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

Raquel Fernández
New York University

Revised November 2010†

Abstract

This paper investigates the role of changes in culture in generating the dramatic increase in married women's labor force participation over the last century. 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 endogenously 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. 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. I examine the model's cross-sectional and dynamic testable implications and compare them with the data. 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).

*Other affiliations: CEPR, NBER, IZA, and ESOP. I thank Liz Potamites and Joyce Cheng Wong 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 in the US. The pace of change, however, has been markedly uneven. As shown in figure 1, married women's labor force participation increased very slowly from 1880 to 1920, averaging 1.0 percentage point per decade. It grew somewhat more rapidly between 1920 and 1950 (on average 4.9 percentage points per decade), and then took off between 1950 and 1990, increasing on average 12.9 percentage points per decade. Since then, it has stayed relatively constant.[1]

Many explanations have been proposed for this transformation. Depending on the particular time period under consideration, potential causal factors have included structural change in the economy, technological change in the workplace and in the household, medical advances, decreases in discrimination, institutional changes in divorce law, and the greater availability of childcare.[2]

A striking fact that none of these theories have addressed is the accompanying revolution in social attitudes towards married women working. This social transformation can be seen everywhere, from changes in laws governing women's work (e.g. the "marriage bar" which made it difficult for schools to hire married women or to keep women as teachers once they married) to the depiction of married women in literature and the popular press (e.g. Sarah Hale, editor of Godey's from 1836-1877, lectured her readers about "the importance of sticking to home and hearth" so that they would be able to "influence, and ennoble, the entire world" and the popular novelist Grace Greenwood wrote that "true feminine genius is ever timid, doubtful and clingingly dependent, a perpetual childhood").[3]

Quantitative evidence for the dramatic changes in social attitudes is provided by polls. Figure 2 plots the evolution over time of the percentage of the sample that answered affirmatively to the question "Do you approve of a married woman earning money in business or industry if she has a husband capable of supporting her?"[4] As shown by the circles (green), in 1936 fewer than 20% of individuals sampled agreed with the statement; in 1998 fewer than 20% of individuals disagreed with it.

To understand why social attitudes (i.e., culture) and labor force participation moved in tandem requires a framework that is able to address both phenomena. The objective of this paper is to provide such a framework and to examine whether the suggested mechanism

[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, in non-agricultural occupations and living in non-farm, non-institutional quarters.

[2] The classic source for an economic history of female labor force participation is Goldin (1990). For various explanations for this change see, among others, Goldin (1990), Galor and Weil (1996), Costa (2000), Goldin and Katz (2002), Jones, Manuelli, and McGrattan (2003), Greenwood, Seshadri, and Yorukoglu (2005), Gayle and Golan (2010), Albanesi and Olivetti (2009a, 2009b), and Knowles (2007).

[3] One way to follow this transformation is by examining the evolution of perceptions of women's role by the influential Godey's Lady's Book, a popular periodical. The quotes are taken from Collins (2003), p. 86-87.

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

may be quantitatively significant. Taking inspiration from the fact that the path of female labor force participation follows an "S-shape" over time, suggesting a process of information diffusion, this paper develops a model in which cultural change is the result of a rational, intergenerational *learning* process in which individuals are endogenously learning about the long-run payoff to married women working.[5] In this process, female labor force participation (LFP) and culture are co-determined, giving rise to an S-shaped process of aggregate labor supply and social attitudes.

In this paper I develop a simple model of a married woman's work decision. Using a framework broadly similar to Vives (1993) and Chamley (1999), I assume that women possess private information about the long-run costs of working (e.g., about the severity of the consequences of working for a woman's marriage, her 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 women's decisions in the preceding generation and they in turn make their own work decision. Thus, beliefs evolve endogenously via a process of intergenerational learning.

The model has several attractive features. First, the model *generically* generates an S-shaped figure for female labor force participation and for social beliefs. Second, the model introduces a new role for changes in wages (or technological change). Unlike traditional models in which changes in women's wages affect female LFP solely 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. Thus, wages affect the pace of learning and consequently have dynamic effects on female LFP. Third, the model generates a path for culture, i.e., for the evolution of social beliefs.

To evaluate whether the proposed learning mechanism has the potential to be quantitatively significant, I calibrate the model to a few key statistics for the last decades of the sample (1980-2000). I find that the calibrated model does a fairly good job of replicating the dynamic path of married women's LFP from 1880 to 2000. 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 -1970, both the dynamic and static effects of wage changes played a role in increasing female LFP, and from 1970 -1990 in what Goldin (2006) has called the "revolutionary" phase, the dynamic effect on beliefs of changes in earnings is critical in accounting for the large increase in the proportion of working married women. Thus, the model resurrects the importance of wage changes in explaining the dynamic path of female LFP with a novel mechanism.[6] A welfare

[5] A curve that rises slowly at first, then rapidly, and then flattens is called an "s-shaped" diffusion shape even if it doesn't, strictly speaking, look like an S. These curves are common in the technology and epidemiology literature. See Geroski (2000) for a review of this literature.

[6] It should be noted that Jones, Manuelli, and McGrattan (2003) show that changes in the gender gap can

analysis of the cost of imperfect information indicates, not surprisingly, that the latter was high.

A unique feature of the model is that it endogenously generates a path for culture or beliefs. A comparison of the predictions of the calibrated model with the poll data shows that the model does a good job in generating the basic features of the data. I also derive other testable implications of the model regarding elasticities, cross-sectional predictions across women of different education levels, and the dynamics of LFP and beliefs. In particular, cross-state variation in the importance of World War II as a shock that generated exogenous variation in female labor supply can be used to study whether the intergenerational implications of the model are consistent with the data.

The paper is organized as follows. Section 2 reviews recent literature on the role of culture and the evolution of married women's LFP. Section 3 presents the learning model and Section 4 derives the main results. Section 5 presents the calibration of the model and decomposes the changes in LFP into a beliefs component, a static wage component, and a dynamic wage-belief component to assess their respective contributions to the evolution of LFP. Section 6 examines other testable implications of the model and conducts a welfare analysis. Section 7 discusses the roles of various assumptions and concludes.

2 The Literature on Culture and Married Women's LFP

The next two sections briefly review some of the pertinent literature on culture and married women's labor supply and discuss the reasons for using a learning model to think about cultural change.

2.1 Culture and Married Women's LFP

There is a by now a large literature on married women's labor supply.[7] This literature has attempted to explain the change in married women's labor supply primarily with changes in standard economic variables such as women's education, earnings and the gender gap, fertility, and marriage/divorce prospects. It has sometimes included factors that are more difficult to measure such as changes in the cost of childbearing, culture or social norms, or technological change in household production. Overall, this literature has concluded, often by default, that this last category of difficult-to-measure variables plays an important quantitative role in accounting for the large changes seen in married women's LFP.

For example, a very interesting recent paper by Eckstein and Lifshitz (2009) develops a quantitative dynamic stochastic model that embeds all the variables described above. In their discrete-choice labor supply model, they take the initial state (i.e., schooling, marital status, employment, wage, fertility, husband's employment and wages) of a woman at age 22 as given and have the woman decide in each period until the age of 65 whether

account for much of the changes in married women's labor supply from 1950 to 1990 in a model with home production and leisure. The new channel for wages that I explore here does not require a home-production sector and in this way the papers can be seen as complementary.

[7] This literature is too large to be reviewed here. See, e.g., the survey by Blundell and MaCurdy (2000).

to work. They assume that wages follow standard Mincer functions over time whereas the remaining variables evolve following a simple state-dependent stochastic process. The model is estimated using March CPS (1963-2007) data for the 1955 cohort and then the estimated parameters are used to assess the quantitative importance of different variables in the labor supply of cohorts born every 5 years between 1925 and 1975. The authors find that schooling, wages, and the "other explanations" category have significant explanatory power in accounting for the differences in labor supply across cohorts, contributing respectively 35%, 20%, and 38% to the observed differences. Moreover, they find that by changing only two parameters which capture the utility cost of working for all women and the cost of working for women with young children allows them to obtain a good fit for the "unexplained" portion of labor supply of women from all cohorts.

Lee and Wolpin (2010) use a similar approach to investigate the importance of demand versus supply side factors in the evolution of the labor market over the last three decades. They also find that changes in preferences for leisure are quantitatively important in accounting for changes in the employment ratio of women to men. Heathcote, Storesletten, and Violante (2010) calibrate a dynamic general equilibrium model to investigate the macroeconomic and welfare consequences of rising wage inequality in the US. One important exercise they do is to feed the historical (1967-2005) wage series into their parametrized model to investigate the relative importance of the changing skill premium, wage gender gap, and the changes in the variances of the permanent and transitory components of the wage residual (i.e., that part unexplained by education, demographics, etc.) in explaining some of the cross-sectional and time-series dimensions of the data. They find that the changing wage gender gap is an important driver for the increase in ratio of female to male hours, but they too conclude that factors such as changes in culture or technology are responsible for the sizeable unexplained differences between the model predictions and the data on women's relative labor supply over this period.[8]

The papers discussed above conclude that simple changes in preferences (culture) can help explain a large share of the changes in married women's LFP over the last few decades but do not provide direct evidence on the role of culture. There is another strand of literature that provides micro-based evidence on the role of culture and married women's labor supply using the "epidemiological" approach of studying culture.[9] This approach attempts to separate the influences of culture from those of institutions and traditional economic variables by studying the descendants of immigrants in a given country. The basic argument is that whereas immigrants tend to transmit the preferences and beliefs of their country of origin to their offspring, the new economic and institutional environment in which these descendants make choices will be the same across the different immigrant groups (controlling for geographic variation within the country). One can then examine

[8] There exist, of course, many older papers that concluded that changes in preferences or difficult to measure changes in technology must play an important role in explaining the evolution of female LFP. See, e.g., Smith and Ward (1985) or Pencavel (1998). A recent paper that does an excellent job quantifying the contribution of medical and other technological progress is Albanesi and Olivetti (2009b).

[9] See the forthcoming chapter in the Handbook of Social Economics (Fernández (2010)) for a review of the literature using this approach.

whether variables that reflect the beliefs in the country-of-ancestry are able to account for a significant portion of the variance in economic outcomes across the second-generation groups.[10]

For example, Fernández and Fogli (2009) study the work behavior of married second-generation American women in the US. The authors show that past values of female LFP in the country in which a woman's father was born – a variable that reflects culture as well as institutions and economic conditions in the home country – help explain the variation in how much married women from different ancestries work in the US even after controlling for a series of covariates (e.g. measures of human capital, area of residence, and several husband's characteristics such as education and earnings).[11] They conclude that culture plays a significant role in explaining cross-country patterns of female LFP.

Using a similar approach, Fernández (2007) shows that attitudes towards women's work in the father's country of birth, as expressed in answers to pertinent poll questions in the World Value Survey (WVS), likewise are significant in explaining the variation in the amount of market work performed by second-generation married American women in the US. Burda, Hamermesh, and Weil (2007) study instead first-generation American women and the *total* sum of hours in which they engage in work across the market and at home. They likewise find that cross-national variation in attitudes expressed in the WVS, and thus culture or social norms, is significant in explaining the variation in total number of hours worked by these women in the US.

