-wage elasticity of 0.30 and the cross-elasticity (husband's wage) of -0.13 for our preferred specification in the year 2000 and a cross elasticity of -0.14 in 1990.[29]

To calculate the probability that a woman worked in 1980 conditional on her mother working, we used the General Social Survey (GSS) from 1977, 1978, 1980, 1982, and 1983.

[28] 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.

[29] Using instead the specification with education controls does not affect our results; the elasticities are very similar to the ones we chose (0.28 and -0.12 for 2000 and -.15 in 1990).

We included in our sample all white married women between the ages 25-45 who were born in the U.S.[30] The GSS asked a variety of questions regarding individual's mother's work behavior. We used affirmative responses 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. Taking all women (i.e., the daughters) with working mothers, we calculated the percentage of daughters who worked for each year in our sample and averaged this percentage across years, yielding 0.60 as the conditional probability.[31]

### 4.2 Calibrating the Model Without Learning

We start out by calibrating the model without learning (which we will also call 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 $\gamma, \beta$, and $\sigma_l$.

Note that the wage elasticity (own or cross) is given by:

$$\varepsilon_k = g\left(l^*\right) \frac{\partial l^*}{\partial w_k} \frac{w_k}{\omega} \tag{24}$$

$k = f, h$. Taking the ratio of the two elasticities and manipulating the expression, yields a closed-form expression for $\gamma$, from which we can obtain a parameter value by using the data in 2000, i.e.,

$$\gamma = \frac{\log\left(1 - \frac{w_f}{w_h}\frac{\varepsilon_h}{\varepsilon_f}\right)}{\log\left(1 + \frac{w_f}{w_h}\right)} = 0.503 \tag{25}$$

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

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

where $\Phi^{-1}$ is the inverse of a standard normal distribution $N\left(0, 1\right)$. After some manipulation of (24), we obtain:

$$\sigma_l = \frac{A}{\exp\left(\frac{\Phi^{-1}(\omega)^2}{2}\right)} = 2.29 \tag{26}$$

where $A = \frac{w_f\left(w_f + w_h\right)^{-\gamma}}{\sqrt{2\pi}\varepsilon_f \omega}$. We can then solve for $\beta$ directly from the definition of $l^*$, yielding $\beta = 0.321$. To interpret this value, note that this is 9.4% of the utility from working in 1880 or 46.8% of the difference in utility between working and not working in that year.

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

This basic inability of the earnings only model to match the historical data is robust to a wide range of values for the elasticities (we explored with values twice and half that of Blau

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

[31] For daughers, a woman was defined as working if she was reported in the labor force.

and Kahn). It is also robust to alternative specifications of the shares of consumption that women obtain from their husband's earnings. In particular, we 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 $\alpha$ that vary from 0.1 to 1. As is clear from this 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 then 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.[32]

### 4.3 Calibrating the Learning Model

We now turn to calibrating the learning model. After some algebra and noting that $\frac{\partial \bar{l}}{\partial w_k} = \frac{\partial \underline{l}}{\partial w_k}$, $k = f, h$, one can show that the ratio of the elasticities in this model can be written as

$$\frac{\varepsilon_{w_f}}{\varepsilon_{w_h}} = \frac{\frac{\partial \underline{l}}{\partial w_f}}{\frac{\partial \underline{l}}{\partial w_h}} \frac{w_f}{w_h} \tag{27}$$

Noting further that $\frac{\partial \underline{l}}{\partial w_k} = \frac{\partial l^*}{\partial w_k}$, this implies that following the same manipulations as in the previous section we obtain (25), and thus the same value of $\gamma$ as in the earnings only model, i.e., $\gamma = 0.503$. We assume throughout that the true state of nature is given by $\beta^* = \beta_L$.

There is one complication in estimating this model–the presence of an aggregate 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 $\eta_t$ of the public shock generates a corresponding different public belief $\lambda_{t+1}$ in the following period, and consequently a different proportion of women who choose to work after receiving their private signals, $\omega_{t+1}$. Note that we cannot simply evaluate the model at the mean of the expected $\eta$ shocks (i.e., at zero) since, although $\lambda_{t+1}$ is linear in $\eta$, the work outcomes $\omega_{t+1}$ are not linear in $\lambda_{t+1}$.

We used the following procedure to deal with the presence of an aggregate shock. For each period $t$, given the labor force participation in the previous period $\omega_{t-1}$, we calculated the proportion of women who would work, $\omega_t$, for each realization of the shock, $\eta_{t-1}$, i.e., for each induced belief $\lambda_t \left(\eta_{t-1}\right)$. Integrating over the shocks, we found the expected value of LFP for that period, $E\omega_t \left(\lambda_t \left(\eta_{t-1}\right)\right)$, and then backed out the public belief (or shock) that would lead to exactly that same proportion of women working, i.e., we found, $\lambda_t^*(\eta_{t-1}^*)$

[32] 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.

