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

Raquel Fernández
New York University, CEPR, NBER, IZA

Revised October 2008†

**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 (LFP) 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 child-care.[2]

An alternative explanation is that changes in culture or in social norms regarding women's role have had a significant impact on the evolution of women's market work. Although economists have tended to have a skeptical attitude towards cultural explanations, there is recent body of work showing that culture helps explain variation in how much women engage in market work across countries (see, for example, Fernández (2007b), Fernández and Fogli (forthcoming), Burda, Hamermesh, and Weil (2007), and Antecol (2000)).[3] There has not been any equivalent work, however, that shows that *changes* in culture can help explain the time series pattern of female market work. In part this is due to the fact that while it is easy to understand why cultures may differ across space (e.g., people have different priors), we do not have many theories that tell us why culture would change over time.[4]

There is, of course, a wealth of anecdotal and historical evidence showing that there has been a large shift in attitudes towards women's work. Poll questions provide useful quantitative evidence for these shifts, particularly if similar questions were asked over a

[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, non-farm occupations, non-group 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), Attanasio et al (2008), 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).

[4] Modeling culture as a selection mechanism among multiple equilibria, for example, 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. For an interesting example of evolutionary theory applied to culture see Bowles (1998). An interesting alternative is that social norms are passed on from parents to children in an optimizing fashion as in Bisin and Verdier (2000) and Tabellini (2007). This alternative does not tell us why new beliefs develop, however.

long time period. An example of this is shown in figure 2 that plots the evolution over time of the percentage of the sample of the US 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?"[5] In 1936 fewer than 20% of individuals sampled agreed with the statement; in 1998 fewer than 20% disagreed with it.[6]

Merely stating that society has changed the way in which it regards women, however, is not very enlightening on its own. It begs the question as to why culture changed and why these changes affected work behavior in such a gradual and uneven fashion as shown in figure 1. Indeed, one might be tempted to dismiss the opinions expressed in polls, novels, or personal anecdotes as mere reflections of technological advances that subsequently changed women's work behavior. Viewed from this perspective, as technology changed it altered the material costs and benefits of women's participation in the formal labor market and then people simply automatically voiced different opinions; changes in culture played no fundamental role. This paper will provide a different perspective.

The objective of this paper is to provide a theoretical framework in which culture and female labor participation interact and to show that this framework, in which changes in culture affect women's market work, may be quantitatively useful. Furthermore, this model will provide a different channel through which wages affect female labor supply and allow us to quantify the importance of this new role for wages over time. Taking inspiration from the fact that the path of female labor force participation follows an "S-shape" over time, the model explores the idea that cultural change may be the result of a rational, intergenerational *learning* process.[7]

S-shaped paths are commonly found in models of technological diffusion where the latter is either specified in an ad hoc fashion or through a learning mechanism. The S-shaped path of female LFP may thus 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.[8] If this is so, 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 (market) 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. A growing, more affluent middle class and a new ideal of womanhood gave rise to what has been called "the cult of domesticity." This process left the wife in charge of the domestic realm and her husband in charge of supporting the

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

[6] 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).

[7] 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 very different since they are interested in information cascades in which individuals stop learning.

[8] 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.

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.[9] 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 family and, consequently, most women exited the workplace upon marriage. 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.[10] Women were considered both physiologically and intellectually inferior to men. More recently, public anxiety regarding working women centers around the effect of a working mother on a child's intellectual achievements and emotional health [11] 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.[12] Using a framework broadly similar to Vives (1993) and Chamley (1999), I assume that women possess private information 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. 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.[13]

[9] For a fascinating account of this process see Woloch (1984) and Goldin (1990).

[10] Over 80% of married women, not employed in 1939 but who 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.

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

[12] Whether preferences or beliefs changed is often impossible to distinguish and, in a reduced-form setup, it is also unnecessary. In a model of cultural change, however, the assumption that changes in beliefs are driven by learning is important 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).

[13] Fogli and Veldkamp (2007) independently develops a related idea to explain the LFP of women with children from 1940-2000. They assume that women learn about the ability cost to a child from having a working mother from viewing a small sample of outcomes. Whereas in my model actions change because people modify their beliefs about the cost of working, in their model actions 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

The model has some very attractive features. First, the model *generically* generates an S-shaped figure for female labor force participation. Second, the model introduces a new role for changes in wages (or technological change). Unlike in traditional models in which changes in women's wages affect female LFP by changing the payoff from working, in this model they also affect the informativeness of the public signal and hence the degree of intergenerational updating of beliefs, i.e., changes in wages affect the pace of learning. Thus wage changes have dynamic effects on female LFP. Third, the calibrated model shows that this new role for wages was quantitatively important in increasing female labor supply, particularly in the decades of largest changes in female LFP – 1970-1990 – which Goldin has called the "revolutionary" phase.[14] Prior work on intergenerational changes in female labor supply during those decades had concluded that "changes in market work opportunities as manifested in wage changes are not enough to account for the profound inter-cohort changes that have taken place in the market work of women."[15] This model resurrects the importance of wage changes in explaining the dynamic path of female LFP but in a novel fashion.

To evaluate the model's ability to explain the quantitative evolution of female LFP, I calibrate it to a few key statistics for the last few decades of the sample (1980-2000). The model does a fairly good job of replicating dynamic path of female LFP from 1880 to 2000. I show that this is due to learning, i.e., to the endogenous evolution of beliefs; the same wage series with exogenously specified beliefs does a terrible job in reproducing the female LFP path.

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 the learning model and explains why the intergenerational evolution of beliefs naturally generates an S-shaped curve for female LFP. Section 3 calibrates the model and decomposes the changes in LFP into a beliefs component, a static wage component, and a dynamic wage-belief component. Section 4 discusses the roles of various assumptions and concludes.

## 2 A Simple Learning Model of Work

This section develops a simple model of a (married) woman's work decision that incorporates the two 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. The

about a family-specific component of infant mortality risk.

[14] See Goldin (2006).

[15] Pencavel (1998), p. 802.

disutility from working, in addition to reflecting a woman's known labor-leisure trade-off, is also assumed to capture her unknown long-run welfare consequences from working (i.e., the long-run consequences of working on her identity, marriage or her children, as discussed in the introduction). As these payoffs are revealed gradually over a long period of time, this uncertainty cannot be resolved by short-run experimentation. Thus, whether to work or not is modeled as a one-time decision.[16]

## 2.1 The Work Decision

A woman makes her work decision to maximize:

$$U(w_f, w_h, v_i) = \frac{c^{1-\gamma}}{1-\gamma} - \mathbf{1}(E_{it}v_i), \qquad \gamma \geq 0 \tag{1}$$

where $\mathbf{1}$ is an indicator function that takes the value one if she works and zero otherwise. A woman's consumption $c$ 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 given by:

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

where the first component of $v_i$ is a known idiosyncratic component that has a distribution in the population given by $G(l)$ which I assume is $N(0, \sigma_l^2)$. The second component is the individual realization of a random variable $B$. The value of this variable is revealed only if the woman works – it is the long-run disutility from working.

Women are assumed to be uncertain about the mean value of $B$, $\beta$. For simplicity, I assume that $\beta$ can take on only two values, high ($H$) and low ($L$), i.e., $\beta \in \{\beta_H, \beta_L\}$. Note that $\beta_L$ is the good state of nature in which working is on average not so costly, i.e., $\beta_H > \beta_L \geq 0$. Thus, a woman expected disutility from working is

$$E_{it}v_i = l_i + E_{it}(\beta_i) = l_i + E_{it}(\beta) \tag{4}$$

where $E$ is the expectations operator. That is, women know their own value of $l$ prior to making their work decision. They have (possibly different) expectations about the value of $\beta$ and only learn their individual realization of $B$, $\beta_i(\beta)$ once they work.[17]

Consider a woman in period $t$ who has a prior belief about the value of $\beta$ 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 obtains or inherits a private signal $s_{it}$ about the true value of $\beta$. This signal can be

[16] For simplicity, furthermore, we only consider the extensive margin, i.e., she either works or not, as this is the one that has seen the largest changes over time.

[17] Since utility is linear in $\beta_i$ and women obtain an *iid* draw from the distribution of $B$, only its mean matters.

thought of as arising from many sources (e.g., a draw from the scientific literature that exists at that moment regarding the welfare consequences of a woman working) or, in the preferred interpretation used in the quantitative section, it can be inherited directly from the woman's mother. The private signal is informative, i.e., it yields information about $\beta$. In particular, the value of the signal is given by:

$$s_{it} = \beta + \varepsilon_{it} \tag{5}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and its cumulative and probability distribution functions are denoted by $F(\cdot; \sigma_\varepsilon)$ and $f(\cdot; \sigma_\varepsilon)$, respectively.[18] The private signals are assumed to be i.i.d across women.

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

$$\begin{aligned}
\lambda_{it}(s) &= \lambda_t + \ln\left(\frac{Pr(s|\beta = \beta_L)}{Pr(s|\beta = \beta_H)}\right) \\
&= \lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma_\varepsilon^2}\right)\left(s - \bar{\beta}\right) \qquad (6)
\end{aligned}$$

where $\bar{\beta} = (\beta_L + \beta_H)/2$.[19].