2.2 Learning and Cultural Change

Neither the epidemiological approach to culture nor the dynamic quantitative models discussed above have explicit mechanisms that explain why culture might vary, whether across countries or over time. At least in part this absence has been due to the paucity of models of cultural *change*. The few papers that model the dynamics of culture do so either in an evolutionary theory framework (see, e.g. Bowles (1998)) or in utility-maximizing models in which the proportions of people with different *fixed* ideas (e.g., the proportion that is Catholic versus Jewish) evolve in the population over time. The latter is the route taken by Bisin and Verdier (2000) and Tabellini (2007) in which parents optimally choose whether or not to transmit *given* ideas. In those models individuals "believe" something with either probability one or zero, unlike a model of learning in which the probabilities that individuals assign to different views (ideas) change with the information that is available to them.

An exception to this is Hazan and Maoz (2002) and Fernández, Fogli, and Olivetti (2004). Hazan and Maoz are also inspired by the S-shape in the path of women's LFP to develop a model in which social norms regarding women's work evolves over time. In their model, working incurs the psychic cost of violating the social norm, which is assumed to decrease with the proportion of women who worked in the previous generation. For

[10] See Fortin (2005) for evidence that differences in attitudes help explain the variance in how much women work across OECD countries.

[11] See also Antecol (2000).

some parameters, the model can exhibit multiple locally-stable steady states – one with low participation and one with high participation. Which steady state is chosen depends on initial conditions, i.e. on the proportion of women working initially. To obtain the S-shape, the authors assume that there is normally distributed heterogeneity in work disutility among women. Thus, it is possible for few women to work early on and for LFP to evolve gradually over time giving rise to the shape of LFP found in the data.

In the model of Fernández et al (2004), working mothers are assumed to pass on to their male children a more positive view of working women, making these boys more amenable later on to having a working wife. This in turn renders it more attractive for girls invest in market-work human capital (as opposed to house-work human capital), as they know that men are more likely to be receptive to them working in the market. Their model does not necessarily imply an S shape for married women's LFP, however.

In contrast with the literature above, the present paper explicitly models changes in culture as arising from a process of learning.[12] In this process, the probability individuals assign to different views of the long-term consequences of married women working is updated in a Bayesian fashion as new information endogenously becomes available.[13] Although there is no direct evidence showing that learning is responsible for the change in social attitudes, as noted in the introduction the path followed by married women's LFP is the commonly denoted S-shape that arises in processes of technological diffusion.[14] Thus, the very shape of the LFP path may itself constitute a clue that a similar mechanism of information diffusion is also at play in this context, though on a very different time scale. Furthermore, using a learning mechanism to think about cultural change has the advantage, from an economist's perspective, that it lends itself to standard welfare analysis unlike in environments in which preferences themselves are changing.

For an approach that emphasizes learning to make sense, women need to have been uncertain about the consequences of working. Thus, one must come to grips with the question of what is it that women could have been learning about over the last 120 years. It is not an exaggeration to state that, throughout the last century, the consequences of women's (market) work have been a subject of great contention and uncertainty. In the US, as in other countries, a process of specialization accompanied industrialization and urbanization. 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

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

[13] A recent paper by Fogli and Veldkamp (2007) independently (but subsequent to the public presentation of the learning model developed in this paper), develops a related idea. They study the labor force participation of women with children for a substantially shorter period (1940-2000) and assume that women learn about the ability cost to a child from having a working mother. Learning occurs through sampling the ability outcomes of children of a small number of other women. Whereas in my model actions change because people modify their beliefs about the cost of working, in their model actions change only because of a reduction of uncertainty about the cost.

[14] See, e.g., 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.

separate. This process left the wife in charge of the domestic realm and her husband responsible for supporting the family, and spawned a debate on the consequences of a working wife 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.[15]

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 (around 80% prior to 1940) exited the workplace upon marriage.[16] Over time, the debate shifted to the effect of a married woman working on family stability and to the general suitability of women for various types of work and careers.[17] More recently, public anxiety regarding working women centers around the effect of a working mother on a child's intellectual achievements and emotional health.[18] 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 married woman working has been the subject of an evolving process of discovery and debate.

3 A Simple Learning Model of Work and Culture

This section develops a very simple model of a married woman's work decision in which she trades off between her enhanced consumption possibilities and her disutility from working. The disutility from working, in addition to reflecting a woman's known labor-leisure trade-off, is assumed to also 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 previously discussed). 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.[19]

The next two sections describe first the maximization problem faced by a woman in time t, given her prior beliefs and her private information and then, how these beliefs evolve across generations over time.

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

[16] 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.

[17] For example, at the turn of the 20th century, work and marriage were not seen as very compatible for educated middle-class women. Jane Addams noted that men "did not want to marry women of the new type, and women could not fulfill the two functions of profession and home making until...public opinion tolerated the double role." Cited in Collins (2003), p.294.

[18] See, for example, Bernal (2008), Keane and Bernal (2009) and Ruhm (2004) for reviews and recent findings of this literature.

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

3.1 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:

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

where the first component l_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-level realization of a random variable B that is iid across women. The value of this variable is revealed only if the woman works – it is the *long-run* disutility from working.[20]

Note that since v_i enters linearly in utility, only the expected value of B matters for women's work decisions, rather than its individual realization, B_i.[21] Women are assumed to be uncertain about the mean value of B, denoted by β. For simplicity, I assume that β can take on only two values, high (H) and low (L), i.e., $\beta \in \{\beta_H, \beta_L\}$. Note that β_L is the good state of nature in which working is on average not so costly, i.e., $\beta_H > \beta_L \geq 0$. B_i is assumed to be equal to β plus some noise term whose mean is zero. Thus, a woman's expected disutility from working is given by:

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

where E is the expectations operator. That is, women know their own value of l prior to making their work decision but have (possibly different) expectations about β and only learn their individual realization of B_i once they have worked for the entire period.

Consider a woman in period t who has a prior belief about the value of β as summarized in the log likelihood ratio (LLR) $\lambda_t = \ln \frac{Pr(\beta=\beta_L)}{Pr(\beta=\beta_H)}$. Prior to making her work decision, she inherits her mother's private signal s_i about the true value of β.[22] The private signal

[20] Note that a woman who obtains a bad realization of B does so at the end of her work life and thus she does not exit the labor market. A multi-period model would allow B_i to be learned slowly over time. In that case, one would also have women exiting the labor market if their expectations become more pessimistic. This would complicate the algebra but not otherwise change the main conclusions of the model.

[21] A simple linear specification allows us to abstract from risk aversion (the only force at work in Fogli and Veldkamp (2007)) and focus purely on learning.

[22] It is equivalent from a modelling perspective to have the signal inherited from one's mother or drawn afresh (e.g., a draw from the scientific literature that exists at that moment regarding the welfare conse-

is informative, i.e., it yields information about β. In particular, the value of the signal is given by:

$$s_i = \beta + \varepsilon_i \tag{5}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. We denote ε's cumulative distribution and probability density functions by $F(\cdot;\sigma_\varepsilon)$ and $f(\cdot;\sigma_\varepsilon)$, respectively.[23] The private signals are assumed to be i.i.d across women.

After 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

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

where $\bar{\beta} = (\beta_L + \beta_H)/2$.[24]

It is worth noting a few properties of $\lambda_{it}(s)$:[25]

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

(*ii*) $\frac{\partial^2 \lambda_{it}(s)}{\partial s \partial \sigma_\varepsilon^2} > 0$.

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 λ is decreasing with the variance of the noise term, σ_ε^2, since a greater variance lowers the informativeness of the signal.

Furthermore, from the assumption that $s \sim N\left(\beta, \sigma_\varepsilon^2\right)$, we can use (6) to derive the distribution of beliefs (strictly speaking, the distribution of the LLR) at any point in time:

$$\lambda_{it}(\beta) \sim N\left(\lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma_\varepsilon^2}\right)\left(\beta - \bar{\beta}\right), \frac{\left(\beta_H - \beta_L\right)^2}{\sigma_\varepsilon^2}\right) \tag{7}$$

Having derived the distribution of beliefs, we next turn to a woman's work decision. Assume that women share a common prior in period t, λ_t.[26] What proportion of women will choose to work that period? From (1) it follows that woman i will work iff

$$\frac{1}{1-\gamma}[(w_{ht} + w_{ft})^{1-\gamma} - w_{ht}^{1-\gamma}] - E_{it}(\beta) \geq l_i \tag{8}$$

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

quences of a woman working).

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

[24] 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 $N(\beta, \sigma_\epsilon^2)$.

[25] These follow directly from taking the appropriate derivative of equation 6.

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

$\frac{1}{1-\gamma}[(w_{ht}+w_{ft})^{1-\gamma}-w_{ht}^{1-\gamma}]$ by $W(w_{ht},w_{ft})$.

Note first that given $\{\beta_H,\beta_L\}$ and earnings (w_{ht},w_{ft}), irrespective of their beliefs and thus of the signal they receive, women with very low l ($l\leq \underline{l}(w_{ht},w_{ft})$) will always choose to work and women with very high l ($l\geq \bar{l}(w_{ht},w_{ft})$) will never choose not to work, where

$$\underline{l}(w_{ht},w_{ft}) \equiv W(w_{ht},w_{ft})-\beta_H \quad (9)$$

$$\bar{l}(w_{ht},w_{ft}) \equiv W(w_{ht},w_{ft})-\beta_L \quad (10)$$

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 $s_j^*(\lambda)$ that she would need to inherit from her mother in order to be indifferent between working and not working. Thus, for any $s\leq s_j^*$, given her prior belief λ, she would be willing to work. Let $p=Pr(\beta=\beta_L)$ and let p_j^* be the critical probability such that a woman of type l_j is indifferent between working and not, i.e.,

$$p_j^*\beta_L+(1-p_j^*)\beta_H=W(w_{ht},w_{ft})-l_j \quad (11)$$

and thus $p_j^*(w_{ht},w_{ft})=\frac{\beta_H+l_j-W}{\beta_H-\beta_L}$ or, using (9),

$$\ln\frac{p_j^*}{1-p_j^*}=\ln\frac{l_j-\underline{l}}{\bar{l}-l_j} \quad (12)$$

Thus, the critical value, s_j^*, of the private signal a woman of type l_j must inherit in order to be indifferent between working and not, given a prior of λ_t, is given by $\lambda_t(s_j^*)=\lambda_t-\left(\frac{\beta_H-\beta_L}{\sigma_\varepsilon^2}\right)(s_j^*-\bar{\beta})=\ln\left(\frac{l_j-\underline{l}}{\bar{l}-l_j}\right)$ and hence

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

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 given the true value of β and a prior of λ_t, $L_{jt}(\beta;\lambda_t)$, is the proportion of them that receive a signal lower than $s_j^*(\lambda_t)$, i.e.,

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

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

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

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 $\beta=\beta_H$. How $L_t(\beta_L;\lambda_t)-L_t(\beta_H;\lambda_t)$ evolves as λ changes will play an important role in the results that follow.

Before specifying how beliefs are transmitted across generations, it is worth noting a few features of $s_j^*(\lambda)$.[27]

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

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

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 states that increases in own earnings make women more willing to work (and thus willing to do so at a higher signal) whereas the opposite holds for increases in husband's earnings. The proposition below follows directly:

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 satisfies the traditional (desirable) comparative statics results with respect to wages, which is not surprising as these properties follow entirely from the specified preferences.

3.2 Intergenerational Transmission of Beliefs

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 generation $t+1$ inherits the prior of generation t (its "culture"), λ_t, which each individual then updates with her own idiosyncratic private signal, s_i, inherited from her mother. An entirely equivalent assumption is that each woman transmits her belief to her daughter, i.e., a child inherits $\lambda_{it}(s)$ from her mother. In either case, a woman in $t+1$ ends up with the belief $\lambda_{it}(s)$. If solely this information were transmitted intergenerationally, then $\lambda_{it}(s) = \lambda_{it+1}(s)$, $\forall i$. In that case, only changes in wages could lead to changes in work behavior. There is, however, potentially an additional source of information available to women in $t+1$ that was unavailable to women at time t — the aggregate proportion of women who worked in period t. The women making a work decision in period $t+1$ will want to use this information to update $\lambda_{it}(s)$ if it is available.