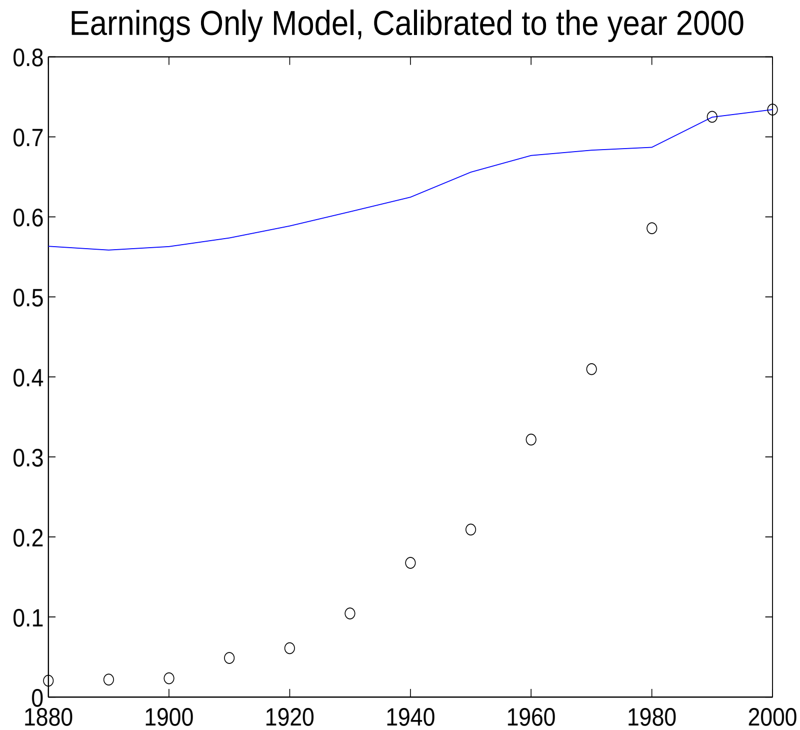


Figure 6: Parameters: $\gamma = 0.503$, $\beta = 0.321$, and $\sigma_L = 2.293$

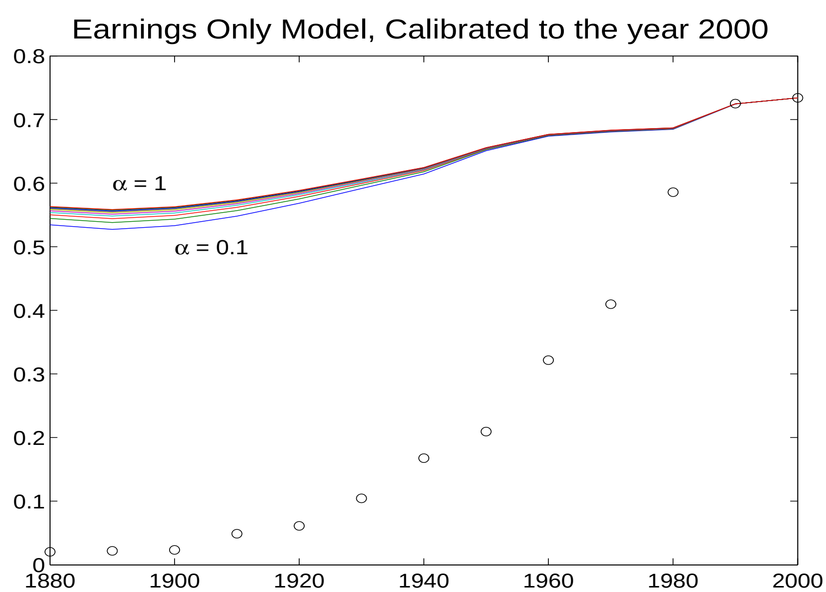


Figure 7: $\alpha$ is the fraction of husband's earnings that enter a wife's utility via consumption.

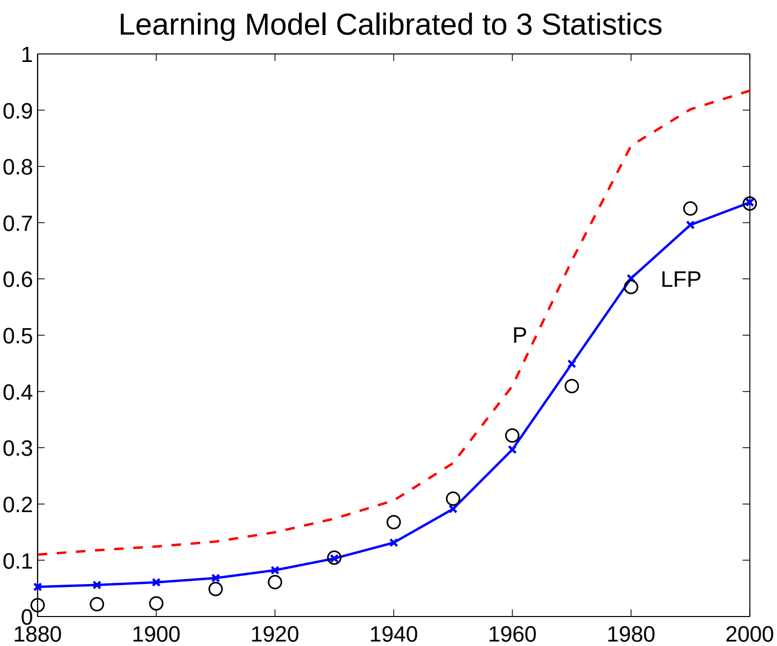


Figure 8: Dashed red line (p) is belief path. Sum of squared errors (distance of predicted LFP from actual LFP) is 0.009.