It is worth noting a few properties of $\lambda_{it}(s)$:[20]

(*i*) $\frac{\partial \lambda_{it}(s)}{\partial s} < 0$;

(*ii*) $sign\left(\frac{\partial \lambda_{it}(s)}{\partial \sigma_\varepsilon^2}\right) = sign\left(s - \bar{\beta}\right)$;

(*iii*) $\lambda_{it}(s; \beta) \sim N\left(\lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma_\varepsilon^2}\right)\left(\beta - \bar{\beta}\right), \frac{(\beta_H - \beta_L)^2}{\sigma_\varepsilon^2}\right)$.

The first property follows from the fact that higher realizations of $s$ increase the likelihood that $\beta = \beta_H$. The second property implies that the updating of $\lambda$ is decreasing with the variance of the noise term, $\sigma_\in^2$, since a greater variance lowers the informativeness of the signal. The last property follows from the fact that $s \sim N\left(\beta, \sigma_\varepsilon^2\right)$.

Assume that women share a common prior in period $t$, $\lambda_t$.[21] What proportion of women will choose to work that period? From (1) it follows that a woman will work iff

$$\frac{1}{1-\gamma}[(w_{ht} + w_{ft})^{1-\gamma} - w_{ht}^{1-\gamma}] - E_{it}(\beta) \geq l_i \tag{7}$$

that is, her net expected benefit from working must exceed her 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 $W\left(w_{ht}, w_{ft}\right)$.

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

[19] To obtain (6) one uses the fact that $Pr(s|\beta)$ is equal to the probability of observing a signal $s$ generated by a normal distribution $N(\beta, \sigma_\epsilon^2)$.

[20] Each property follows directly from taking the appropriate derivative of equation 6.

[21] 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 $l$ $(l \leq \underline{l}(w_{ht}, w_{ft}))$ will work and women with very high $l$ $(l \geq \bar{l}(w_{ht}, w_{ft}))$ will not work, where

$$\begin{aligned} \underline{l}(w_{ht}, w_{ft}) &\equiv W(w_{ht}, w_{ft}) - \beta_H \qquad (8) \\ \bar{l}(w_{ht}, w_{ft}) &\equiv W(w_{ht}, w_{ft}) - \beta_L \qquad (9) \end{aligned}$$

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 $\lambda$, 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.,

$$p_j^* \beta_L + (1 - p_j^*)\beta_H = W(w_{ht}, w_{ft}) - l_j \qquad (10)$$

Using (8), we obtain $p_j^*(w_{ht}, w_{ft}) = \frac{l_j - \underline{l}(w_{ht}, w_{ft})}{\beta_H - \beta_L}$ and hence,

$$\ln \frac{p_j^*}{1 - p_j^*} = \ln \frac{l_j - \underline{l}}{\bar{l} - l_j} \qquad (11)$$

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

$$\lambda_t(s_j^*) = \lambda_t - \left( \frac{\beta_H - \beta_L}{\sigma_\varepsilon^2} \right) (s_j^* - \bar{\beta}) = \ln \left( \frac{l_j - \underline{l}}{\bar{l} - l_j} \right) \qquad (12)$$

and hence

$$s_j^*(\lambda_t; w_{ht}, w_{ft}) = \bar{\beta} + \left( \frac{\sigma_\varepsilon^2}{\beta_H - \beta_L} \right) \left( \lambda_t + \ln \left( \frac{\bar{l}(w_{ht}, w_{ft}) - l_j}{l_j - \underline{l}(w_{ht}, w_{ft})} \right) \right) \equiv s_j^* (\lambda_t) \qquad (13)$$

We can conclude from the derivation above that the proportion of women of type $l_j$, $\underline{l} < l_j < \bar{l}$, that will choose to work in time $t$ given a prior of $\lambda_t$ and given the true value of $\beta$, $L_{jt}(\beta; \lambda_t)$, is the proportion of them that receive a signal lower than $s_j^* (\lambda_t)$, i.e.,

$$L_{jt}(\beta; \lambda_t) = F(s_j^*(\lambda_t) - \beta; \sigma_\varepsilon) \qquad (14)$$

Thus, integrating over all types, the total fraction of women who will work in period $t$ is given by:

$$L_t(\beta; \lambda_t) = G(\underline{l}) + \int_{\underline{l}}^{\bar{l}} F(s_j^*(\lambda_t) - \beta; \sigma_\varepsilon) g(l_j) dl \qquad (15)$$

where $g(\cdot)$ is the pdf of the $l$ distribution $G(\cdot)$. Note that $L_t$ can take on only two values: one if $\beta = \beta_L$ and another if $\beta = \beta_H$.

Before specifying how beliefs are transmitted across generations, it is worth noting a few features of $s_j^* (\lambda)$.[22]

[22] Each property follows directly from taking the appropriate derivative of equation 13.

$(i) \qquad \frac{\partial s_j^*}{\partial \lambda} < 0;$

$(ii) \qquad \frac{\partial s_j^*}{\partial w_f} > 0; \ \frac{\partial s_j^*}{\partial w_h} < 0.$

The first property states that if, ceteris paribus, women hold more optimistic priors, they are willing to work at higher values of $s$. The second property is that increases in own earnings makes women more willing to work whereas the opposite holds for increases in husband's earnings. This same property yields the next result

**Proposition 1** $\frac{\partial L_t}{\partial w_f} > 0$ *and* $\frac{\partial L_t}{\partial w_h} < 0.$

**Proof.** The proof follows directly from property (ii) of $s_j^*$ above. ■

Thus the model yields the traditional (desirable) comparative statics results with respect to wages.

## 2.2 Intergenerational Transmission

The model thus far is purely static. To incorporate dynamics we need to specify how the state variable (beliefs) changes over time. To do this, we need to be explicit as to what information is passed on from generation $t$ to generation $t+1$.

I assume that each woman transmits her prior belief to her child, i.e., a child inherits $\lambda_{it}(s)$. An equivalent assumption is that generation $t+1$ inherits the prior of generation $t$ (its "culture"), $\lambda_t$, which each individual then updates with the private signal, $s$, inherited from her mother. If solely this information were transmitted intergenerationally, then $\lambda_{it}(s) = \lambda_{it+1}(s)$. In that case, only changes in wages could lead to changes in aggregate work behavior. There is, however, an additional source of information available to women in $t+1$ that was unavailable 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 $t$, $L_t$, they would be able to back out $\beta$ as a result of the law of large numbers (i.e., using equation (15)). While assuming that information about how many women worked in the past is totally unavailable seems extreme,[23] the notion that this knowledge is completely informative seems equally implausible and is merely an artifact of the simplicity of the model. In particular, with greater sources of heterogeneity in the model, backing out the true value of $\beta$ would require agents to know the geographic distribution of male earnings and female (potential) earnings and how they were correlated within marriages, the distribution of preferences, the geographic distribution of shocks to technology and preferences, etc. I employ, therefore, the conventional tactic in this literature and assume that women observe a noisy function of the aggregate proportion of women worked.[24]

[23] This is the assumption in Fogli and Veldkamp (2007).

[24] This is the strategy used in finance, for example, by introducing noise traders. 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)) which is imperfectly correlated with that of other social circles. This yields similar results. It has the advantage, for the calibration, of

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

$$y_t\left(\beta;\lambda_t\right)=L_t\left(\beta;\lambda_t\right)+\eta_t \tag{16}$$

and where $\eta_t \sim N(0,\sigma_\eta^2)$ with a pdf denoted by $h(\cdot;\sigma_\eta)$.

Alternatively, one could assume that individuals perfectly observe LFP, but are uncertain about the distribution of a parameter that affects the idiosyncratic disutility of work. The realization of that parameter (e.g., the mean of the distribution of $l$ which is now set at zero) would change randomly every period (for example, by depending on an unobservable aggregate factor in the economy).[25] One very simple formulation of this alternative is to assume that a proportion $\omega > 0$ of the population is subject to extreme shocks so that a random proportion $z$ of them work and a proportion $1-z$ don't work, independently of wages. Thus the proportion of women who work $\widetilde{L}_t\left(\beta;\lambda,z\right)$ would be given by:

$$\widetilde{L}_t\left(\beta;\lambda,z\right)=\left(1-\omega\right)L_t\left(\beta;\lambda\right)+z\omega$$

Assuming that $z$ is distributed normally would then yield an expression equivalent to (16).

The assumption that $\eta$ (or $z$ in the example above) is distributed normally should be taken as an approximation made for analytical simplicity.[26] One can make alternative assumptions about this distribution (e.g., an appropriately truncated normal pdf) but this renders the analytical expressions and computations considerably more cumbersome.

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

$$\begin{aligned}\lambda_{t+1}(\lambda_t,y_t) &= \lambda_t+\ln\frac{h(y_t|\beta^*=\beta_L)}{h(y_t|\beta^*=\beta_H)} \\ &= \lambda_t+\left(\frac{L_t(\beta_L;\lambda_t)-L_t(\beta_H;\lambda_t)}{\sigma_\eta^2}\right)\left(y_t-\overline{L}_t(\lambda_t)\right) \end{aligned} \tag{17}$$

where $\overline{L}_t(\lambda_t)=\frac{L_t(\beta_L;\lambda_t)+L_t(\beta_H;\lambda_t)}{2}$. Note that (17) is the law of motion of aggregate beliefs which we can think of as "culture."