If generation $t+1$ were able to observe perfectly L_t, the aggregate proportion of women who worked in period t, L_t, they would be able to back out β 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, the notion that this variable is completely informative seems equally implausible[28] – it is merely an artifact of the simplicity of the model. In particular, with additional sources of heterogeneity in the model, backing out the true value of β would require agents to know the geographic distribution of male earnings and female (potential) earnings and how they were correlated

[27] Each property follows directly from taking the appropriate derivative of equation (13).

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

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.[29]

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

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

and where $\eta_t \sim N(0, \sigma_\eta^2)$ with a pdf denoted by $h(\cdot; \sigma_\eta)$. The assumption that η (or z in the example developed below) is distributed normally should be taken as an approximation made for analytical simplicity.[30] 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.

Alternatively, one could assume that individuals perfectly observe aggregate LFP, but are uncertain about the distribution of a parameter that affects the distribution of the idiosyncratic disutility of work, $G(l)$. 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 shock in the economy).[31] One very simple formulation of this alternative is to assume that a fixed fraction $\omega > 0$ of the population is subject to extreme preference shocks such that a *random* proportion z_t of this fraction works, whereas the remaining proportion $1 - z_t$ does not work, independently of wages. Thus the aggregate proportion of women who work, $\widetilde{L}_t(\beta; \lambda_t, z_t)$, would be given by:

$$\widetilde{L}_t(\beta; \lambda_t, z_t) = (1 - \omega) L_t(\beta; \lambda_t) + z_t \omega$$

where $L_t(\beta; \lambda_t)$ is given again by equation (15). Assuming that z is distributed normally then yields an expression equivalent to (16).

Returning to the original model, given a common inherited prior of λ_t implies that after observing last period's signal of aggregate female LFP, y_t, using Bayes law generates an

[29] This is the strategy used in finance, for example, by introducing noise traders. An alternative assumption 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 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 (2009) 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.

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

[31] See Chamley (1999).

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

$$\lambda_{t+1}(\lambda_t, y_t) = \lambda_t + \ln \frac{h(y_t|\beta = \beta_L)}{h(y_t|\beta = \beta_H)}$$

$$= \lambda_t + \left(\frac{L_t(\beta_L; \lambda_t) - L_t(\beta_H; \lambda_t)}{\sigma_\eta^2} \right) \left(y_t - \overline{L}_t(\lambda_t) \right) \quad (17)$$

where $\overline{L}_t(\lambda_t) = \frac{L_t(\beta_L;\lambda_t)+L_t(\beta_H;\lambda_t)}{2}$. Thus, observing values of y_t that lie above the mean value that LFP would take under the two states of nature, $\overline{L}_t(\lambda_t)$, updates λ upwards and the opposite if y_t lies below $\overline{L}_t(\lambda_t)$. As before, a noisier signal (in this case, a larger value of σ_η^2), implies less updating of the common prior.

Figure 3 summarizes the time line for the economy. Individuals start period t with a common prior, λ_t. Each woman updates the common prior with her inherited private signal and makes her work decision. This generates 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 (λ_t), generating λ_{t+1} – the "culture" of generation $t+1$.[32] It should be noted that instead of assuming that women in $t+1$ inherit λ_t (or λ_{it}) which they update with the information contained in y_t, we can equivalently assume that women always have the same common prior of λ_0 and access to the entire history of y_τ, $\tau = 0, 1, 2..., t$ which they use to update λ_0. This would yield the same value of λ_{t+1} (or of λ_{it+1}) as in the first interpretation.

Note that (17) is the law of motion of the commonly held prior, λ, which we can think of as "culture." That is, culture at time t is given by the core beliefs, encapsulated in λ_t, common to all women at that point in time. This does not mean that women share exactly the same beliefs (as they do not in any society). Heterogeneity in beliefs at a point in time in this model stems from the private information inherited from one's mother, and thus has a distribution as expressed in (7). Furthermore, culture will change over time as new information becomes available. Sometimes this change will be large and at other times it will be small. This model provides predictions, discussed in the next section, as to how culture will tend to evolve and the ensuing time path of female LFP.

4 The Dynamic Evolution of Culture and LFP

The learning model developed above has several important properties that will prove useful in generating LFP dynamics similar to those in figure 1. It also gives rise to a new role for wage changes and their effect on women's labor supply. Below I discuss both in turn. It is useful to note first that beliefs in this model are unbounded. Hence, in the long-run beliefs must converge to the truth.[33] Since female LFP has been increasing over time, this implies that it is more likely that $\beta = \beta_L$ and we shall henceforth assume that this is the

[32] As explained below, we can think of generation τ as having a shared culture given by λ_τ. Individual deviations around the median λ_τ (given by the normal distribution of $\lambda_{i\tau}$ (s)) constitute the distribution of beliefs induced by different individual's dynastic histories (i.e., by their inheritance of different s).

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

case.

4.1 Generating an S-Shape in Culture and LFP

A key characteristic of this model is that it naturally generates an S-shaped LFP curve. To understand why this is so, we start by noting from (15) that, ceteris paribus, the size of change in λ is positively related to the size of the change in LFP. Thus, understanding the shape of the LFP curve requires us to know how λ varies over time.

To determine how λ tends to evolve, 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 those who work when $\beta = \beta_H$, i.e., in $L_t(\beta_L;\lambda_t) - L_t(\beta_H;\lambda_t)$. Thus, ceteris paribus, there will be significant changes in work behavior over time when $L_t(\beta_L;\lambda_t) - L_t(\beta_H;\lambda_t)$ is large and small changes when it is small.

To see why the above will generate an S-curve requires understanding, therefore, how the size of $L_t(\beta_L;\lambda_t) - L_t(\beta_H;\lambda_t)$ varies over time. Insight into this is easiest by first abstracting from the heterogeneity in l types in the economy and noting that for a given $l_j \in (\underline{l}, \overline{l})$

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

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 maximized at $s_j^* = \overline{\beta}$.

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.[34] Why does this happen? 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 η_t. Thus, the intergenerational updating of beliefs will be slight and 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 $\overline{\beta}$. In that case, the work decision depends entirely on

[34] Note that the distance of s^* from $\overline{\beta}$ depends not only on the value of λ, but also on w_f and w_h as these affect a woman's willingness to work as well.

whether the signal is closer to β_L or to β_H.

The conclusion above follows from the general features of the normal distribution of ε, which is depicted in figure 4.[35] As can be seen in the figure, when $s^* - \overline{\beta}$ is far from zero, the difference in the proportion of women who work in the two states is small. That is, the difference between the values of L_j at $s^* - \beta_L$ and at $s^* - \beta_H$, (i.e., the shaded area) is small, and thus not very informative about the true value of β. 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, if the economy consisted only of individuals of type l_j, beliefs would change substantially from one generation to the next, leading to large changes in behavior.

The intuition above can now be applied to the economy with heterogeneity in l_j types, for which a similar conclusion holds. Taking the derivative of $L(\beta_L; \lambda_t) - L(\beta_H; \lambda_t)$ with respect to λ 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 far from $\overline{\beta}$ for a substantial proportion of individuals, (21) will be small in absolute value, intergenerational updating will be small, and the evolution of LFP over time will be slow. The opposite is true when the critical signal is close to $\overline{\beta}$ for a substantial proportion of individuals.[36]

The S-shape follows from the logic above. If parameter values are such that initially a substantial proportion of women require an extreme signal (i.e. very large negative values of s) in order to be willing to work, then learning will be slow for several generations since little information will be revealed by the aggregate signal y. 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 less extreme values of s to change their work behavior, learning accelerates since the aggregate signal is very informative leading to large changes in aggregate LFP over time. This is the second (rapidly rising) portion of the S curve. Lastly, once beliefs are optimistic and women, on average, require extreme signals (i.e., very large positive values of s) in order to choose *not to work*, learning once again slows down since the aggregate signal becomes relatively uninformative. Changes in aggregate female LFP during this period are 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 LFP result solely from changes in wages.

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

[36] Note that, at each λ, s_j^* ranges from $-\infty$ (for $l_j = \underline{l}$) to $+\infty$ ($l_j = \overline{l}$). Thus, what determines the general shape of the LFP curve is how the value of s_j^* evolves for the range of $l_j \in (\underline{l}, \overline{l})$ where is populated by a substantial mass of individuals. It should also be noted here that the assumption of a normal distribution for l_j does not play an important role.

4.2 A New Role for Wages

The learning model generates a novel role for wages. As in a standard labor supply 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 $E_t(\lambda_{t+1} - \lambda_t)$. This does not happen, as one might initially think, by greater numbers of working women providing more information about the welfare consequences of working.[37] Rather, the mechanism operates by having wage changes affect s^* and, through this, wages affect the *informativeness* of the aggregate work outcome.

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) any force that increases s^* will also increase the difference in the proportion of working women across the two states of nature and hence also the informativeness of the aggregate signal for the next generation. This mechanism can take the form of increases in female wages, decreases in male wages, 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 (2009b)) or policy changes that increase the attractiveness of work.

The above does not imply that increases in women's wages will always increase the pace of learning. The opposite effect would be present if women were initially very optimistic about β and hence required, on average, very high values of s^* in order not to work. In that case, wage increases would tend to decrease the amount learned by the next generation.

Thus, in general, wage changes (or other changes that affect s^*) 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 evaluate the quantitative importance of this effect in explaining the historical evolution of female LFP.

5 Quantitative Analysis

In this section I examine the quantitative potential of the simple learning model. Given a path of wages, the degree of heterogeneity in the taste for work/leisure (σ_l), an initial probability that $\beta = \beta_L$, a distribution of private information (σ_ε), and a variance for the noise in the observation of aggregate LFP (σ_η), the model delivers predictions for female LFP and for the path of beliefs. The model also delivers predictions with respect to the quantitative importance of wages and beliefs in changing female LFP over time. This will allow us to assess the significance of the new dynamic learning role of wage changes (discussed in section 4.2) over the last century.

The analysis presented below neither constitutes a "test" of the model nor nests the learning mechanism within a larger model that allows channels other than wages and learn-

[37] It would be easy though to incorporate this additional channel which is standard in the literature.

ing to operate. Both alternatives would require one to restrict attention to a significantly shorter time period for which more detailed data is available (in most cases, after 1940). Furthermore, considering alternative mechanisms requires using more complicated multi-period dynamic models of female labor supply as in Eckstein and Lifschitz (2009) or Attanasio, Low, and Sanchez-Marcos (2008). Note that by including wages, the model incorporates a variable that has been found to be an important determinant of the increase in married women's LFP by several authors (see, e.g., Jones, Manuelli, and McGrattan (2003) or Heathcote, Storesletten, and Violante (2010)). As this is the first paper to explore the dynamic contribution of the evolution of cultural beliefs to women's work, assessing its quantitative potential in a simple model that allows both the theory and the calibration to be fairly transparent is an important first step to subsequently developing more complicated quantitative dynamic models.

5.1 Calibration Strategy

In the model, married women decide whether to engage in market work taking their husbands' earnings and their own potential earnings 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 is available as of 1890.[38] 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. After 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. The details for the construction of the earnings series are given in the Appendix.

Figure 5 shows the evolution of female and male median earnings over the 120 year period 1880-2000 (with earnings expressed in 1967 dollars). In order to compare data sets, 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.[39]

The evolution of married women's 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

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

[39] 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.