such that:[33]

$$E\omega_t\left(\lambda_t\left(\eta_{t-1}\right)\right) = \omega_t\left(\lambda_t^*(\eta_{t-1}^*)\right) \tag{28}$$

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

Before turning to the remaining calibration targets, it may be useful to first examine the maximum potential of this model by first calibrating it solely to the same set of statistics from 2000 as the earnings only model. As the earnings only model is 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 we show below, it greatly improves the ability of the model to match the data.

The results of the estimation are shown in figure 8; table 1 reports the parameter values. The blue line in figure 8 shows the evolution of the expected value of female LFP and the red dashed line(labelled $P$) shows the evolution of public beliefs, i.e., the belief, $p_t\left(\eta_{t-1}\right)$, that the true state is $\beta_L$ in period $t$ (derived from $\lambda_t$) As can be seen from figure 8, the partially calibrated model does an excellent job of replicating the LFP time series.

We now return to the full calibration of the learning model in order to impose more discipline on the free parameters of the model. In addition to the statistics discussed above, we choose to also match the cross-wage elasticity in 1990, female LFP in 1990, the conditional probability a woman works in 1980 if her mother worked, and female LFP in 1980.

[33] We take a large number of draws of entire histories for $\eta$ (500 histories). See the Appendix for details on the calibration and estimation methods.

In order to calculate the conditional probability of working for the daughter of a working mother, we need to specify an inherited characteristic, otherwise the conditional probability will be 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 passed on across generations whereas the $l_j$ type is a random draw from the normal distribution $G(\cdot)$ that is *iid* across generations.[34]

Thus, given a signal $s$ we can define $l_s$ as the $l_j$ type that is just indifferent between working and not (i.e., $s^*_{l_s} = s$). Hence, the probability that a woman with signal $s$ works is $G(l_s)$. Rearranging the expression for $s^*_j$ in (15), we obtain

$$l_{st} = \frac{\underline{l}_t + \bar{l}_t \exp(\lambda_t - \frac{\beta_H - \beta_L}{\sigma_\epsilon^2}(s - \bar{\beta}))}{1 + \exp(\lambda_t - \frac{\beta_H - \beta_L}{\sigma_\epsilon^2}(s - \bar{\beta}))} \tag{29}$$

And using Bayes rule and $\beta^* = \beta_L$, we can calculate,

$$\begin{aligned} Pr(DW_t|MW_{t-2}) &= \frac{Pr(DW_t \text{ and } MW_{t-2})}{P(MW_{t-2})} \\ &= \frac{\int_{-\infty}^{\infty} Pr(DW_t \text{ and } MW_{t-2}|s) f(s - \beta_L) ds}{\omega_{t-2}(\beta_L)} \\ &= \frac{\int_{-\infty}^{\infty} G(l_{st}) G(l_{s(t-2)}) f(s - \beta_L) ds}{\omega_{t-2}(\beta_L)} \end{aligned} \tag{30}$$

where $DW$ and $MW$ stand for daugher works and mother worked. We use the predicted LFP from two periods earlier to calculate the probability that the mothers worked (hence the $t-2$ in expressions such as $G(l_{s(t-2)})$). Note that in (30), the probability that both mother and daugher worked, $Pr(DW_t \text{ and } MW_{t-2}|s)$, is multiplied by $f(s - \beta_L)$ as this is the proportion of mothers (or daughters) who have a private signal $s$ in any time period.

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 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 $\beta_L$. The calibrated model does a good job of replicating the historical path of female LFP.[35] It over predicts LFP in the first few decades and under predicts it in 1950-1970, hitting the calibration LFP targets of 1980 and 2000, as well as 1990. See table 2 for a comparison of the calibration targets and the model's predicted values.

Individuals start out in 1880 with pessimistic beliefs about how costly it is to work. They assign around a 5% probability to the event $\beta^* = \beta_L$ and this belief evolves very slowly over the first seventy years or so (remaining no higher than 10% for this period).

[34] See Farré-Olalla and Vella (2007) for recent evidence on the correlation of mother's and daughter's attitudes towards work. Vella finds that a woman's attitudes towards work (instrumented by whether her mother worked) have important explanatory power for the variance in work outcomes.

[35] The sum of squared errors (between actual and model predicted LFP) is 0.035.