Figure 3 summarizes the time line for the economy. Individuals start period $t$ with a common (updated) prior, $\lambda_t$. Each woman updates the common prior with her inherited private signal and makes her work decision, generating an aggregate $L_t$ and a noisy public signal $y_t$. Generation $t+1$ observes $y_t$ and uses it to update the old common prior ($\lambda_t$),

---

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.

[25] See Chamley (1999).

[26] It must be taken as an approximation since otherwise it implies that some observations of $y_t$ could be negative and some greater than one.

generating $\lambda_{t+1}$ – the "culture" of generation $t+1$.[27] It should be noted that instead of assuming that women in $t+1$ inherit $\lambda_t$ (or $\lambda_{it}$) which they update with the information contained in $y_t$, we can equivalently assume that women always have the same common prior of $\lambda_0$, they observe the entire history of $y_\tau$, $\tau = 0, 1, 2..., t$ and they update accordingly. This would yield the same value of $\lambda_{t+1}$ (or of $\lambda_{it+1}$).

## 2.3 Some Properties of the Learning Model

In additional to generating qualitatively similar comparative statics as in traditional labor supply models without learning, 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.[28] 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, first note that given $\beta = \beta_L$, we can rewrite (17) as

$$\lambda_{t+1} = \lambda_t + \left( \frac{L_t(\beta_L; \lambda_t) - L_t(\beta_H; \lambda_t)}{\sigma_\eta^2} \right) \left( \eta_t + \frac{L_t(\beta_L; \lambda_t) - L_t(\beta_H; \lambda_t)}{2} \right) \tag{18}$$

Taking the expected value of (18), it is easy to see that $E_t \lambda_{t+1} - \lambda_t$ is increasing in the difference between the aggregate proportion of women who work when $\beta = \beta_L$ relative to when $\beta = \beta_H$, i.e., in $L_t(\beta_L; \lambda_t) - L_t(\beta_H; \lambda_t)$. A large change in the LLR, ceteris paribus, implies that the proportion of women who change their work decisions from one generation to the next will also be large. Thus, there will be significant changes in behavior over time when $L_t(\beta_L; \lambda_t) - L_t(\beta_H; \lambda_t)$ is large and the opposite when it is small.

To understand why an $S$-curve will be generated requires understanding, therefore, how the size of $L_t(\beta_L; \lambda_t) - L_t(\beta_H; \lambda_t)$ varies over time. We start by noting that for a given $l_j \in \left(\underline{l}, \bar{l}\right)$

$$L_{jt}(\beta_L; \lambda_t) - L_{jt}(\beta_H; \lambda_t) = F(s_j^*(\lambda_t) - \beta_L; \sigma_\varepsilon) - F(s_j^*(\lambda_t) - \beta_H; \sigma_\varepsilon). \tag{19}$$

where $L_{jt}$ is the proportion of women who work of type $l_j$. Taking the derivative of (19) with respect to $s_j^*$ yields the f.o.c.

$$f\left(s_j^* - \beta_L\right) - f\left(s_j^* - \beta_H\right) = 0 \tag{20}$$

Recalling that $f\left(s_j^* - \beta\right) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp - \left\{ \left( \frac{(s_j^* - \beta)^2}{2\sigma_\varepsilon^2} \right) \right\}$, this implies that (19) is minimized at $s_j^* = \pm\infty$ and it is at its maximum at $s_j^* = \overline{\beta}$.

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

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

Thus, if the critical signal $s_j^*$ is far from $\overline{\beta}$ in absolute value, (19) will be small, the updating of beliefs will be small, and the change in work behavior next period will also be small.[29] Why? When women require an extreme signal in order to be willing to work, the difference in the proportion of woman who work across the two states $L, H$ is small. This renders the aggregate signal $y_t(\beta; \lambda)$ less informative as its variance across the two possible states will be swamped by the variance of the aggregate noise term $\eta_t$. Thus, the amount of intergenerational updating of beliefs will be slight and hence the change in the proportion of women who work that period, ceteris paribus, will likewise be small. The opposite is true when $s_j^*$ is close to $\overline{\beta}$.

The conclusion above follows from the general features of the normal distribution of $\varepsilon$, which is depicted in figure 4.[30] As can be seen in the figure, when $s^* - \overline{\beta}$ is far from zero, the difference in proportion of women who work in the two states is small, i.e., the difference between $L_j$ 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 value of $\beta$. The opposite is true at $s^{*\prime}$. As shown in the figure, when $s^{*\prime} - \overline{\beta}$ is close to zero (i.e., the point midway between $s^{*\prime} - \beta_L$ and $s^{*\prime} - \beta_H$), the difference between $L_j'$ at the two states of nature is large. Hence beliefs will change substantially from one generation to the next, leading to large changes in behavior.

Note that a similar conclusion holds once we aggregate over the $l_j$ types. Taking the derivative of $L(\beta_L; \lambda_t) - L(\beta_H; \lambda_t)$ with respect to $\lambda$ (so as to change all $s_j^*$) and using (13), we obtain:

$$\frac{\partial}{\partial \lambda_t}\left(L(\beta_L; \lambda_t) - L(\beta_H; \lambda_t)\right) = \left(\frac{\sigma_\varepsilon^2}{\beta_H - \beta_L}\right) \int_{\underline{l}}^{\overline{l}} \left[ f\left(s_j^*(\lambda_t) - \beta_L\right) - f\left(s_j^*(\lambda_t) - \beta_H\right)\right] g(l_j) dl_j \tag{21}$$

Thus, if the critical signal $s_j^*(\lambda_t)$ is, for the average individual in $(\underline{l}, \overline{l})$, far from $\overline{\beta}$, (21) will be small in absolute value, intergenerational updating will be small, and the evolution of LFP over time will be slow.[31] The opposite is true when the critical signal is close to $\overline{\beta}$ for the average individual.

The S-shape follows from the logic above. If parameter values are such that initially women, on average, require extreme (negative) signals in order to be willing to work, then learning will be slow for several generations (as little information will be revealed by the aggregate signal). During this time, there will be little change in the aggregate proportion of women who work. This is the first (slowly rising) portion of the S curve for LFP. Once beliefs have become more moderate and a large proportion of women require moderate values of $s$ to change their work behavior, learning accelerates and the change in work

[29] Note that $s^*$ can be far from $\overline{\beta}$ for a variety of reasons (e.g., because $\lambda$ is low or because $w_f$ is low or $w_h$ is high).

[30] As will be clear from the intuition that follows, a normal distribution of the noise term $\epsilon$ is not critical. Rather, the distribution needs to be able to give rise to a cdf that increases slowly at the beginning, rapidly towards the middle, and then slowly once again towards the end.

[31] The assumption of heterogeneous types complicates matters since one must also be concerned about the size of $g(\underline{l})$. Thus, in order for difference in $L$ across the two states of nature to be large, we need $s_j^*$ to be close to $\overline{\beta}$ for types with a large frequency not only in $(\underline{l}, \overline{l})$ but overall.

behavior over time is large (the aggregate signal is very informative). This is the second (rapidly rising) portion of the S curve. Lastly, once beliefs are optimistic and women, on average, require extreme (positive) signals in order to choose not to work, learning once again slows down since the aggregate signal becomes relatively uninformative. The change in work behavior over time is correspondingly small. This is the third and last (slowly rising) portion of the S curve. As the time horizon goes off to infinity, beliefs converge to the truth, so any further changes in female LFP result solely from changes in wages.

### 2.4 Wages and Learning

The learning model generates a novel role for wages. As in a model without learning, an increase in women's wages increases female LFP that period. Learning, however, introduces an additional dynamic effect. In particular, wage changes affect the pace of intergenerational learning, i.e., the magnitude of $\lambda_{t+1} - \lambda_t$. Note that this effect is not because having a greater proportion of working women increases the information about the welfare consequences of working.[32] Rather, the channel is through the effect of wage increases on $s^*$. As explained in the preceding section, if women require on average an extreme private signal to be willing to work (i.e., $s^*$ is very negative for the average individual) anything that serves to increase $s^*$ will also increase the difference across states in the proportion of women who work and hence increase the informativeness of the aggregate signal for the next generation. This mechanism is present for increases in female wages, decreases in male 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)). The opposite would be true if women were initially very optimistic and hence required, on average, very high values of $s^*$ in order not to work. In that case, wage increases would have the effect of decreasing the amount learnt by the next generation. Thus, in general, wage changes have a dynamic externality in this model that is not present in more traditional settings. The model yields a very different perspective on how one should evaluate the effects of changes in wages, technology, and policy and one of the main objectives of the next section will be to ask whether this effect could be quantitatively important in explaining the historical evolution of female LFP.

## 3 Quantitative 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. It should be noted from the outset that the empirical analysis is not a "test" of the model as it does not attempt to quantify the contributions of other potentially important factors discussed in the introduction except insofar as they are reflected in wages (e.g., as would be the case for many forms of technological change or changes in wage discrimination).