(non-farm occupations and non-group quarters). I calibrate the model to match female LFP in 1980, 1990 and 2000. The remainder of the LFP series is generated endogenously by the model. 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 calibration targets.

I take the elasticity estimates from Blau and Kahn (2007) who 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. For the year 2000, Blau and Kahn estimate an own-wage elasticity of 0.30 and a cross-elasticity (husband's wage) of -0.13. The cross elasticity in 1990 is -0.14. See the Appendix for a discussion of the exact specification chosen. If instead of matching the cross-wage elasticity in 1990 one matches the own-wage elasticity in that year (0.44), very similar results are obtained. In general, the model is not very sensitive to the choice of calibration targets.

To calculate the probability that a woman worked in 1980 conditional on her mother's work behavior, I use the General Social Survey (GSS) from 1977, 1978, 1980, 1982, and 1983.[40] 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 between the probability of a woman working given that her mother worked, to the probability of a woman 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. The GSS sample for this calculation was restricted to all white married women between the ages 25-45 who were born in the U.S.[41] 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).[42]

The calibration targets are summarized in table 1.

5.2 Calibration Results

Before turning to the calibration of the full model, it is instructive to calibrate a simpler version of it in which β is known so there is no learning.

[40] 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.

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

[42] 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 (2009).

5.2.1 The Model without Learning

Shutting down the learning channel implies that only changes in wages (male and female) can explain why labor supply changes over time. The unknown parameters are now three instead of seven (γ, β, and σ_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, β, and about the necessary dispersion of l types 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, β, note that this is equivalent to 4.7% of the utility obtained from consumption in 2000 if a woman worked 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.

This basic inability of the model absent learning 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 α 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 not help matters as it would require women to have obtained a much larger share of husband's earnings in the past in order to explain the much lower participation rates 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., greater bargaining power and hence a higher share of male earnings 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

[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.

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 wildly overpredicted LFP.

It is also important to note that the failure of the model without learning does not follow from the structure of the model. If, instead of calibrating, one chooses parameter values that minimize the sum of squared errors between the predicted LFP path and the data, this actually results in a better fit (the sum of squared errors is 0.021) than what we obtain with the fully calibrated model below.

5.2.2 The Full Model

We now turn to calibrating the full model (i.e., β is unknown and women endogenously learn its value over time). As married women's 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 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 of η_t generates a corresponding different public belief λ_{t+1} in the following period, and consequently a different proportion of women who choose to work after receiving their private signals. Note that we cannot simply evaluate the model at the mean of the expected η shocks (i.e., at zero) since, although λ_{t+1} is linear in η, the work outcomes L_{t+1} are not.

I deal with the aggregate shock in the following way. For each period $t+1$, given L_t, I calculate the proportion of women who would work, L_{t+1}, for each possible realization of the shock, η_t, i.e., for each induced belief $\lambda_{t+1}(\eta_t)$. Integrating over the shocks, I find the expected value of LFP for that period, $E_t(L_{t+1}(\lambda_{t+1}(\eta_t)))$, and then back out the particular public belief (or shock) that would lead to exactly that same proportion of women working, i.e., I solve for $\lambda^*_{t+1}(\eta^*_t)$ such that:[44]

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

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

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

As shown in the Appendix, manipulating the ratio of elasticities in this model yields the same value of γ as absent the learning channel, i.e., $\gamma = 0.503$. As in the simplified model, the elasticities and female LFP levels yield information both about how bad women believe it is to work and how much heterogeneity there is across women are in their willingness to work at given wages. Unlike before, however, this dispersion is a function not only of the distribution of the l types, σ_l, but also of the distribution of private information, σ_ε. Furthermore, women's beliefs about the value of β are evolving over time. The values of LFP from 1980-2000 yield information as well on how rapidly the commonly held portion of beliefs, λ_t, needs to evolve over these decades and hence on the variance, σ_η, of the noise η in the public signal.

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 λ and how different the values of β_H and β_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.[45] 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 the prior calibration exercise in the model with no learning.

[45] 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.

Table 1

Calibration Targets		Model with known β	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			
γ		0.503	0.503
σ_L		2.293	2.085
β		0.321	
β_H			4.935
β_L			0.001
$P_0(\beta = \beta_L)$			0.086
σ_ε			5.288
σ_η			0.055

All elasticities are from Blau & Kahn (2007). The work risk ratio uses data from GSS (see text). The values **in bold** are the model's predicted values for its calibration targets.

5.3 The Calibrated Model: LFP

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 β_L that is held by the median woman.

As can be seen from the figure, the calibrated model on the whole does a fairly good job of replicating the historical path of married women's LFP.[46] It under-predicts LFP from 1930 to 1970, however, and slightly over-predicts it from 1880 to 1900.

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 the war

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

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 γ. 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 β in 2000 is not very different from the value of β obtained when learning is eliminated: the individual with median beliefs has an expected value of β of 0.26 whereas without learning $\beta = 0.32$.[47] Thus, it is the endogenous evolution of the expected value of β in the learning model that is responsible for the difference in the LFP behavior observed over time across the two models. Whereas by construction the expected value of β remains constant when learning is eliminated, the median expected value of β in the learning model is close to 4.51 in 1880 and evolves over time to 0.26 in 2000. This allows LFP to respond in dramatically different ways over time.

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.

5.4 The Calibrated Model: Beliefs

As noted previously, the model is able to generate not only predictions for female LFP, but also for the evolution of social attitudes. Individuals start out in 1880 with pessimistic beliefs about how costly it is to work.[48] 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 β_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%.

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.[49]

It would be interesting to contrast the model's prediction about social attitudes with those of the data, especially as this is a unique feature of this model. Doing so, however, is complicated by the fact that there is no data that reports directly on women's beliefs about the cost of working. An imperfect indication however can be obtained by using poll data, as discussed below.

As mentioned previously, over some period of time polls asked a few variants of the question "Do you approve or disapprove of a married woman earning money in business or

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

[48] It is worth noting that this pessimism is fully congruous with social views concerning married women working (see the quotes in the introduction and the poll data).

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

industry if she has a husband capable of supporting her?".[50] The solid line in Figure **??** shows the percentage of white married women between the ages of 25-45 who agreed with the question. To compare the model's predictions and the poll data, one needs to map the distribution of probabilities at any point in time to the binary variable ("approve" versus "disapprove") available in the data. To do this, I use equation (7) to find the cutoff belief ($\widetilde{\lambda}$) in the model's predicted belief distribution in 1940 such that the proportion of women with more optimistic beliefs (i.e. those with $\lambda \geq \widetilde{\lambda}$) is equal to the proportion of women who approved of women working in the poll data at that time, 0.27.[51] This allows me to calculate from the model, for every period, the proportion of women who would approve of married women working, i.e., those with beliefs above $\widetilde{\lambda}$ and contrast it with the data.

The dashed line in Figure **??** shows the model's predictions for the proportion of women who approve of married women working. As can be seen from the figure, the shape of the beliefs path is quite similar to the data path but the level is higher throughout by some 10 percentage points.

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.

5.5 The Quantitative Contributions of Wages and Beliefs

To investigate the quantitative contributions of earnings and beliefs, we start by not allowing public beliefs to evolve (i.e., shutting down the public signal). First, we can freeze beliefs at the 1880 level (with the median individual's prior that $\beta = \beta_L$ at 9%) and ask how labor force participation would have evolved in the absence of any updating of beliefs. Thus, women have private information which they transmit to their descendants, but λ does not evolve over time. As show by the bottom line (with the caption "LFP if no public updating") in Figure 9, female LFP would have barely exceeded 10% by the year 2000.

Alternatively, one can ask what LFP would have been if, throughout the entire time period, agents had known the true value of β, i.e., $\beta = \beta_L$. This scenario is shown for the parameters of the calibrated model= by the top (red) line (with the caption "full information LFP"). It 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 either of the two extreme scenarios 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 9). The model with dynamics induced solely by changes in male and female earnings (with unchanged beliefs) grossly under or over estimates women's labor supply over the entire time period.

To distinguish between the static and dynamic effects of wage changes on female LFP we can perform the following instructive decomposition. First, as before, we can keep

[50] See the Appendix for details.

[51] There is no poll data for 1940. Instead I use a linear interpolation between the 1938 and 1945 points to obtain it.

wages constant at their initial 1880 levels, henceforth denoted by w_0, and let beliefs change endogenously over time given these wages. The LFP path obtained in this fashion, denoted $L(\overline{\lambda}(w_0), w_0)$ in Figure 10, results only from the changes in beliefs that would have occurred had earnings stayed constant at their 1880 levels.[52] 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 10. Hence, the difference between the level of LFP in 1880 (given by the dotted horizontal line) and $L(\overline{\lambda}(w_0), w_0)$ measures the contribution of changes in beliefs to the historical evolution of female LFP.

Next we can disentangle the dynamic from the static effect of wages by confronting women with the actual historical earnings path, $\overline{w}$, but endowing them with the belief path obtained from the exercise above, $\overline{\lambda}(w_0)$. In this exercise, changes in wages affect the attractiveness of working vs not working for the usual reasons, but their dynamic effect on the intergenerational evolution of beliefs is eliminated since, by construction, these beliefs are derived from a constant (1880) wage path. We denote the LFP obtained this way by $L\left(\overline{\lambda}(w_0), \overline{w}\right)$ and it is shown with (red) x's in the figure. The difference between $L(\overline{\lambda}(w_0), w_0)$ and $L\left(\overline{\lambda}(w_0), \overline{w}\right)$ 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 $L\left(\overline{\lambda}(w_0), \overline{w}\right)$).

Lastly, we allow wages to also influence intergenerational learning and thus beliefs; let the LFP path obtained this way be denoted by $L\left(\overline{\lambda}\left(\overline{w}\right), \overline{w}\right)$).[53] It is the top (blue) curve shown in Figure 10. The difference between $L\left(\overline{\lambda}\left(\overline{w}\right), \overline{w}\right)$ and $L\left(\overline{\lambda}(w_0), \overline{w}\right)$ measures the dynamic contribution of wages through its effect on beliefs (i.e., both series have the same historical earnings series, $\overline{w}$, but $LFP\left(\overline{\lambda}\left(\overline{w}\right), \overline{w}\right)$ allows beliefs to respond to these changes whereas $L\left(\overline{\lambda}(w_0), \overline{w}\right)$ keeps the belief path corresponding to constant wages at their 1880 level).

As can be seen in Figure 10, 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, $\overline{\lambda}\left(\overline{w}\right)$ and $\overline{\lambda}(w_0)$, depicted in figure 11.[54] 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

[52] In this section I use a bar to denote a variable that is changing over time. In this case, although wages are fixed, beliefs change over time as individuals learn.

[53] Note that this LFP path is the one predicted by the model and depicted previously in Figure 7.

[54] The figure depicts the evolution of p, i.e., the probability that $\beta = \beta_L$.

dynamic effect of earnings on beliefs.

It should be noted that there is not a unique way to decompose LFP in order to measure the quantitative importance of wages and beliefs. One could alternatively eliminate the $L\left(\overline{\lambda}\left(w_0\right), \overline{w}\right)$ curve and replace it with the LFP path that would result if beliefs followed the path obtained from the historical earnings series, $\overline{w}$, but wages were kept constant at their 1880 levels. This curve is shown in Figure 12 as $L\left(\overline{\lambda}\left(\overline{w}\right), w_0\right)$. The dynamic effect of wages is now given by the difference between $L\left(\overline{\lambda}\left(\overline{w}\right), w_0\right))$ and $L\left(\overline{\lambda}\left(w_0\right), w_0\right)$. These paths are obtained using the same constant 1880 earnings, but in the first trajectory beliefs evolve as they would with the historical earnings profile, whereas in the second 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 $L\left(\overline{\lambda}\left(\overline{w}\right), w_0\right)$ and $L\left(\overline{\lambda}\left(\overline{w}\right), \overline{w}\right)$, 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 changes in beliefs induced by higher earnings were 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.