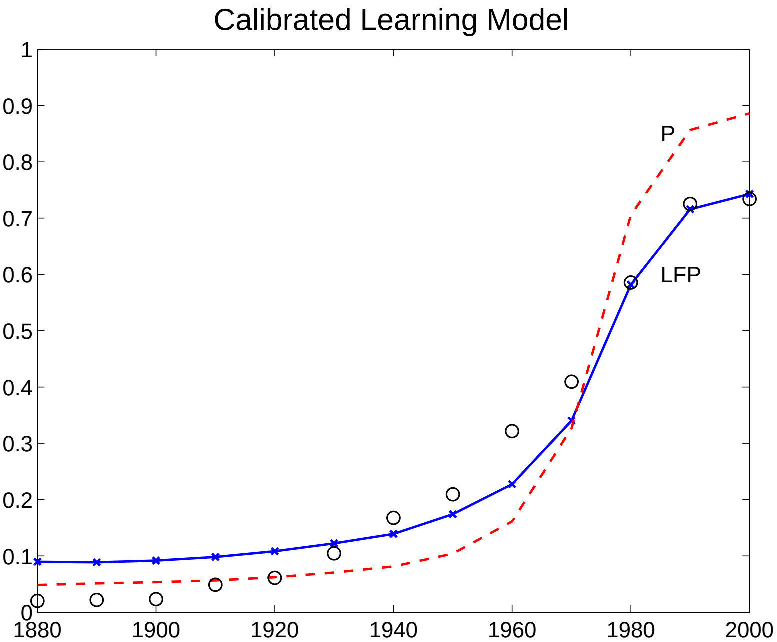


Figure 9: Dashed red line (p) is belief path. Sum of squared errors (distance of predicted LFP from actual LFP) is 0.035.

Then, as of 1950, the change in beliefs accelerate, jumping from 16.1% in 1950 to 32.7% in 1960 to 70.5% in 1970. By 2000, the public belief assigns a probability of 88.6% to $\beta^* = \beta_L$. Figure 10 shows the path of beliefs once again, but this time for the median or mean individual, as well as the individuals two standard deviations below and above the mean.[36]

It is also of interest to examine the pattern of own and cross wage elasticities predicted by the model. These are shown in figure 11. As can be seen from the picture, over time both elasticities are first increasing (in absolute value) and then decreasing. This pattern makes sense, and corresponds to the historical one reported in Goldin (1990) with respect to women's own wage elasticity. One can speculate that this pattern reflects the unwillingness of women to work in the early decades unless required by a husband's low income, and then, over time, increasing the sensitivity to her own wage as a higher payoff is associated with working. Towards the later decades there is a widespread belief that it is not bad for a woman to work (recall $\beta_L = 0$) and thus women's work behavior becomes less sensitive to both her and her husband's wages.[37]

It is interesting to compare the earnings only model with the learning model. As noted previously, the calibration implies that both models must have the same value of $\gamma$.

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

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

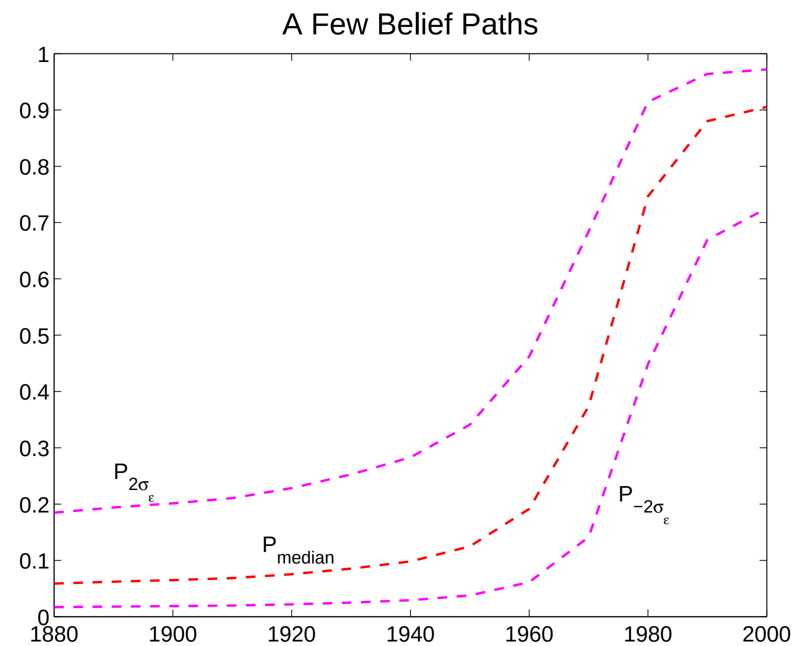


Figure 10: This shows $P(\beta^* = \beta_L)$ for agents with $s = \beta^*$ and $s = \beta^* \pm 2\sigma_\varepsilon$.