[32] It would be easy though to incorporate this additional channel.

The contribution of this quantitative 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. This is a novel application of a learning model to explain how changes in beliefs may diffuse very slowly over time and it is the first model to theoretically and empirically explore this new role for wages. It is therefore useful to first understand the model's theoretical and quantitative implications in a simple setting before developing a more complicated quantitative dynamic model. Furthermore, it should be noted that many 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.[33]

## 3.1 Calibration Strategy

In the model, married women decide whether to engage in market work taking their husbands' earnings and their own potential wages as given. Thus, calibrating the model 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 use the (median) earnings of full-time white men and women for which some data was available as of 1890.[34] This choice exaggerates the earnings of working women in general, as some work less than full time (although part time work is not quantitatively important until after 1940). 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.[35] 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.[36] As is commonly done,

[33] To give 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. The literature tends to ignore the effect of beliefs (culture) on the demand for household technological innovation.

[34] I restrict the sample to white women as black women have had a different LFP trajectory with much higher participation rates earlier on.

[35] 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.

[36] 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.

observations that report weekly earnings less than 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.[37]

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.[38]

The evolution of female LFP from 1880 to 2000 is shown in figure 1. These percentages are calculated from the US Census 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). I calibrate the model to match female LFP in 1980-2000. The remainder of the LFP series is generated endogenously by the model.

In addition to matching female LFP over the last three decades, I also require the model to match the own and cross-wage elasticity in 2000, the cross-wage elasticity in 1990 and the relative probability of a woman working in 1980 (conditional on whether her mother worked). See table 1 for a list of the targets. 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.[39] 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.[40]

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

[37] 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.

[38] 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.

[39] 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).

[40] Using the elasticities estimated from a specification 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).

1983.[41] 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. This procedure was done for the subsample consisting of all white married women between the ages 25-45 who were born in the U.S.[42] In the calibration each period is a decade and, for the purpose of computing the work risk ratio, daughters are assumed to make their work decisions two periods after their mothers (i.e., a separation of 20 years).

### 3.2 Calibration Results

Before turning to the full calibration of the model, it is instructive to first calibrate a simpler version of it in which $\beta$ is known so that there is no learning. Thus, only changes in wages (male and female) can explain why labor supply changes over time. The unknown parameters are now three (instead of seven) – $\gamma, \beta$, and $\sigma_l$ – allowing us to match a restricted set of targets. In particular, I choose to match 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 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, $\beta$, and about how dispersed the $l$ types must be in order to generate a given response to a change in wages.

The simplicity of the model allows one to solve for the parameter values analytically. In particular, as shown in the Appendix, the ratio of the two elasticities can be manipulated to yield $\gamma = 0.503$ and the use of an elasticity and the value of LFP yields $\sigma_l = 2.29$ and $\beta = 0.321$. To interpret the magnitude of the common expected disutility of working, 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 model with no learning does a terrible job of matching the female LFP data. The data is shown in small circles and the topmost line (with the label $\alpha = 1$) is the model's predicted LFP. The model grossly overpredicts female LFP in all decades other than 1990 and, by construction, the calibrated target in 2000.

[41] I use 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 the lenient work requirement definition) than what would be consistent with the Census numbers.

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

This basic inability of the no learning 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. The results obtained from recalibrating the model using values of $\alpha$ that vary from 0.1 to 1 is shown in figure 6. As is clear from the figure, this modification 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.[43]

The failure of the model without learning is also robust to the exact choice of earnings series. For example, 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 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 path.

We now turn to calibrating the full model (i.e., $\beta$ is unknown and women endogenously learn its value over time). As female LFP has been increasing throughout and, from the results of simplified no-learning model, changes in wages alone cannot replicate this phenomenon, I assume that the true state of nature is $\beta = \beta_L$. 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 needs to be addressed—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 $\eta_t$ 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. 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 $L_{t+1}$ are not.

I deal with the aggregate shock in the following way. For each period $t+1$, given LFP

[43] 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.

in the previous period $L_t$, I calculate the proportion of women who would work, $L_{t+1}$, for each possible realization of the shock, $\eta_t$, i.e., for each induced belief $\lambda_{t+1}(\eta)$. Integrating over the shocks, I find the expected value of LFP for that period, $E_t L_{t+1}(\lambda_{t+1}(\eta))$, 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 $\lambda^*_{t+1}(\eta^*)$ such that:[44]

$$\int_\eta L_{t+1}(\lambda_{t+1}(\eta);\lambda_t)\, h(\eta) d\eta = L_{t+1}\left(\lambda^*_{t+1}(\eta^*)\right) \tag{22}$$

Performing this exercise in each period determines the path of beliefs.[45]

To continue with the calibration exercise, as shown in the Appendix, manipulating the ratio of elasticities in this model yields the same value of $\gamma$ as without learning, i.e., $\gamma = 0.503$. As in the simplified model, the additional elasticities targets and values of female LFP yield 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, $\sigma_l$, but also by the dispersion of private information, $\sigma_\varepsilon$. Furthermore, as the expected value of $\beta$ is evolving over time with the beliefs $\lambda$, the values of LFP from 1980-2000 yield information as well on how rapidly $\lambda$ needs to evolve over these decades and hence on how noisy the signal $\eta$ should be (i.e., on $\sigma_\eta$).

Lastly, as mothers and daughters share the same private information, the conditional probability that a woman works as a function of her mother's work behavior (the work risk ratio, $R$) also yields information on the evolution of $\lambda$ and how different the values of $\beta_H$ and $\beta_L$ should be. I take a mother and daughter to be separated by twenty years. The work risk ratio is thus given by

$$R_t = \frac{\Pr(DW_t|MW_{t-2})}{\Pr(DW_t|MNW_{t-2})} \tag{23}$$

and in the benchmark I assume that daughters inherit perfectly their mother's private signal whereas their $l_j$ type is a random draw from the normal distribution $G(\cdot)$ that is *iid* across generations.[46] The details of the calculation are shown in the Appendix.

Table 1 below shows the calibration targets (column 1) and the values obtained in the calibrated learning model (column 3). The second column reports the values obtained in

[44] For the computation, I take a large number of draws of entire histories for $\eta$ (500 histories) in order to calculate the expected value of $L$. See the Appendix for details.

[45] 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 $y_{t,k} = L_t + \eta_{t,k}$ where $\eta$ is an iid draw from the normal distribution $N(0, \sigma^2_\eta)$. Given a common prior, $\lambda_t$ (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 (22), i.e., $\int_{\eta_k} L_{t+1}\left(\lambda_{t+1,k}\left(\eta_{t,k}\right)\right) = L_{t+1}\left(\lambda^*_{t+1}(\eta^*)\right)$. To maintain the common prior assumption, one needs 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 $\lambda^*_{t+1}(\eta^*)$.

[46] 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.

the prior calibration exercise in the model with no learning.

Table 1

| *Calibration Targets* | | Model with known $\beta$ | Learning Model |
|---|---|---|---|
| Own-Wage Elasticity (2000) | 0.30 | **0.30** | **0.29** |
| Cross-Wage Elasticity (2000) | -0.13 | **-0.13** | **-0.13** |
| Female LFP (2000) | 0.734 | **0.734** | **0.744** |
| Female LFP (1990) | 0.725 | 0.725 | **0.716** |
| Cross-Wage Elasticity (1990) | -0.14 | -0.13 | **-0.14** |
| Female LFP (1980) | 0.586 | 0.687 | **0.585** |
| Work Risk Ratio (1980) | 1.13 | 1 | **1.13** |
| *Parameters* | | | |
| $\gamma$ | | 0.503 | 0.503 |
| $\sigma_L$ | | 2.293 | 2.085 |
| $\beta$ | | 0.321 | |
| $\beta_H$ | | | 4.935 |
| $\beta_L$ | | | 0.001 |
| $P_0(\beta = \beta_L)$ | | | 0.086 |
| $\sigma_\varepsilon$ | | | 5.288 |
| $\sigma_\eta$ | | | 0.055 |

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

The LFP predictions from the calibrated model are shown in figure 7. 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 $\beta_L$ that is held by the median woman.

As can be seen from the figure, the calibrated model on the whole does a good job of replicating the historical path of female LFP.[47] It under-predicts LFP from 1930 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. The median individual assigns around a 9% probability to the event $\beta = \beta_L$. Beliefs evolve very slowly over the first seventy years (remaining below 20% for the median individual during this period). Then, as of 1960, the change in beliefs accelerates. The median individual jumps from assigning a probability of 26.0 to $\beta_L$ in 1960, to 48.3% in 1970, to 83.8% in 1980. By 2000, the median probability assigned to $\beta = \beta_L$ is 94.7%.

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

Individual beliefs are very dispersed as the private signal has a large variance. Figure 8 shows the equilibrium path of beliefs once again, for the individual with the median signal as well as for the individuals with private signals two standard deviations below and above this mean.[48]

The fact that the model's LFP predictions are too low in the period 1930-1970 may indicate that another factor, such as technological change in the household or less employment discrimination on the part of firms, 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 would have had repercussions in later decades through the dynamic impact of technological change on learning discussed earlier. World War II may also have played a role by making women more willing to work during that war years and this, in turn, increased the pace of intergenerational learning.