6 Testable Implications of the Model and Welfare

The model can also be used to derive several implications regarding elasticities, future levels of female LFP, cross-sectional predictions, and inter-generational dynamics. It can also be used to quantify the welfare costs of imperfect information.

6.1 Elasticities and Future LFP Predictions

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 evolve with changes in wages and beliefs. The own and cross wage elasticities predictions are shown in Figure **??** (given by the highest and lowest lines, labelled LM). Recall that the model is calibrated to match both elasticities in 2000 and the 1990 cross elasticity. As can be seen from the picture, over time the two elasticities are first increasing (in absolute value) and then decreasing. This pattern is similar to the historical one reported in Goldin (1990) with respect to women's own-wage elasticity. Speculating, in the early decades these elasticities may reflect 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. By the 1960s, there is a much more widespread belief that it is not bad for a woman to work and there is a large drop with respect to the sensitivity to both her own and her husband's wages.[55]

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 **??**.

As an additional 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.[56]

6.2 Cross-Sectional Implications

The model can also be used to generate LFP paths for women of different education levels, e.g. college vs high school. Doing this requires earnings data by women's education types and for their husbands. Unfortunately, the Census does not provide either wage or education data prior to 1940. Thus, I instead examine the implications of the calibrated model from 1940 onwards.

In particular, for each decade, I divide the women from the original sample into two categories: those with at most a high-school degree (labelled "high school") and those with at least some college education (labelled "college"), and then match these women with their husbands. Making use of the same procedure to calculate median earnings for employed people as in the original sample (see the Appendix), one can obtain an earnings series for both types of women and their husbands (see Figure **??**) which can then be used along with calibrated model to generate the path of LFP for both skill types.

The only remaining obstacle is the absence of a prior belief for women in 1940. To deal with this, one can create a prior for 1940, by skill type, by running the model from 1880 onwards using the same initial prior for both types of women as in the calibrated model ($p_{1880} = 0.086$), but with a different wage series for each type. Since there is no wage-education data until 1940, I assume that for each decade prior to that year, the ratio of the wages of each education type relative to the wages of the female population at large is at the 1940 level.[57] This implies a skill premium for women of 1.4, which seems a reasonable figure. I perform a similar procedure for these women's husbands (i.e., I assume that the ratio of median wages for the husbands of each type of woman to the median wages of men

[55] 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.

[56] The recession, however, may well invalidate these predictions.

[57] This is probably not such a bad assumption. During the 1940-1970 period, the ratio of married college women's median earnings to the median woman's earnings was relatively constant as was the same ratio for married high school women. For college women the mean of this ratio was 1.21 with a standard deviation of 0.04; for high school women the mean of this ratio was 0.88 with a standard deviation of 0.04. This relatively tight band is also found for each type's husbands.

was at the same ratio as in 1940). This calculation implies that the 1940 prior of college women is 0.148 and that of high-school women is 0.089.

The results of this exercise are given in Figure **??** . As can be seen from the figure, the model does a surprisingly good job at capturing the general shape of the increase in LFP for both college and high school women. It also does a very good job at predicting LFP levels for college women for the four decades from 1970-2000. It doesn't do a very good job for the LFP of high-school women, however, underpredicting the LFP by a substantial amount for every decade prior to 2000. On the other hand, it consistently predicts college women working more than high school women. Note that the model does not imply this automatically since a woman's work decision depends on the wage levels of both wives and husbands as well as on beliefs. Thus, it is a result of the calibrated model rather than a qualitative prediction.[58]

6.3 Dynamic Implications

The model also has implications concerning the intergenerational speed of learning. As explained in section 4.2, the model implies that wage changes, technological change in the household, or changes in policies that influence women's willingness to work, all have dynamic consequences as they affect the rate at which the next generation learns and consequently the path of LFP. Thus, if one could find exogenous variation in the magnitude of these changes one might be able to use it to examine the dynamic implications of the learning model.

One historical episode that potentially created variation across US states in women's willingness to work is World War II (WWII). As shown by Acemoglu, Autor, and Lyle (2004), women worked more in 1950 (but not in 1940) in those states that had a greater mobilization rate of men during WWII, even after controlling for other differences across these states (e.g., differences in age, education, or racial composition, differences in the importance of farming or occupational structure, or Southern vs non-Southern state differences).[59] They attribute the cause of the variation in female labor supply across states in 1950 to the greater participation of women during the war years, with some women staying in market work after the war ended, in states with higher mobilization rates. This allows them to interpret the mobilization rate as a source of exogenous variation in female LFP across states and to use it to analyze the effect of the latter on male and female wages.

For our purposes here, we can make use of the same source of exogenous variation in female LFP to examine the intergenerational implications of the model. In particular, the model implies that if states with greater mobilization rates had higher female LFP during the war years either because aggregate demand for women's work was greater (i.e., women's

[58] It should also be noted that we are making a number of data and modelling assumptions that may make it more difficult for the model's predictions to match the data, including the assumption that the unknown cost of working is the same for both types of women and that there is no selection into education (e.g., by belief).

[59] The mobilization rate is the number of men of age 18-44 who served divided by the number registered in each state.

wages were higher) or because these women needed to make up for the decline in household income (i.e. husbands' wages were lower) or in response to the campaign to get women to work as part of the war effort (recall, e.g., the iconic image of "Rosie the Riveter"), ceteris paribus this should result in the next generation of women learning and working more. According to the model, the pessimistic beliefs held in 1940 about the payoff to women's work implies that changes in variables that positively affect women's willingness to work will have positive repercussions on the amount learned by the next generation. In particular, changes in these variables will increase the variation in the proportion of women who work in 1940 under the two possible states of nature, leading to more rapid learning, and thus ultimately to the next generation working more. Consequently, at the empirical level, one can examine whether US states with greater mobilization rates during WWII had not only greater female labor supply in 1950, but also greater female labor supply in the next generation. That is, can one also find a positive effect of WWII on female LFP in the following generation?

For rather different purposes, Fernández, Fogli, and Olivetti (2004) showed that states with higher mobilization rates tended to have higher labor supply of white married women within a narrow age range (those likely to have children), even after controlling for other sources of inter-state variation as in Acemoglu et al (2004). They then showed that the effect of the mobilization rate persists for the white married women of next generation. This is what they call an "echo" effect of the war on women who would have been too young to be affected by the war directly but old enough to have seen the change in female LFP of their mothers' generation.[60]

Fernández et al (2004) interpreted their finding as being consistent with the hypothesis that working mothers passed on to their male children a better view of working women which made them more amenable to having a working wife. This then in turn made it more attractive for the girls in that generation to prepare to work (e.g., by investing in market-specific, rather than home-specific, human capital) as they knew that men would tend to be more receptive to them working once married. Their finding, however, is also consistent the hypothesis that women in the post-war generation changed their behavior because their beliefs about the payoff to working changed as a result of the more precise information available in the signal. Note that both interpretations necessitate a change in culture: in one case through a change in men's preferences (or beliefs) and in the other primarily through a change in women's beliefs resulting from learning. In reality, both channels are likely to have played a role and the available evidence does not permit one to distinguish between the two.

6.4 International Evidence

Do other countries' path of married women's LFP also look S-shaped? If they did, this would lend greater support to the hypothesis that a process of cultural change driven by

[60] For greater depth, the interested reader is referred to Fernández et al (2004), and in particular to tables III and IV in that paper.

learning took place. Unfortunately there does not appear to be readily accessible data on most other countries' evolution of female LFP over long periods of time and even those that do, tend not to distinguish between farm and non-farm workers, marital status, nor among different age groups. The restrictions to non-agricultural, married, and within a narrow age group are important as one does not wish the LFP series to be driven by the transition from an agricultural to an industrial economy or by demographic changes in the population.

I have, however, been able to plot married women's LFP for two countries for which relatively lengthy data series exist in tables of published papers: France for 1921-1981 and Great Britain between 1911-1998. As shown in Figure **??**, female LFP for these countries seem to follow a general S-shaped curve.[61] These countries are likely to have had a similar culture to the one in the US and thus this could have led to a similar evolution of LFP.

6.5 The Welfare Costs of Imperfect Information

Another interesting exercise is to quantify the welfare costs of imperfect information. While this is not a policy-relevant calculation in that the government, for example, is not assumed to be better informed than any individual, it gives an idea of how costly mistaken beliefs were and how this cost evolved over time.

To quantify the losses from imperfect information we start by noting that if women were given the true value of β, only those individuals of type $l_j \in (\underline{l}_t, \bar{l}_t)$ who as a result of their private information did not work in time t would change their decisions (and thus their utility).[62] All women with $l_j \leq \underline{l}_t$ worked, and all those with $l_j \geq \bar{l}_t$ would choose not to work even if they knew the truth.[63]

Thus, to quantify the loss of welfare, we can calculate at each moment in time, for each $l_j \in (\underline{l}_t, \bar{l}_t)$ type, the amount of consumption, z_{jt}, that a woman would have to be given in order to make her as well off as she would have been had she worked:

$$\frac{(w_{ht} + z_{jt})^{1-\gamma}}{1-\gamma} = \frac{(w_{ht} + w_{ft})^{1-\gamma}}{1-\gamma} - \beta_L - l_j \tag{24}$$

To interpret equation (24), note that the right-hand side is the utility enjoyed by a woman of type l_j who works. Thus, the left-hand side solves for the consumption equivalent of that utility.

The proportion of a given l_j type who made the wrong decision is given by those whose private signals lay above s^*_{jt} (as expressed in equation (13)), i.e., a fraction $1-F(s^*_{jt}-\beta_L;\sigma_\varepsilon)$. Integrating over the l_j types yields the aggregate welfare loss for these women at time t:

$$Z_t \equiv \int_{\underline{l}}^{\bar{l}} z_{jt}(1 - F(s^*_{jt} - \beta_L; \sigma_\epsilon))g(l_j)dl$$

[61] See the notes accompanying the figure for the details on how these were constructed.

[62] The top (red) line in Figure 9 shows what the evolution of female LFP would have been with full information.

[63] Note that $\underline{l}_t, \bar{l}_t$ have time subscripts since their values depend on wages which are changing over time.

In order to contrast this with the welfare enjoyed by working women, we need to translate the utility of an l_j type who worked into consumption units. This is given by finding x_{jt} such that:

$$\frac{(w_{ht}+w_{ft}+x_{jt})^{1-\gamma}}{1-\gamma}=\frac{(w_{ht}+w_{ft})^{1-\gamma}}{1-\gamma}-\beta_L-l_j \tag{25}$$

The proportion of a given l_j women who made the correct decision is given by $F(s^*_{jt}-\beta_L;\sigma_\varepsilon)$ if $l_j\in\left(\underline{l}_t,\bar{l}_t\right)$ and equals one if $l_j\le\underline{l}_{jt}$. Thus, the aggregate welfare of working women, expressed in consumption units, is given by:

$$W_t\equiv\int_{\underline{l}}^{\bar{l}}(w_{ht}+w_{ft}+x_{jt})F(s^*_{jt}-\beta_L;\sigma_\epsilon)g(l_j)dl+\int_{-\infty}^{\underline{l}}(w_{ht}+w_{ft}+x_{jt})g(l_j)dl \tag{26}$$

Lastly, the total welfare of non-working women expressed in units of consumption, is given by:

$$N_t\equiv(1-L_t)\,w_{ht}$$

Thus, the welfare lost as a result of imperfect information as a proportion of married women's welfare at time t, translated into consumption units, is given by $\widehat{c}_t$:

$$\widehat{c}_t\equiv\frac{Z_t}{W_t+N_t} \tag{27}$$

Note that $\widehat{c}_t$ gives the proportion of married women's average consumption lost as a result of imperfect information, when all utility is expressed in consumption units.