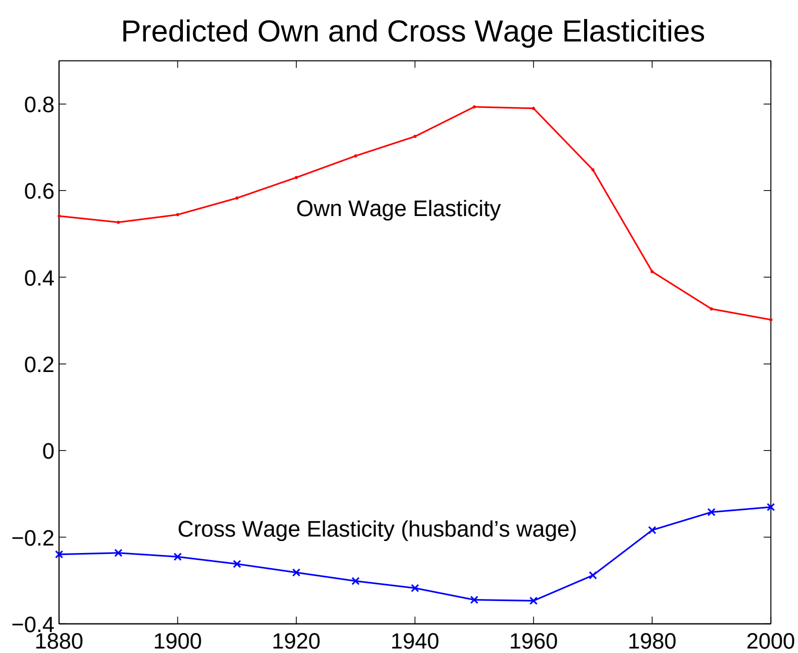


Figure 11: Parameter values from calibrated model. See appendix for how elasticities were calculated.

Furthermore, the difference in the standard deviation of the normal distribution of types is relatively small: 2.29 versus 2.05. Lastly, the expected value of $\beta$ in 2000 (a constant, of course, in the earnings only model) is also not very different across models: 0.32 as opposed to 0.42 in the learning model.[38] Thus, it is the endogenous evolution of the expected value of $\beta$ 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 $\beta$ is about 3.52 in 1880 and then evolves over time to 0.42 in 2000.

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, both models over-predict female LFP in the first few decades, though the error is greater for the calibrated model. The main decades in which the partially calibrated model does significantly better are 1960 and 1970. The requirement that the parameters be able to match the conditional probability that a woman works appears to be mainly responsible for this. In the partially calibrated model, this conditional probability is quite a bit higher than the target for the calibrated model (0.71 rather than 0.60).[39]

Table 2

| Calibration Targets | | Earnings Model | Partially Calibrated | Learning Model |
|---|---|---|---|---|
| Own-Wage Elasticity (2000) | 0.30 | **0.30** | **0.30** | **0.30** |
| Cross-Wage Elasticity (2000) | -0.13 | **-0.13** | **-0.13** | **-0.13** |
| Female LFP (2000) | 0.734 | **0.734** | **0.736** | **0.743** |
| 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 | **0.582** |
| $Pr(DW\|MW)$ in 1980 | 0.595 | 0.687 | 0.706 | **0.602** |

Statistics in bold are the calibration targets used for the model.

As a last exercise, we can use the calibrated learning model to generate a prediction for future female LFP. 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 76.7% of women would work in 2010.

### 4.4 The Roles of Earnings and Beliefs

To investigate the roles of earnings and 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

[38] Note that the calibration does not require both models to have the same values of $\sigma_l$ and $\beta$ (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.

[39] 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).

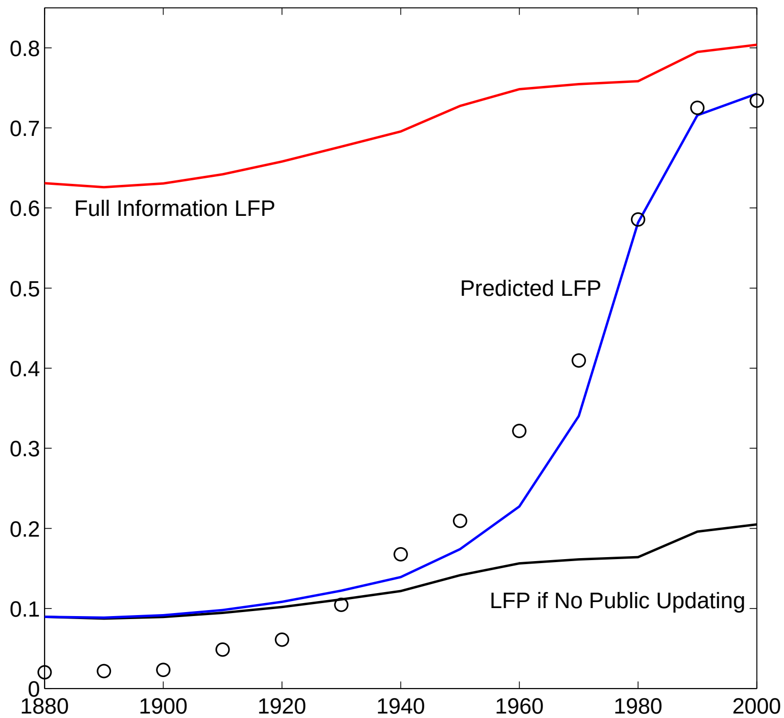


Figure 12: Uses solution parameters from calibrated model.