It is instructive to ask why the learning model yields such a different time path for female LFP than the model with no learning. As noted previously (see also Table 1), the calibration implies that both models must have the same value of $\gamma$. Furthermore, the difference in the standard deviation of the normal distribution of types is relatively small: 2.29 versus 2.09. Lastly, the expected value of $\beta$ in 2000 is not very different from the known value of $\beta$ when learning is eliminated: 0.32 as opposed to the 0.26 assigned by the median individual in the learning model.[49] 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 when learning is eliminated, in the learning model the median expected value of $\beta$ is close to 4.51 in 1880 and then evolves over time to 0.26 in 2000 This allows LFP to respond in dramatically different ways over time.

It is also 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)). The learning model's elasticities predictions are shown in figure 9 (given by the highest and lowest lines). 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.[50] 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

[48] Using (6), note that the median individual has a LLR given by $\lambda_t + \frac{(\beta_L - \beta_H)^2}{2\sigma_\epsilon^2}$.

[49] 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.

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

woman to work (recall that we find that indeed $\beta_L$ 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.[51]

As seen in the figure, the model is able to generate large changes in wage elasticities. This is mostly due to the fact that beliefs are changing over time, thus changing women's response to wages. Eliminating learning, as in the earlier calibration, gives rise to much smaller time variation in these paths, as can be seen by the middle two lines in figure 9.

As a last exercise, one can use the calibrated learning model to generate a prediction for future female LFP and 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), the 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. I 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.

### 3.3 The Quantitative Contributions of Wages and Beliefs

To investigate the quantitative contributions 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 level (i.e., at the median individual's prior of approximately 9% 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 10, female LFP would barely exceeded 10% 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 top (red) line (with the caption "full information LFP"). It too predicts a very different trajectory, 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 10). 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.[52]

Next, we can distinguish between the static and dynamic effects of wage changes on female LFP by performing the following instructive decomposition. First, as before, we

[51] 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.

[52] This is simply a repetition, with slightly different parameter values, of the finding that, without learning, the model does a very bad job of replicating the LFP trajectory.

can keep wages 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 11, 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 11. Hence, the difference between the level of LFP in 1880 (given by the dotted horizontal line) and $LFP(p_{1880}, w_{1880})$ measures the 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 wages 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 $LFP\left(p_{1880}, \overline{w}\right)$ and it is shown with (red) x's in the figure. The difference between $LFP(p_{1880}, w_{1880})$ and $LFP\left(p_{1880}, \overline{w}\right)$ 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 $LFP\left(p_{1880}, \overline{w}\right)$).

Lastly, we allow wages to also influence intergenerational learning and thus beliefs and denote the LFP path obtained this way $LFP\left(\overline{p}, \overline{w}\right)$). Note that this LFP path is the one predicted by the model and depicted previously in figure 7. It is the top (blue) curve shown in figure 11. The difference between $LFP\left(\overline{p}, \overline{w}\right)$ and $LFP\left(p_{1880}, \overline{w}\right)$ measures the dynamic contribution of wages through its effect on beliefs (i.e., both series have the same historical earnings series, $\overline{w}$, but $LFP\left(\overline{p}, \overline{w}\right)$ allows beliefs to respond to these changes and thus affect LFP whereas $LFP\left(p_{1880}, \overline{w}\right)$ keeps the belief path that would have occurred had wages remained at their 1880 level).

As can be seen in figure 11, 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 the decades of largest 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 12. Note that the difference in the probability assigned to $\beta = \beta_L$ is especially large in 1980 and 1990; these probabilities (as held by the median woman) would have been 31.5% and 49.4% for these two decades if wages had remained constant rather than 83.8% and 92.9% 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.

It should be noted that the decomposition of LFP is not unique. One could alternatively eliminate the $LFP\left(p_{1880}, \overline{w}\right)$ curve and replace it with the LFP path that would result if beliefs followed the path obtained from the historical earnings series, $\overline{p}$, but wages were kept constant at their 1880 levels. This curve is shown in figure 13 as $LFP\left(\overline{p}, w_{1880}\right)$. The effect on LFP of beliefs with unchanged earnings ($LFP\left(p_{1880}, w_{1880}\right)$) remains as before, but the dynamic effect of wages is now given by the difference between $LFP\left(\overline{p}, w_{1880}\right)$ and $LFP\left(p_{1880}, w_{1880}\right)$. 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 effect of earnings is now measured as the difference between $LFP\left(\overline{p}, w_{1880}\right)$ and $LFP\left(\overline{p}, \overline{w}\right)$, 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.

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, changes in beliefs, both those that would have occurred even had wages remained constant and those induced by changes in wages, played the largest quantitative role in generating the changes in female LFP over the last 120 years.

# 4 Discussion and Conclusion

This paper modeled the dynamics of married women's labor force participation as reflecting a process of cultural change (i.e., changes in societal beliefs) brought about by intergenerational learning. In this process, married women compared the benefits of increased consumption from labor earnings with the expected utility cost of working. This cost was unknown and women's beliefs about it evolved 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, from their mothers, of private information. I showed that a simple model with these features, calibrated to key statistics from the later part of the 20th century, generates a time trend of female labor force participation that is roughly 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 women are, on average, pessimistic about the welfare value of participating in the labor market, 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 beliefs about the welfare consequences of work become more moderate, the information in the signal improves. Once beliefs are sufficiently optimistic 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 calibrated a version of the model without any evolution of beliefs to a few key statistics for the year 2000. In this model, only changes in earnings over time can explain changes in female LFP. I showed 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 from the last few decades of the century greatly improves the model's capacity to replicate the historical path of female LFP. Furthermore, it generates a pattern of own and cross-wage elasticities that varies greatly over time in ways that have been discussed in the literature.

The model also indicates a novel role for increases in women's wages (or for policies or 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 leaving it constant and thus not introducing additional parameters seemed a better choice. 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. Introducing risk aversion (with respect to the uncertainty about the long-run payoff from working) is straightforward and would create an additional reinforcing channel for learning.

The calibrated model finds that at the outset women were pessimistic about the true cost of working. This lack of neutrality may indicate that particular social forces were at play in determining culture initially. 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 might require a political economy framework to explain why certain opinions become dominant.[53]

In future work it would be interesting to investigate quantitatively both the informational role of different social networks and the contribution of social rewards and punishments to changing behavior over time relative to social learning.[54] Some interesting 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 decide whether to adopt modern contraception.[55] At a theoretical level, it would also be interesting to explore further 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.[56] 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 changes on beliefs. Variation across states in the importance of WWII shocks may permit some progress in this direction.

[53]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).

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

[55]In their model, an individual's payoff from using birth control depends on her type (whether she is a "reformer" or not) and the contraceptive choice of a randomly chosen woman with whom she interacts.

[56]See, e.g. Munshi and Myaux (2006). 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.

# 5 Appendix

## 5.1 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.[57] 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).

## 5.2 Calibration of the learning model

### 5.2.1 The model with no learning

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

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

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

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

Next one can use one of the elasticity expressions and the requirement that $G\left(l^*; \sigma_l\right) = L$

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

in 2000 to solve for $\beta$ and $\sigma_l$. Note that since $G$ is a normal distribution, one can write:

$$l^* = \sigma_l \Phi^{-1}(L)$$

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

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

where $A = \frac{w_f(w_f+w_h)^{-\gamma}}{\sqrt{2\pi}\varepsilon_f L}$. One can then solve for $\beta$ directly from the definition of $l^*$, yielding $\beta = 0.321$.

### 5.2.2 The model with learning

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

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

Noting further that $\frac{\partial \underline{l}}{\partial w_k} = \frac{\partial l^*}{\partial w_k}$, this implies that by performing the same manipulations as in the previous subsection one obtains (25), and thus the same value of $\gamma$ as in the earnings only model, i.e., $\gamma = 0.503$.

In order to calculate a daughter's conditional probability of working (as a function of her mother's work behavior), one needs to specify, in addition to how private signals are inherited, how mothers and daughters are correlated in their $l_j$ types. As a benchmark, I assume that the correlation is zero, i.e., the $l_j$ type is a random draw from the normal distribution $G(\cdot)$ that is *iid* across generations.[58] Signals, on the other hand, are perfectly inherited. 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 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 (13), we obtain

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

And, using Bayes rule and $\beta^* = \beta_L$, we can calculate the probability that a daughter works

[58] 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.

given that her mother worked as:

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

where $DW$ and $MW$ stand for daughter works and mother worked, respectively. I 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, $\Pr(DW_t \text{ and } MW_{t-2}|s)$, 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

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

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

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

In order to estimate $\lambda_0, \sigma_\varepsilon, \sigma_\eta, \beta_H, \beta_L$, and $\sigma_l$ I 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], $\sigma_\varepsilon$ 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 15. 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.

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 $\beta_L - 4\sigma_\varepsilon$ to $\beta_L + 4\sigma_\varepsilon$ were used.

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

[7] Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos (2008), "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review*, 98:4, 1517–1552.

[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 (forthcoming), “Culture: An Empirical Investigation of Beliefs, Work, and Fertility,” *American Economic Journal: Macroeconomics*.

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

[41] 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.