Figure **??** shows the evolution of $\widehat{c}_t$ over time. It starts out very high at some 39.29% of average consumption, decreases slowly until 1950, and then decreases dramatically to 0.19% by the year 2000. The very high numbers at the beginning are the consequence of the fact that the calibrated model implies that 63% rather than 2% of married women would have been working in 1880 had they possessed full information. A model in which the real cost of working evolved over time so that it was higher in 1880 than 2000 would imply smaller numbers as would a model in which the cost of working increased with the number of children.[64] Thus, these numbers should be taken as an upper bound to the costs of imperfect information.

7 Discussion and Conclusion

This paper modeled the joint dynamics of married women's labor force participation and cultural change. In a learning 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. I showed that a simple model with these features, calibrated to key statistics from the later

[64] The total fertility rate of white women in 1880, for example, is estimated to have been 4.24 as compared to 2.05 in 2000 (Haines (2008)). On the other hand, fertility is an endogenous variable and women might have chosen to have fewer children had they faced a higher known opportunity cost.

part of the 20th century, generates a time trend of married women's LFP that is roughly similar to the historical one in the US over the last 120 years. A binary mapping of the evolution of social beliefs in the model follows a path that is similar to that seen in the poll data.

This model naturally generates the S-shaped curve of LFP, shown in figure 1. This shape results from the dynamics of learning. When women are, on average, pessimistic about the cost of participating in the labor market, learning is very slow since the noisiness of the public signal swamps the information content given by differences in the proportion of women who would work in the two states of nature. As the beliefs about the welfare consequences of work become more moderate, the information in the signal improves. Once beliefs become sufficiently optimistic though, once again, the informational content in the public signal falls.

To evaluate the ability of the model to explain the quantitative evolution of female LFP, I first calibrated a version of it, in which beliefs did not evolve to a few key statistics for the year 2000. In this version, only changes in earnings over time can explain changes in female LFP. I showed 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 it to additional statistics from the last few decades of the century greatly improves its capacity to replicate the historical path of female LFP. The paper also examines various other testable implications of the model including its predictions for the historical evolution of elasticities and its cross-sectional predictions for the path of LFP for more vs less educated married women. On the whole, the model performs well.

The model indicates a novel role for changes in variables that affect the attractiveness of working (e.g. wages, policies, or technological change). 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 want to work less extreme, and thus render the public signal more informative. This implies that 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 married women's LFP, particularly over the period 1970-1990.

Since the model implies that exogenous variation in women's LFP should have intergenerational implications for the speed of learning and hence for future levels of female LFP, the paper also discusses using variation in the mobilization rate of men during World War II to obtain exogenous variation in LFP. The findings of Fernandez, Fogli, and Olivetti (2004) are shown to be consistent with those of the model: states with higher mobilization rates had more married women working in 1950 and in the following generation.

The model makes some simplifying assumptions, including an unchanged psychic cost of working over 120 years. It would not be difficult to incorporate changes in the cost structure, but without direct empirical evidence on the matter, an unchanged cost means fewer free parameters – a better disciplining device in a model with few data points. 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.[65] 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. 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. Introducing multiple states of nature or several work periods would considerably complicate the algebra but should not affect the main theoretical results.

The calibrated model finds that at the outset women were pessimistic about the true cost of working, consistent with the literature of that period. 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 or unions), 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.[66]

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. 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.[67] At a theoretical level, it would also be interesting to explore further the potential inefficiencies that arise due to learning externalities and to examine the possible role for policy. At the empirical level, it is important to depart from focusing 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.[68]

[65] See Hazan and Maoz (2002) or Munshi and Myaux (2006).

[66] 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 (2010), Edlund and Pande (2002), and Fernández (2010).

[67] 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. See also Mira (2007).

[68] 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 (2007) structurally estimates his model using Malaysian panel data.

8 Appendix

8.1 Data

8.1.1 Earnings

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.[69] As there is no data for earnings in 1880 and 1910, these points are constructed using a cubic approximation with the data from 1890 -1930 (inclusive).

To construct the earnings sample from 1940 onwards we used 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.[70] As is commonly done, observations that report weekly earnings less than a cutoff are excluded.[71]. Prior to 1980, individuals report earnings from the previous year, weeks worked last year, and hours worked last week. 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 education categories, college is defined as having at least one years of schooling after high school. High School is defined as having at most a high school degree or no more than 13 years of education.

Married Women's LFP

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. Since the individual census data is missing for this 1890, we use the midpoint between 1880 and

[69] 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.

[70] The sample is limited to full-time year-round workers because hourly wages are not reported. 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.

[71] 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. 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.

1900. We restricted our sample to married white women (with spouse present), between the ages of 25-44, born in the US, in non-agricultural occupations and living in non-farm, non-institutional quarters.

For the education categories, college is defined as having at least one years of schooling after high school. High School is defined as having at most a high school degree or no more than 13 years of education.

Elasticities

I take the elasticity estimates from Blau and Kahn (2006) who 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. 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 metropolitian 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). 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. 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).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.

Poll Data

I use the Gallup Poll data for years 1936, 1938, 1945 and 1970. From 1972 to 1998, I use data from the General Social Survey (GSS). The exact question varied somewhat over time. It was " Should a married woman earn money if she has a husband capable of supporting her? " in 1936; "Do you approve or disapprove of a married woman holding a job in business or industry if her husband is capable to support her?" in 1945; and " Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?" for the remaining years. Furthermore, in 1936, the possible responses were "Yes" versus "No", whereas in the later years the options included "Yes", "No" or "No Opinion" / "Don't know", depending on the year.

8.2 Calibration of the learning model

The model with no learning

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

$$\varepsilon_k = g\left(l^*\right)\frac{\partial l^*}{\partial w_k}\frac{w_k}{L} \tag{28}$$

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

$$\gamma = \frac{\log\left(1 - \frac{w_f}{w_h}\frac{\varepsilon_h}{\varepsilon_f}\right)}{\log\left(1 + \frac{w_f}{w_h}\right)} = 0.503 \tag{29}$$

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

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

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

$$\sigma_l = \frac{A}{\exp\left(\frac{\Phi^{-1}(L)^2}{2}\right)} = 2.29 \tag{30}$$

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

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 (29), and thus the same value of γ 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.[72] 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

$$l_{st} = \frac{\underline{l}_t + \bar{l}_t \exp\left(\lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma^2_\varepsilon}\right)(s - \bar{\beta})\right)}{1 + \exp\left(\lambda_t - \left(\frac{\beta_H - \beta_L}{\sigma^2_\varepsilon}\right)(s - \bar{\beta})\right)} \tag{31}$$

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

$$\begin{aligned} \Pr(DW_t|MW_{t-2}) &= \frac{\Pr(DW_t \text{ and } MW_{t-2})}{P(MW_{t-2})} \\ &= \frac{\int_{-\infty}^{\infty} \Pr(DW_t \text{ and } MW_{t-2}|s) f(s - \beta_L) ds}{L_{t-2}(\beta_L)} \\ &= \frac{\int_{-\infty}^{\infty} G(l_{st}) G(l_{s,t-2}) f(s - \beta_L) ds}{L_{t-2}(\beta_L)} \end{aligned} \tag{32}$$

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 (32), 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

$$\Pr(DW_t|MNW_{t-2}) = \frac{\int_{-\infty}^{\infty} G(l_{st})(1 - G(l_{s,t-2})) f(s - \beta_L) ds}{1 - L_{t-2}(\beta_L)} \tag{33}$$

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

$$R_t = \frac{\Pr(DW_t|MW_{t-2})}{\Pr(DW_t|MNW_{t-2})} \tag{34}$$

In order to estimate $\lambda_0, \sigma_\varepsilon, \sigma_\eta, \beta_H, \beta_L$, and σ_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 λ_0 ranging between [-10, -.01], σ_ε in [0.1, 5], σ_η in [0.01, 2], σ_l between

[72] 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.

[0.5, 4], β_L in [.01, 1], and β_H to be between [1, 10] units greater than β_L.

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

300 discrete types were assumed between $\underline{l}(w_h, w_f)$ and $\bar{l}(w_h, w_f)$ in each year to approximate the integral in equation 15. Then we average over the η shocks to determine the expected number of women working. We then back out the belief that would lead to exactly that many women working. This determines the path of beliefs.

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] Acemoglu, Daron, David Autor, and David Lyle, "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury," Journal of Political Economy, CXII (2004), 497–551.

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

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

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

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

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

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

[8] 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.

[9] 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.

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

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

[12] 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.

[13] 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.

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

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

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

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

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

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

[20] 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.

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

[22] 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.

[23] 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.

[24] 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.

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

[26] Fernández, Raquel and Alessandra Fogli (forthcoming), "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics.*

[27] 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.

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

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

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

[31] 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.

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

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

[34] 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.

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

[36] 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.

[37] 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.

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

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

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

[41] Haines, Michael, "Fertility and Mortality in the United States". EH.Net Encyclopedia, edited by Robert Whaples. March 19, 2008. URL http://eh.net/encyclopedia/article/haines.demography

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

[43] 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.

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

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

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

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

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

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

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

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

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

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

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

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

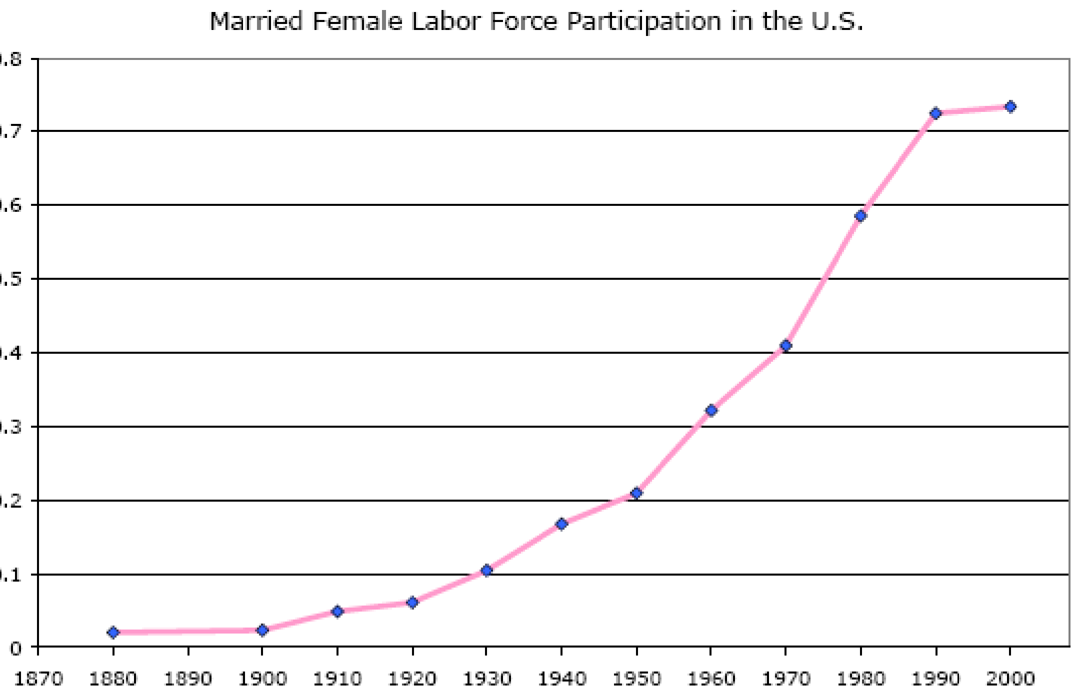

Figure 1: U.S. Census data 1880-2000. Percentage of white, married (spouse present) women born in the U.S., 25-44 years old (non-agricultural, non-group quarters), who report being in the labor force.