level (i.e., a prior of approximately 5% 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, in each period women receive a private signal and decide how much to work but there is no intergenerational evolution of beliefs. As show by the bottom line (with the caption "LFP if no public updating") in figure 12, female LFP would barely exceeded 20% by the year 2000. Alternatively, one can ask what female LFP would have been if, throughout the entire time period, agents had known the true value of $\beta$, i.e., $\beta^* = \beta_L$. This scenario is shown for the parameters of the calibrated model by the red line (with the caption "full information LFP"). It predicts a very different trajectory than the one we estimated, with LFP starting close to 65% in 1880 and slowly evolving to a bit over 80% by 2000. Thus, as can be seen from contemplating either of the two extremes regarding constant beliefs, the actual dynamics of beliefs is central to producing the path of female LFP shown in the figure 12. 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.[40]

Next, we can distinguish between the static effect of earnings on female LFP versus their dynamic effect via the induced change in beliefs by performing the following instructive ex-

[40] 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.

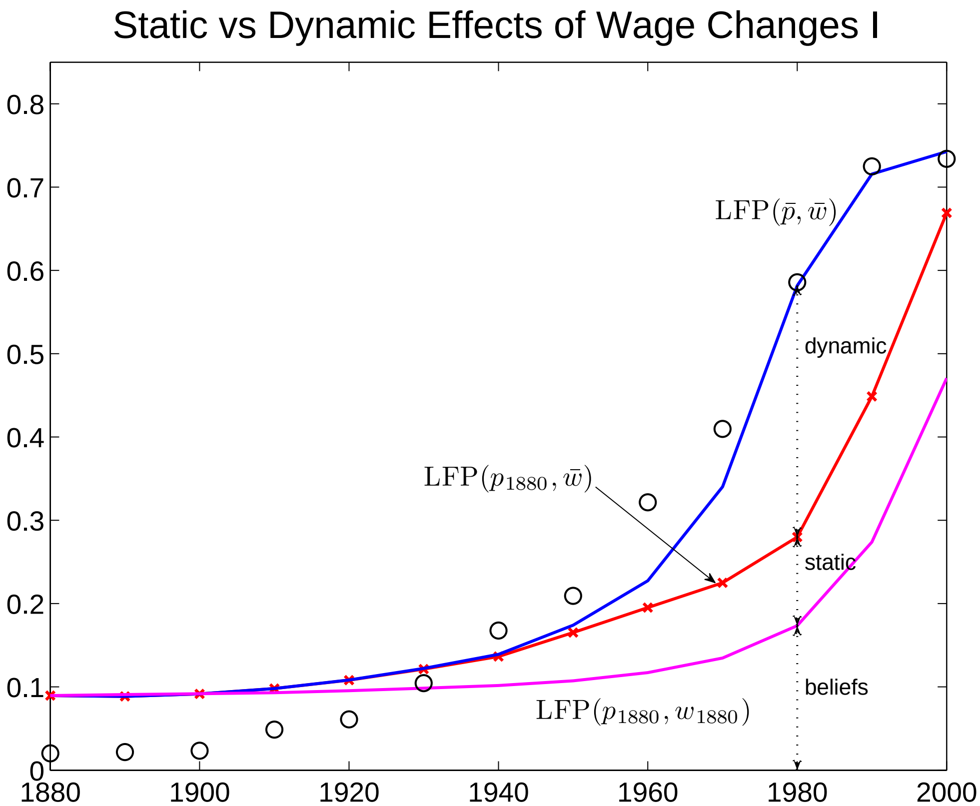


Figure 13: Decomposition of LFP. See text for notation.

ercise. First, as before, we can keep earnings frozen 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 13 below, is thus a result only of the changes in beliefs from constant wages. This measures the importance of beliefs in the LFP evolution for which changes in earnings play no part. Next, we can use the belief path obtained from this exercise and calculate, for each period, the increase in the proportion of women that would have worked for the actual historical earnings in every period. In this way, the 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. We denote the LFP curve obtained from this exercise $LFP\,(p_{1880}, \overline{w})$. Lastly, in figure 13 we also include the true LFP curve, i.e., the one obtained when both earnings and beliefs follow their historical paths (this is denoted $LFP\,(\overline{p}, \overline{w})$). Note that the difference between $LFP(p_{1880}, w_{1880})$ and $LFP\,(p_{1880}, \overline{w})$ is a measure of the static contribution of wages (as beliefs evolve the same way for both curves whereas earnings evolve only in the second curve); the difference between $LFP\,(\overline{p}, \overline{w})$ and $LFP\,(p_{1880}, \overline{w})$ on the other hand is a measure of the dynamic contribution of wages through changing beliefs (i.e., both series have the same historical earnings series but $LFP\,(\overline{p}, \overline{w})$ allows beliefs to respond to these changes and thus affect LFP).