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

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

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

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

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

[47] Pencavel, John (1998), "The Market Work Behavior and Wages of Women: 1975-94," *The Journal of Human Resources*, 33(4), 771-804.

[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.

[53] Wolloch, Nancy (1984), *Women and the American Experience*, New York: Alfred A. Knopf.

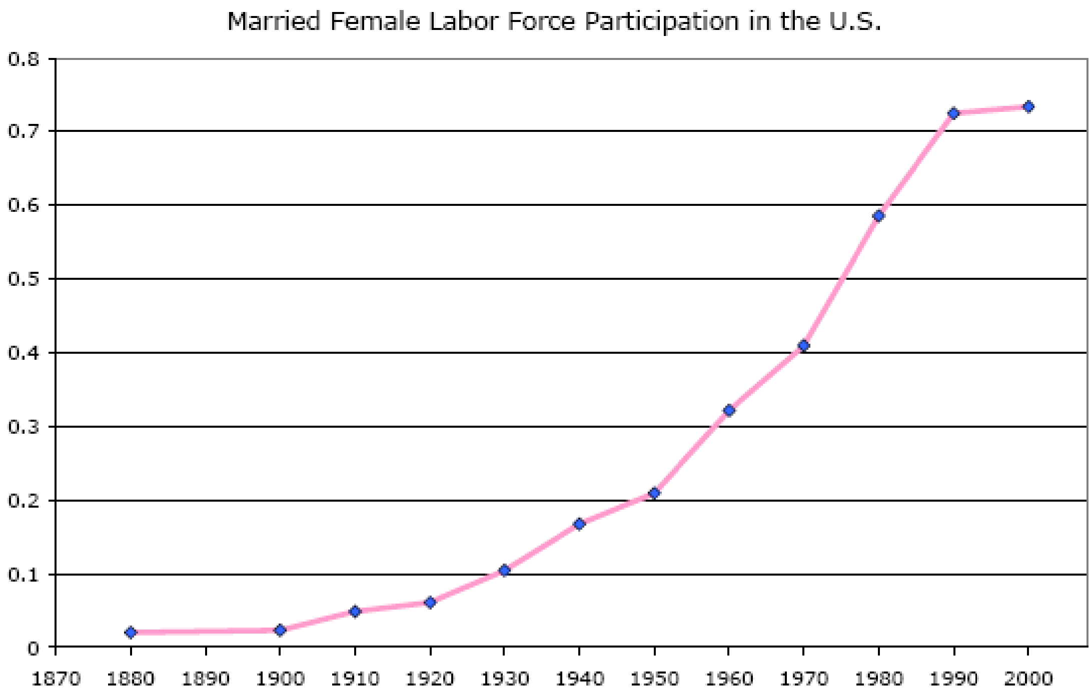


Figure 1: U.S. Census data 1880-2000. Percentage of white, married (spouse present) women born in the U.S., 25-44 years old (non-agricultural, non-group quarters), who report being in the labor force.

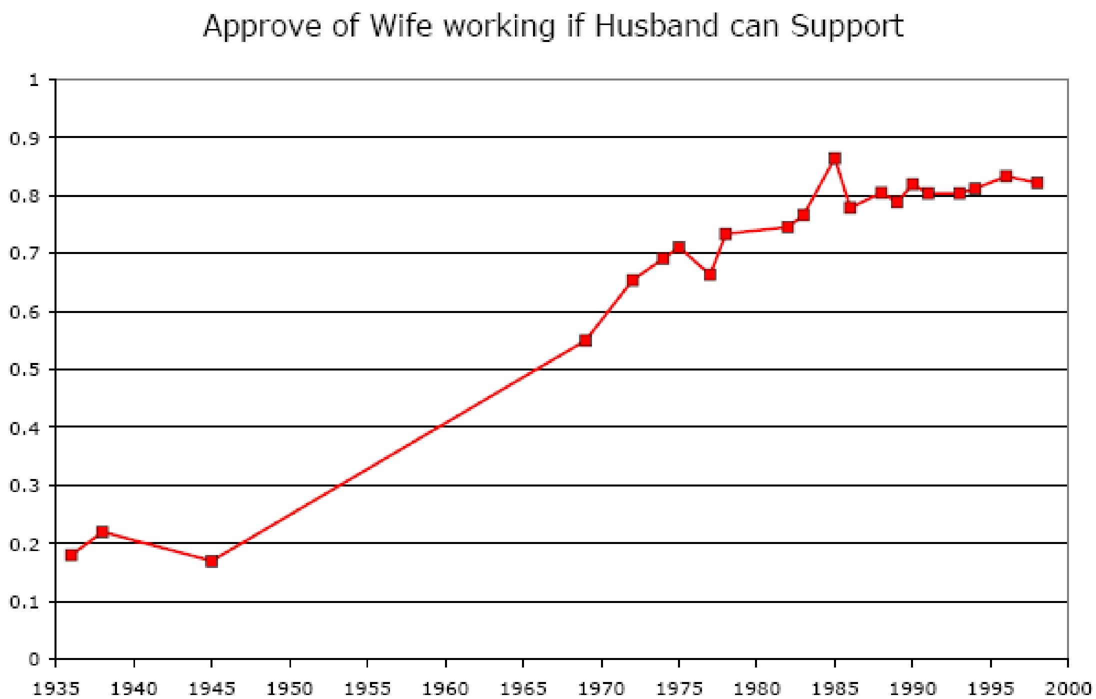


Figure 2: Sources: 1936-1938 and 1969 numbers are from the Gallup Poll (1972), 1945 is from Benjamin I. Page and Robert Y. Shapiro, The Rational Public, University of Chicago Press, 1992; pp. 101, 403-4. 1972 onwards are from the General Social Survey.

Figure 3: Timeline of Learning Model

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

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

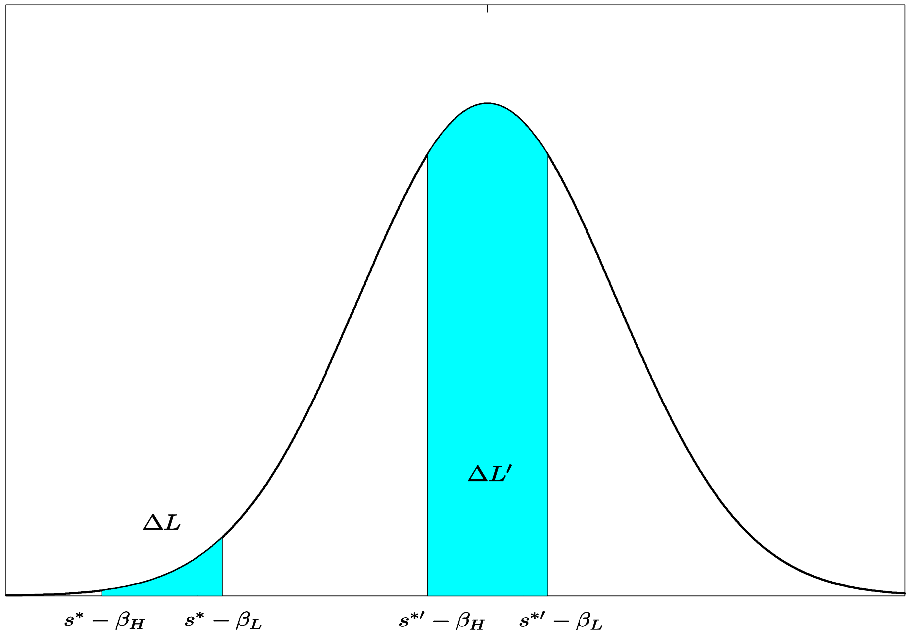


Figure 4: Normal PDF

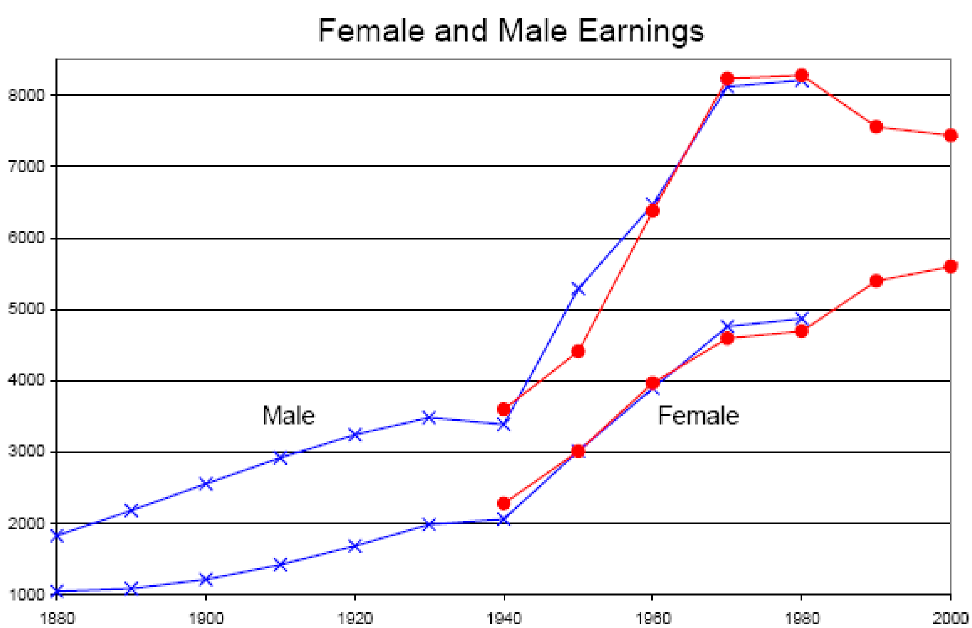


Figure 5: Crosses (blue) represent the yearly median earnings data from Goldin (1990), Table 5.1. Dots represent our calculations using U.S. Census data (red). They are the median earnings of white men and women between the ages of 25-44 in non-farm occupations and not living in group quarters. All earnings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