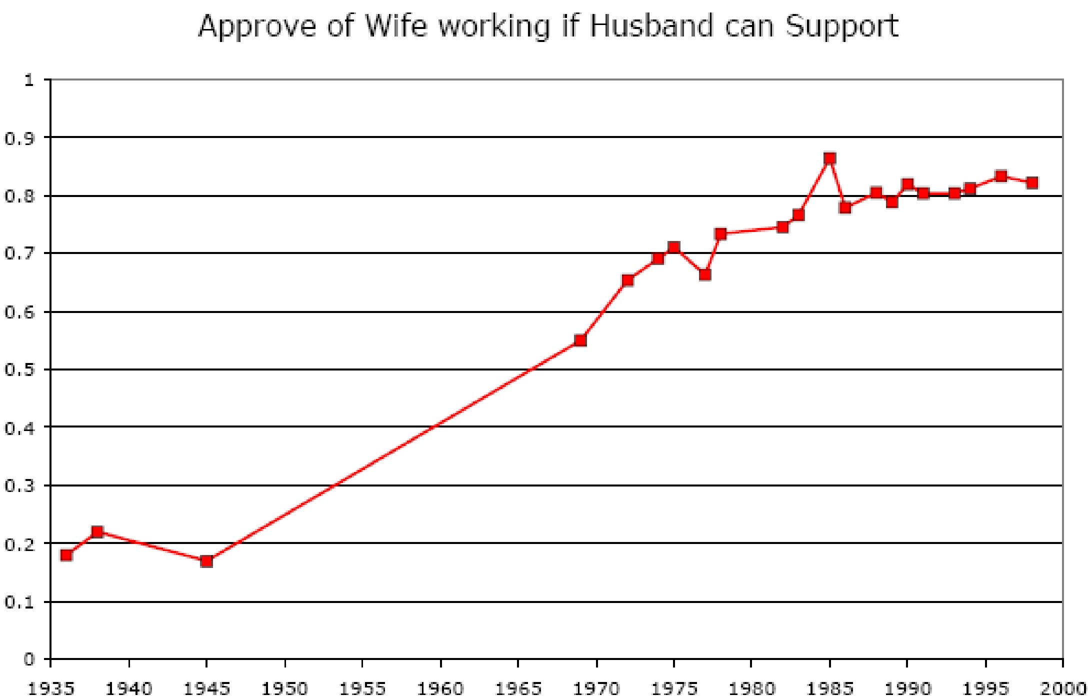

Figure 2: Sources: 1936-1938 and 1969 numbers are from the Gallup Poll (1972), 1945 is from Benjamin I. Page and Robert Y. Shapiro, The Rational Public, University of Chicago Press, 1992; pp. 101, 403-4. 1972 onwards are from the General Social Survey.

Figure 3: Timeline of Learning Model

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

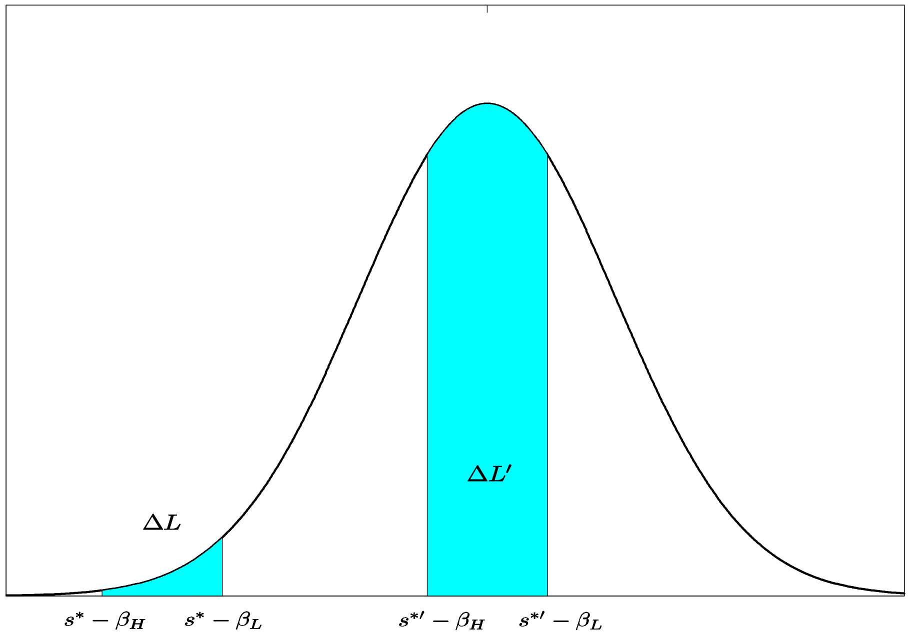

Figure 4: Normal PDF

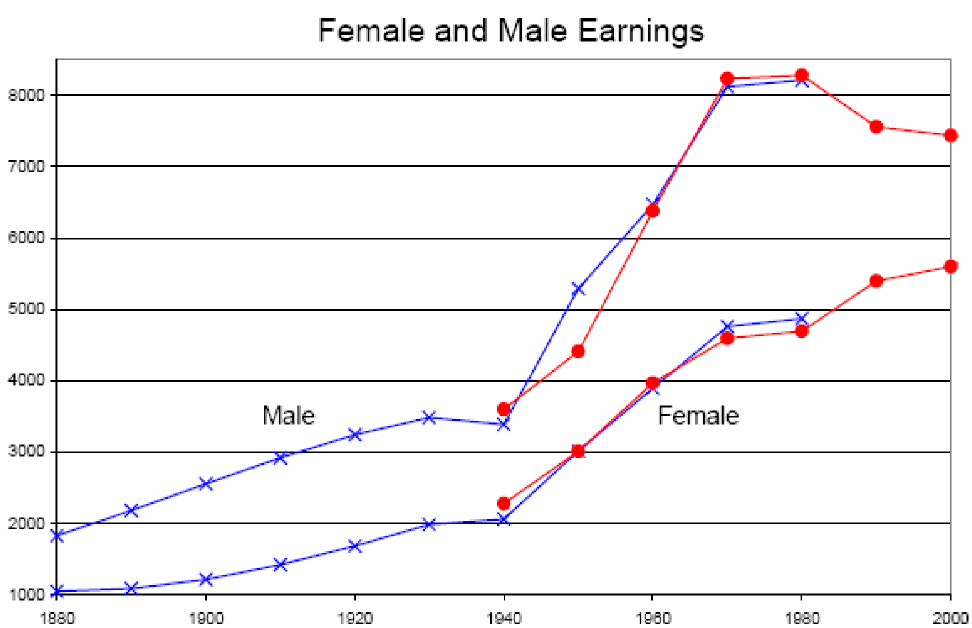

Figure 5: Crosses (blue) represent the yearly median earnings data from Goldin (1990), Table 5.1. Dots represent our calculations using U.S. Census data (red). They are the median earnings of white men and women between the ages of 25-44 in non-farm occupations and not living in group quarters. All earnings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

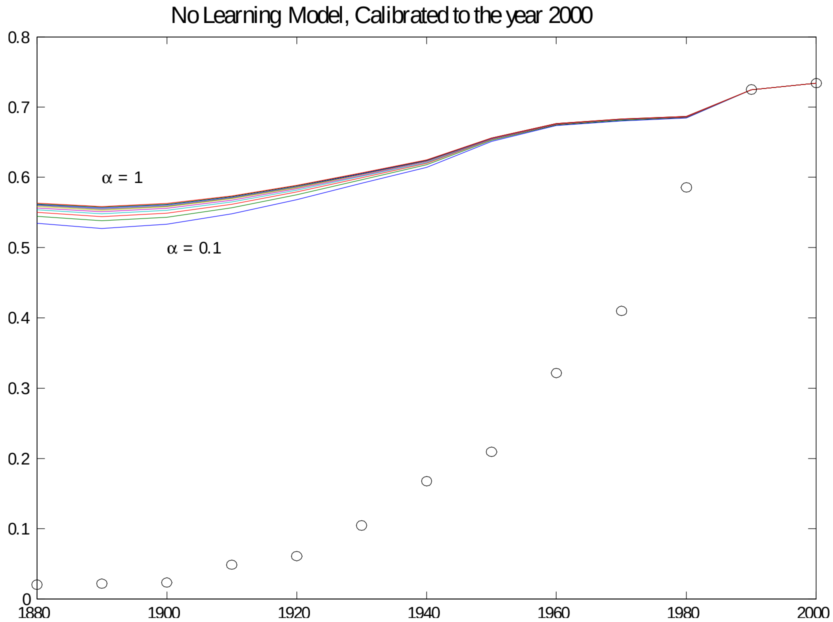

Figure 6: Parameters: $\gamma = 0.503$, $\beta = 0.321$, and $\sigma_L = 2.293$. α is the fraction of husband's earnings that enters a wife's utility via consumption.

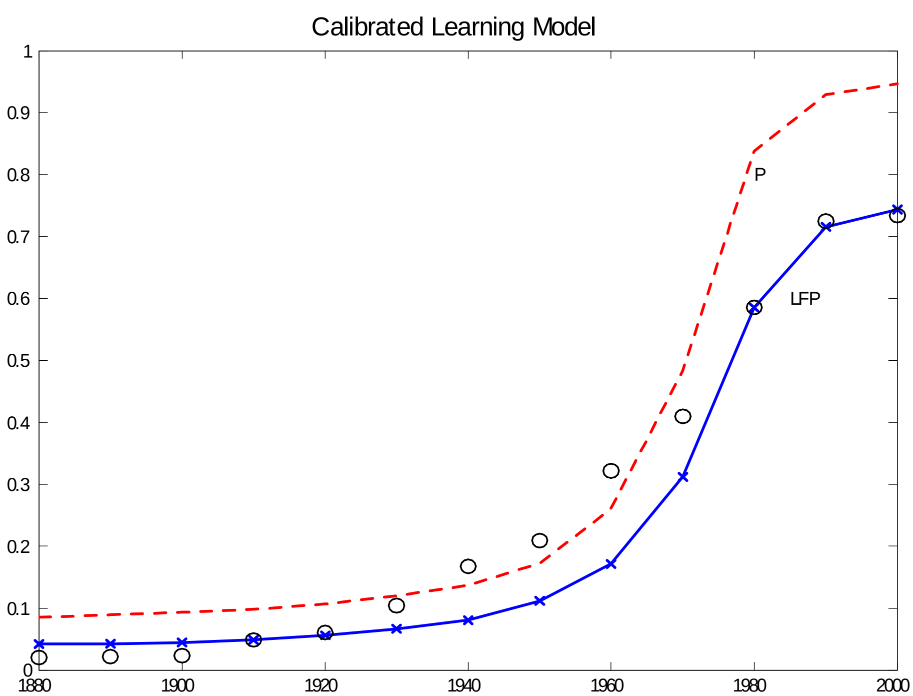

Figure 7: The dashed red line (p) is the belief path of the median individual. The sum of squared errors (distance of predicted LFP from actual LFP) is 0.052.