As can be seen in figure 13, for the first eight decades, over 50% of female LFP is coming

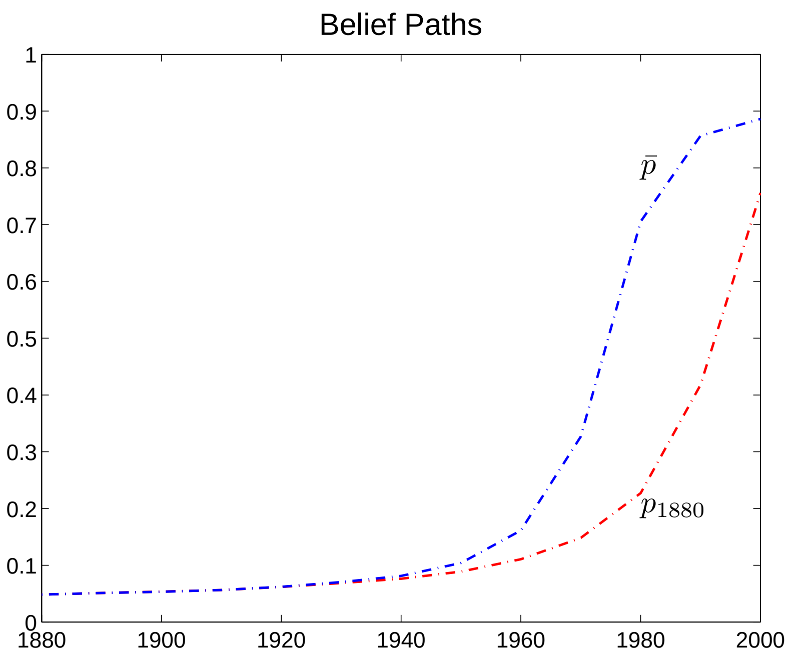


Figure 14: $P(\beta^* = \beta_L)$ for historical earnings series and for earnings constant at the 1880 levels.

from the evolution of beliefs over time independently of what is happening to wages. The dynamic and static effects of wages play an increasingly important role over time, however. In particular the dynamic effect becomes very important from 1970 to 1990, accounting for between a third to a half of female LFP. In 2000, however, the dynamic effect is under 10% and the static one is 27%.

One way to better visualize why the dynamic effect is more important in some decades than others, is to look at the two belief paths, as in done in figure 14 below. In that figure we plot the beliefs from the calibrated model (i.e., those given by $\overline{p}$) and the beliefs that would result if wages remained constant at their 1880 levels (i.e., those given by $p_{1880}$). Note that the differences in the probability assigned to $\beta^* = \beta_L$ is especially large in 1980 and 1990; this probability would have been 0.23 and 0.42 if earnings had not changed rather than 0.71 and 0.86 respectively. By 2000, however, the difference in probability assigned by the two belief paths is relatively much smaller, which explains the diminished importance of the dynamic effect of earnings on beliefs.

The way in which we decompose the wage effect into static and dynamic of course matters, though the basic story remains the same as above. We 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 but wages were kept fixed at their 1880 levels. This curve is shown in figure 15 as $LFP\,(\overline{p}, w_{1880})$. The effect of beliefs with unchanged earnings ($LFP\,(p_{1880}, w_{1880})$) remains as before, but the dynamic effect is now given by the difference between $LFP\,(\overline{p}, w_{1880})$ and $LFP\,(p_{1880}, w_{1880})$ as these have

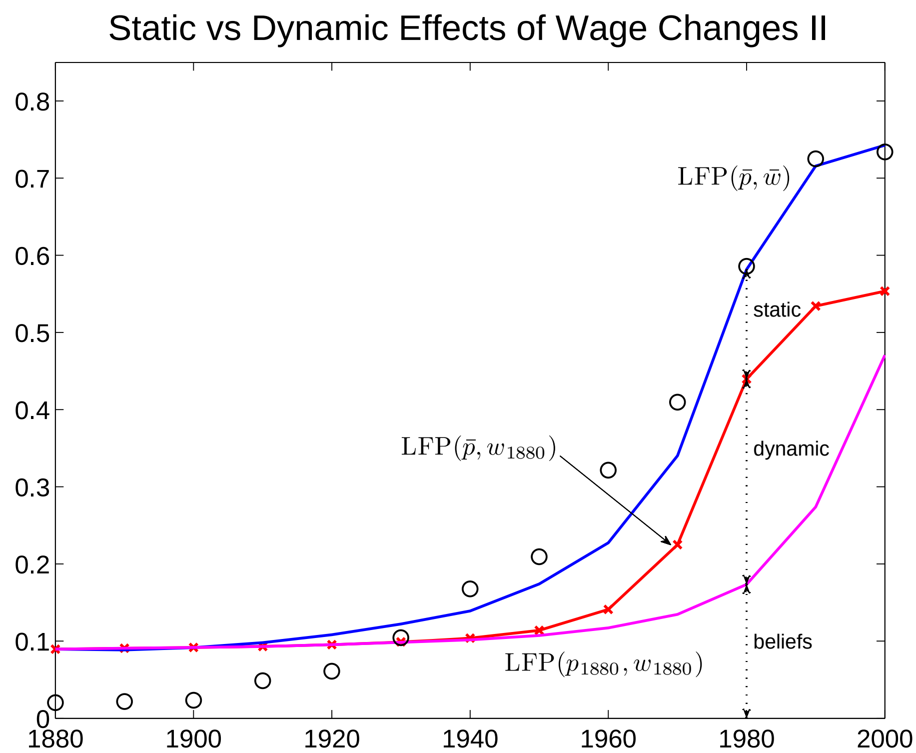


Figure 15: Alternative decomposition of LFP.

the same constant earnings but in the first case beliefs evolve as in the model with the historical earnings profile, whereas in $p_{1880}$ beliefs follow the path they would have taken had wages remained constant. For similar reasons, the static effect is the difference between $LFP\,(\overline{p}, w_{1880})$ and $LFP\,(\overline{p}, \overline{w})$, as now 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 27% to 45% of LFP in the decades 1970-1990.