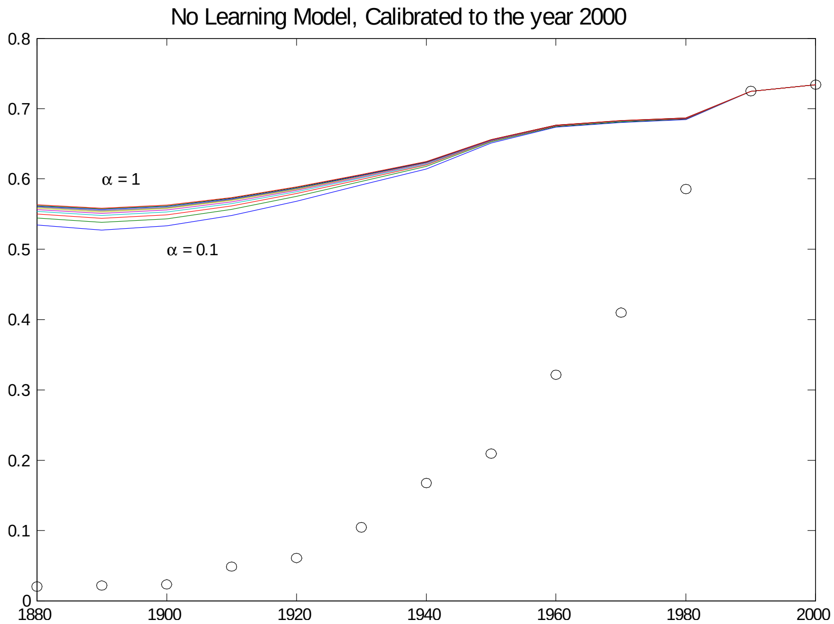


Figure 6: Parameters: $\gamma = 0.503$, $\beta = 0.321$, and $\sigma_L = 2.293$. $\alpha$ is the fraction of husband's earnings that enters a wife's utility via consumption.

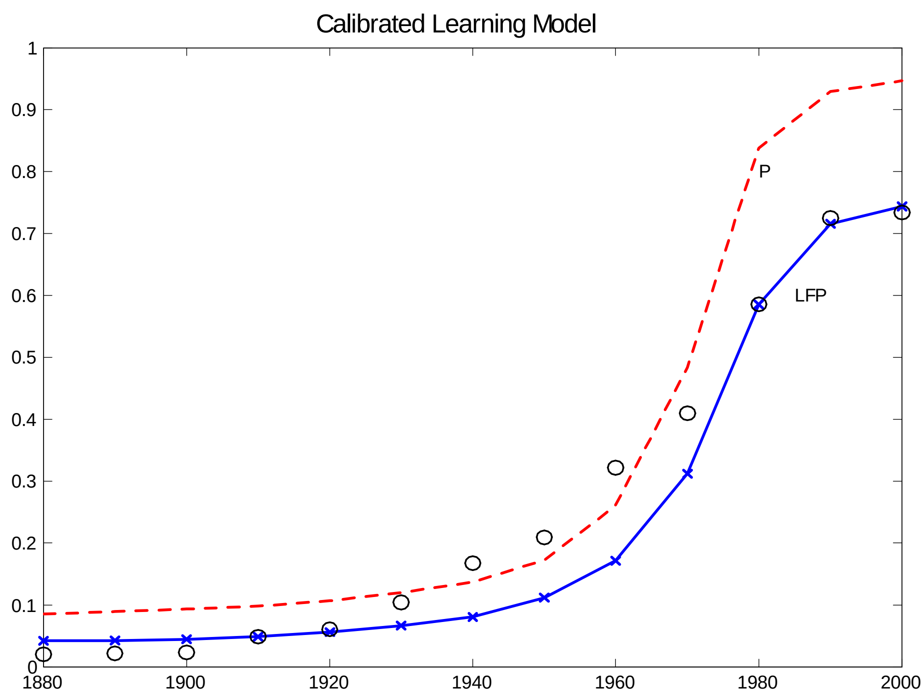


Figure 7: The dashed red line (p) is the belief path of the median individual. The sum of squared errors (distance of predicted LFP from actual LFP) is 0.052.

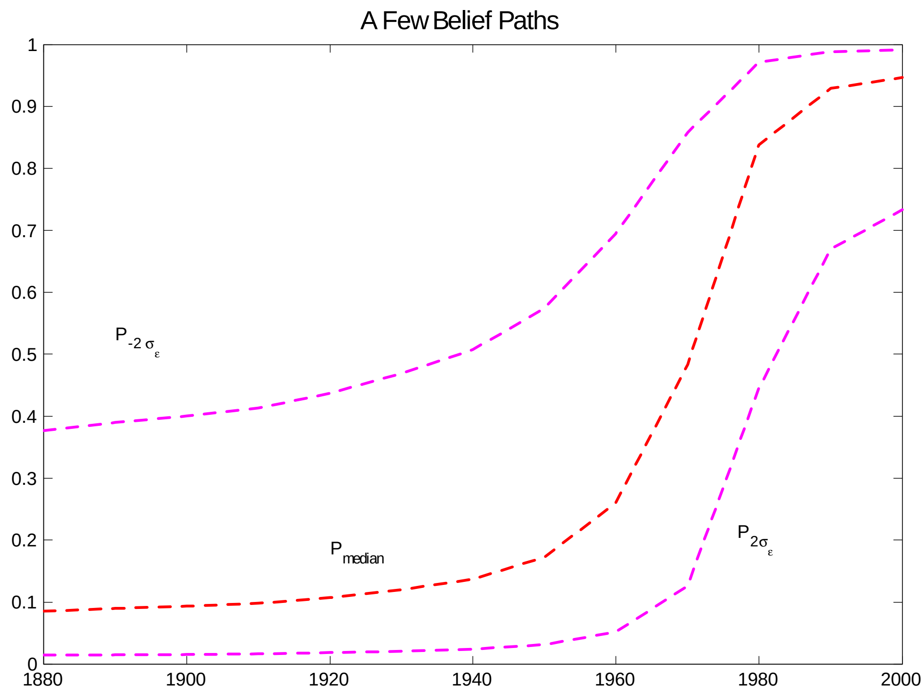


Figure 8: This shows $\Pr(\beta = \beta_L)$ for agents with $s = \beta$ and $s = \beta \pm 2\sigma_\varepsilon$.

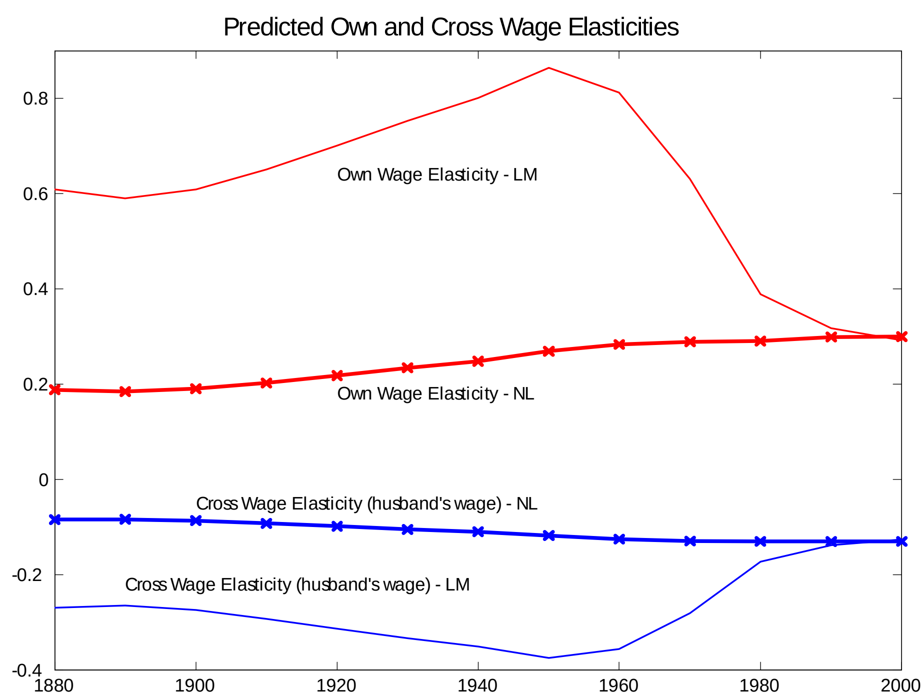


Figure 9: Parameter values from calibrated model. See the Appendix for a description of how the elasticities were calculated. LM = Learning Model, NL = No Learning Model.