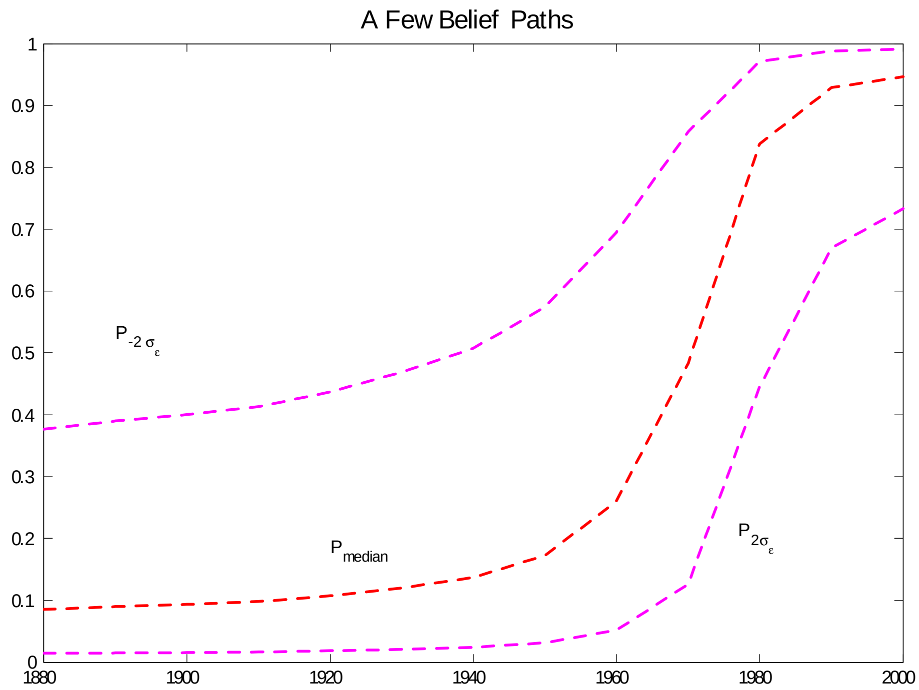

Figure 8: This shows $\Pr(\beta = \beta_L)$ for agents with $s = \beta$ and $s = \beta \pm 2\sigma_\varepsilon$.

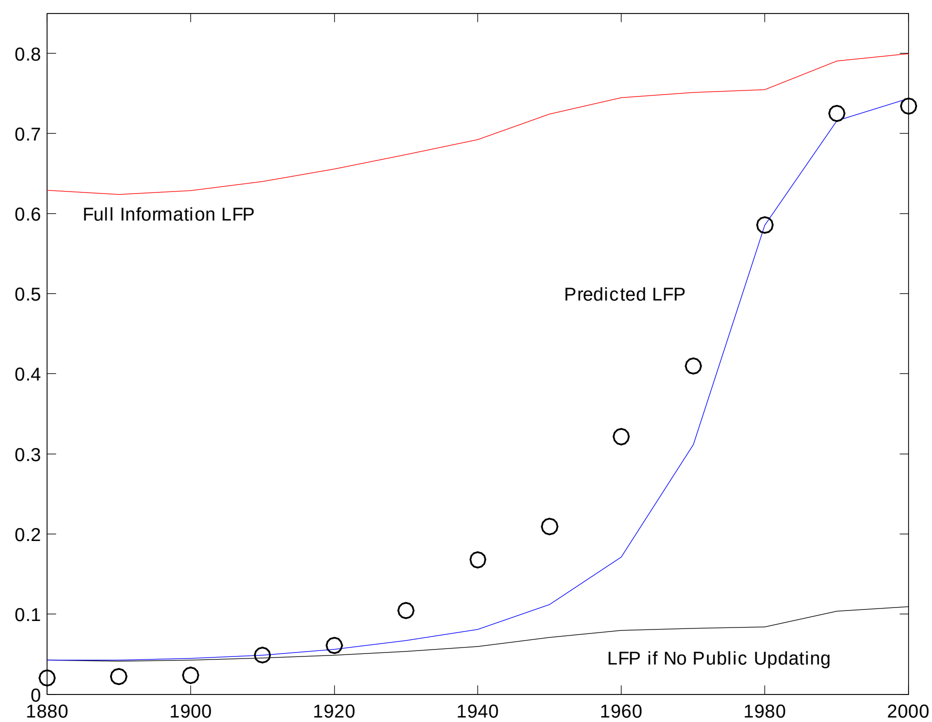

Figure 9: Uses the solution parameters from calibrated model but without public learning.

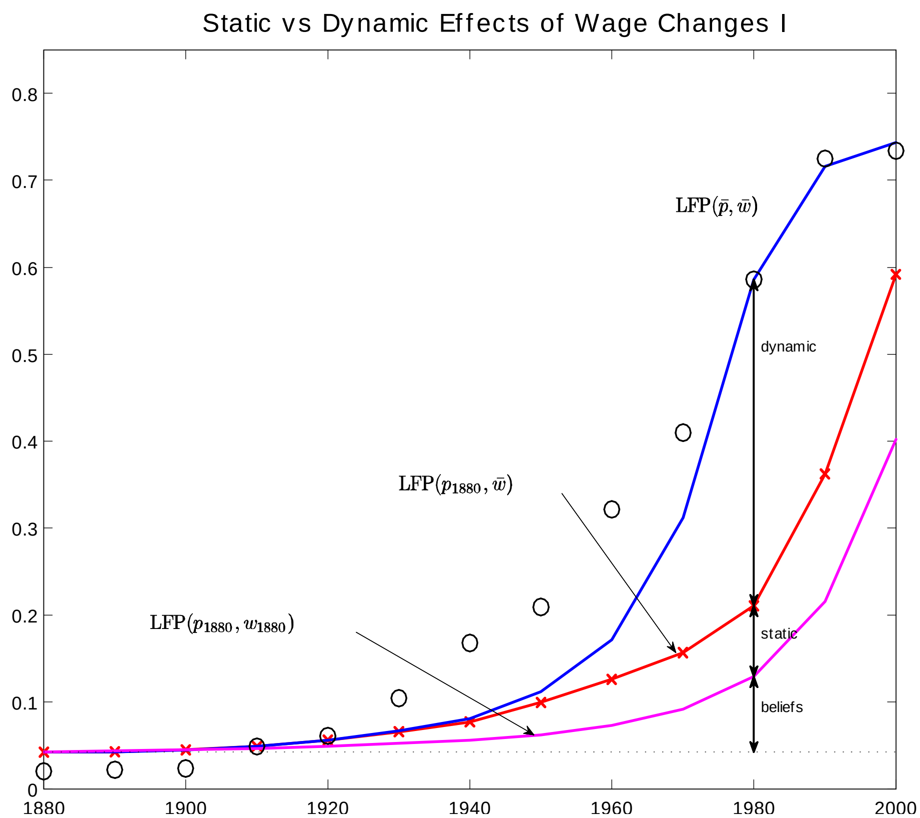

Figure 10: Decomposition of LFP. See the text for notation.

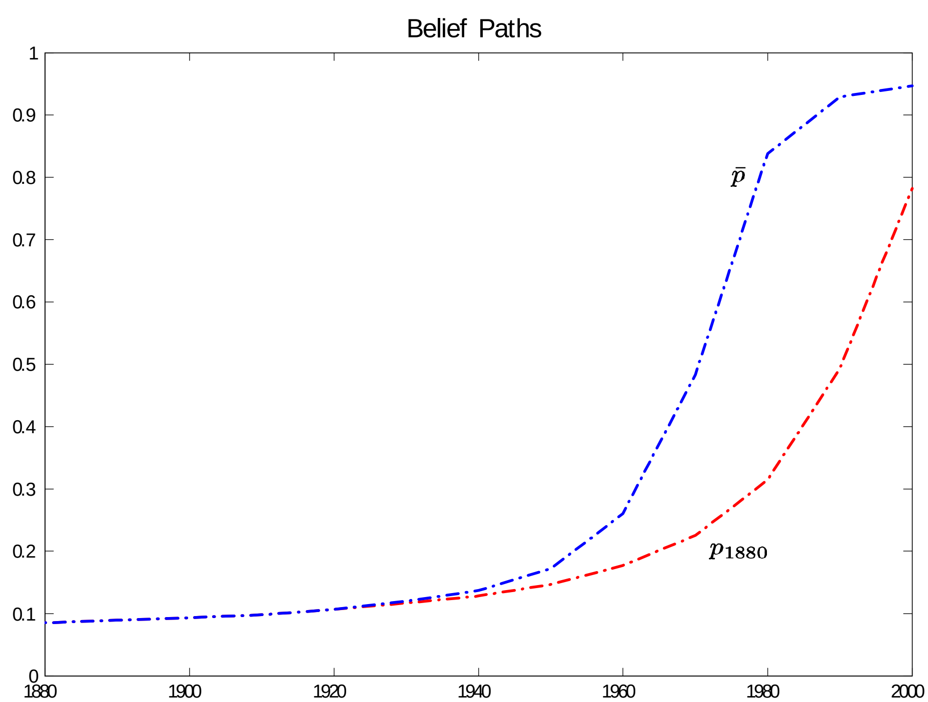

Figure 11: $P(\beta = \beta_L)$ for historical earnings series and for earnings constant at the 1880 levels.

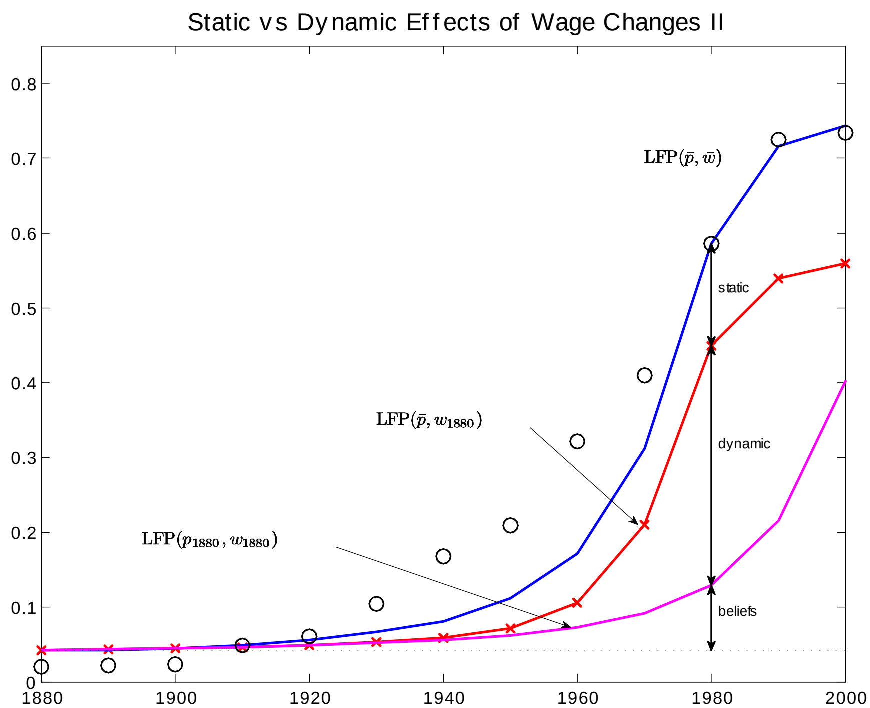

Figure 12: Alternative decomposition of LFP.