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, the dynamic learning effect induced by the changes in earnings and the learning that would have gone on even in the absence of these changes, both contributed significantly to the quantitative evolution of female LFP.

# 5 Conclusion

This paper argues that in some contexts it may be useful to think about cultural change as a process of updating beliefs that occurs as part of a rational intergenerational learning process. In particular, it models the changes in married women's labor supply as the outcome of a Bayesian learning process in which women learn about the long-term implications of working for their marriage and for their children's welfare. We show that a simple model with these features, calibrated to key statistics from the later part of the 20th century, is capable of generating the aggregate timetrend of female labor force participation 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 differences in the proportion of women who would work under different states of the world is swamped by 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 female LFP statistics for the year 2000, namely LFP, and the own and cross-wage elasticities of LFP. In this model only changes in earnings 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 for basically every time period. Introducing learning in this simple model, and calibrating the new parameters to additional statistics, greatly improves the capacity of the model to predict the historical path of female LFP.

The model also indicates a novel role for increased women's wages (or for technological change), beyond the direct effect of making it more attractive for women to work outside the home. In particular, when beliefs are relatively pessimistic, increases in wages will make the signal required by the average woman less extreme and thus render the public signal more informative. Analysis of the calibrated model indicates that the dynamic effect of wages on beliefs played an important role in changing female LFP, particularly over the period 1970-1990. Learning over time, however, independently of any changes in wages, accounts for the bulk of the increase in LFP over the entire time period.

The model makes some heroic simplifications, including learning about an unchanged value of $\beta_L$ and $\beta_H$ over 120 years. It is difficult to believe, however, that only questions about individual payoffs were at stake during this time period, rather than also social reactions to women's roles at home and in the workforce. 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 did potentially changes in economic interests.[41] Thus, in addition to

[41] 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

exploring the informational role of social networks as in Fernández and Potamites (2007), in future work 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. Munshi and Myaux (2006) incorporate social rewards and punishments in the context of a learning model with multiple equilibria in which individuals are deciding whether to adopt modern contraception.[42] Lastly, exploring the potential inefficiencies that arise because individuals do not take into account the effect of their actions on learning and the role that policy may play would also be worthwhile.

Our paper has concentrated mainly on aggregate features of the data over a very long time horizon. It would also be of interest to examine sharper hypothesis about cultural change over a shorter time horizon that would allow a greater use of microdata.[43] Lastly, it would be important to explore whether variation in policies or technologies across space may allow us to empirically identify the dynamic effects of these on beliefs.

(e.g. voting) see, for example, Doepke and Tertilt (2007) and Edlund and Pande (2002).

[42] 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.

[43] 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 our 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.[44] 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 $\sigma_l$ 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 $\lambda_0$ ranging between [-10, -.01], $\sigma_\epsilon$ in [0.1, 5], $\sigma_\eta$ in [0.01, 2], $\sigma_l$ between [0.5, 4], $\beta_L$ in [.01, 1], and $\beta_H$ to be between [1, 10] units greater than $\beta_L$.

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 17. Then we average over the $\eta$ 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.

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

The elasticities were calculated computationally by assuming either a 1% increase in female wages or male wages 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 $\pm$ .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 integral that is needed to compute $\Pr(DW|MW)$, 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 $\sigma_l$ by minimizing the sum of the squared errors between predicted and actual LFP (12 obs).

## 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 (2007), "Gender Roles and Technological Progress," working Paper.

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

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

[5] 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.

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

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

[8] 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.

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

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

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

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

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

[14] 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.

[15] Doepke, Matthias and Michele Tertilt (2007), "Women's Liberation: What Was in It for Men?," working paper.

[16] 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.

[17] 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.

[18] 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.

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

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

[21] 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.

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

[23] Fogli, Alessandra and Laura Veldkamp (2007), “Nature or Nurture? Learning and Female Labor Force Participation,” Federal Reserve Bank of Minneapolis Staff Report 386.

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

[25] 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.

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

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

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

[29] 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.

[30] 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.

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

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

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

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

[35] Hill, Jennifer L., Jane Waldfogel, Jeanne Brooks-Gunn, and Wen-Jui Han (2005), "Maternal Employment and Child Development: A Fresh Look Using Newer Methods," *Developmental Psychology* 41(6), 833-850.

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

[37] 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.

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

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

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

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

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

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

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

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