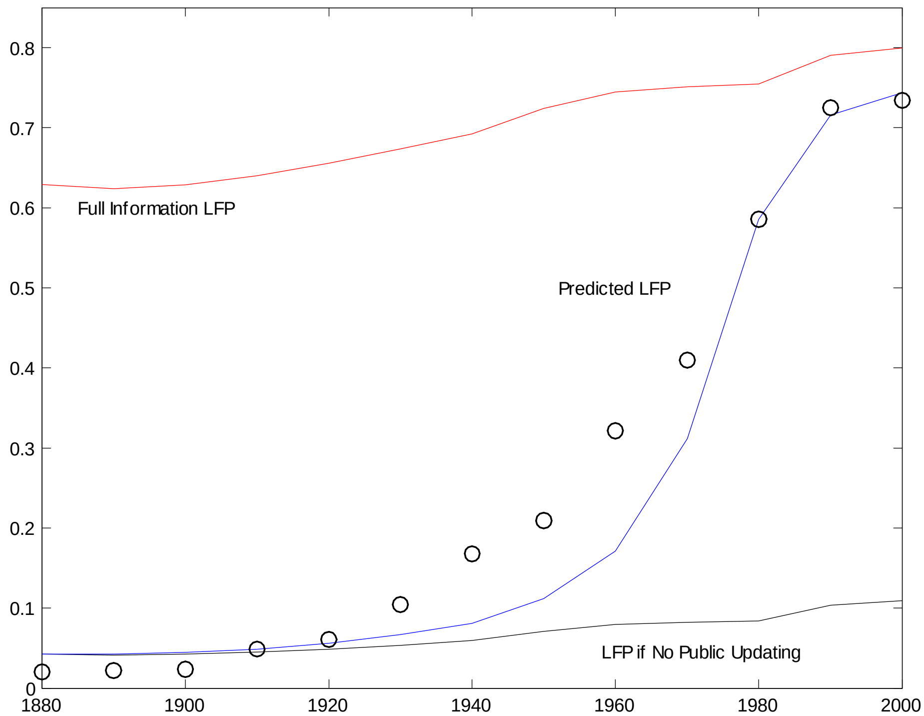


Figure 10: Uses the solution parameters from calibrated model but without public learning.

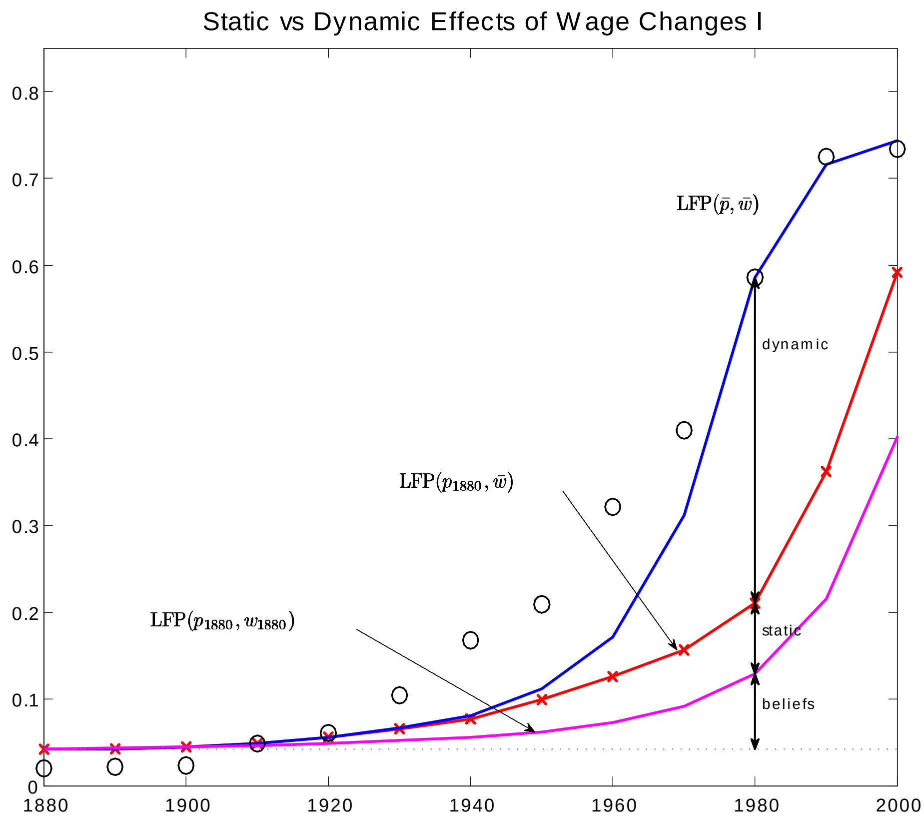


Figure 11: Decomposition of LFP. See the text for notation.

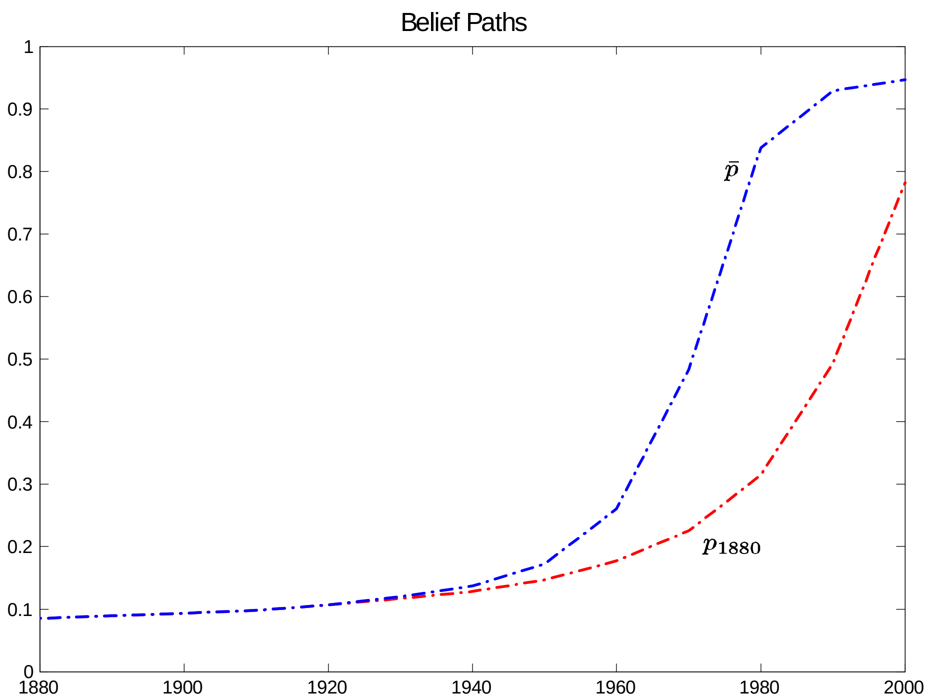


Figure 12: $P(\beta = \beta_L)$ for historical earnings series and for earnings constant at the 1880 levels.

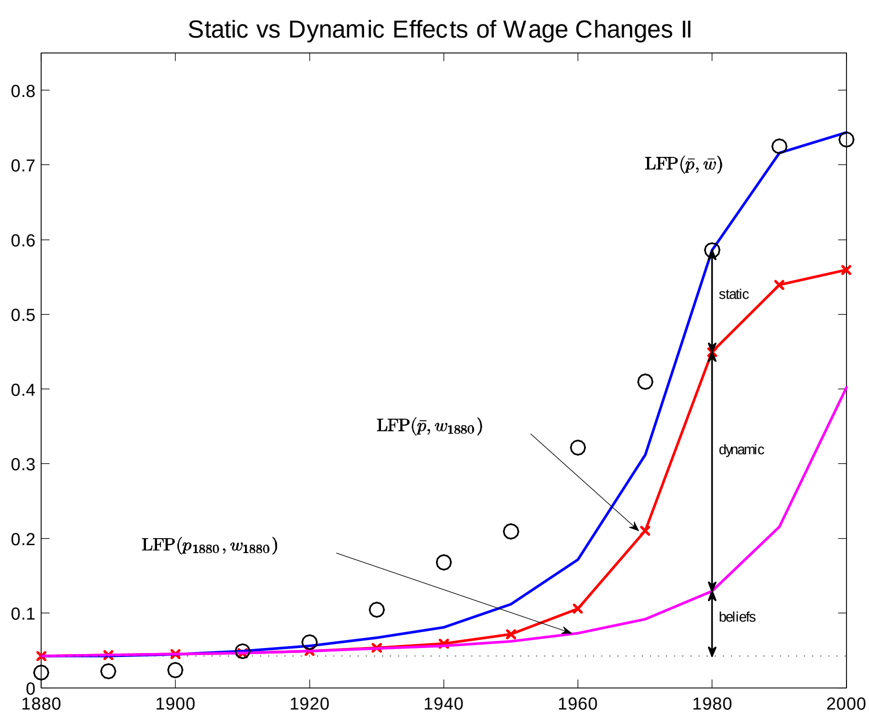


Figure 13: Alternative decomposition of LFP.