THE SUPPLY OF GENDER STEREOTYPES

AND DISCRIMINATORY BELIEFS

by

Edward L. Glaeser[1]

Harvard University and NBER

and

Yueran Ma

Harvard University

Revised Draft, December 2013

Abstract

What determines beliefs about the ability and appropriate role of women? An overwhelming majority of men and women born early in the twentieth century thought women should not work; a majority now believes that work is appropriate for both genders. To explain this change, we present a model where parents perpetuate beliefs out of a desire to encourage the production of grandchildren. Undersupply of female education will encourage daughters' fertility, directly by reducing the opportunity cost of their time and indirectly by leading daughters to believe that they are less capable. Children will be particularly susceptible to persuasion if they overestimate their parents' altruism towards themselves. The supply of persuasion will diminish if women work before child-bearing, which may explain why gender-related beliefs changed radically among generations born in the 1940s.

I. Introduction

Why do gender-related beliefs emerge and shift over time? Changes in these beliefs may have played a role in the secular changes in female labor force participation discussed by Goldin (1990) and Olivetti (this volume). According to the General Social Survey, 47 percent of women

born before 1946 (and 59 percent of men) agree with the statement "It is much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family."[2]   Only 29 percent of women born after 1945 share that view. These perceptions not only vary over time, but also across regions. A full 50 percent of female respondents (from all cohorts) agree with that statement in the West South Central Region, while only 26 percent of New Englanders share the view.

We have less survey evidence on discriminatory beliefs about women's ability in the workforce than we do about women's "proper" role in the home.   Nonetheless, the evidence that does exist also suggests dramatic transformations about beliefs about women's capacity in the workplace during the late twentieth century.   In 1953, Gallup asked "If you were taking a new job and had your choice of a boss, would you prefer to work for a man or for a woman?" In the 1953, 57 percent of women and 79 percent of men expressed a preference for a male boss, as opposed to only eight percent of women and two percent of men who expressed a preference for a female boss.   By 1987, the share of female and male respondents expressing a preference for a male boss had dropped to 37 and 29 percent respectively, with men now preferring a female boss (Simon and Landis, 1989).

Moreover, an abundance of personal histories, ethnographic work and field-specific statistical research suggests that men and often women as well, have often believed that women are less capable in many workplace relevant tasks (e.g. Lerner, 1987).   The literature on women and perceived math ability is voluminous, and suggests that men and women often believe that women are less able in mathematics (see Gunderson et al. 2012).   The women who pioneered their way up corporate ladders have often described a common male presumption that their talents were limited.   Major thinkers from Aristotle to Freud have often depicted women as severely lacking in vital decision-making areas.

Section II further discusses the survey, ethnographic and literary sources that attest to the existence of patriarchal, discriminatory beliefs against women at various points in history.   This section also argues that these gender-related stereotypes cannot be understood as a purely Bayesian response to commonly available facts, but that they are instead a product of persuasion. For example, the surveys discussed above are taken in the same year, by respondents who observe the same labor markets, and yet respondents born before and after 1945 have markedly different opinions about the working women, suggesting that an impact of upbringing on beliefs is far stronger that it should be in a perfectly rational world.

Our view complements Goldin (this volume) who argues that discrimination against women in previously male jobs may reflect some aspects of reality.   We do not mean to suggest that reality is irrelevant, but rather that there are many cases in which beliefs about women do not

---

[2] To increase the sample size, we combine the waves of 2003, 2004 and 2007 of the General Social Survey to compute these statistics.

correspond to reality. Instead, as in Glaeser (2005), we assume that beliefs reflect persuasion rather than reality, and we focus on the supply of persuasion.

To understand the supply of erroneous beliefs, we must understand the incentives to spread falsehood. After discussing several possible alternative sources in section III, we focus our attention on parents. Parents with a strong preference for own grandchildren will have an interest in persuading daughters to forgo work in the formal labor force (Gunderson et al. 2012). Moreover, parents have far greater resources available with which to influence the beliefs of their children than do co-workers, spouses or other possible sources of beliefs. Parents have some control over children's time and experiences for many years, during periods where children are less likely to have strong alternative sources of information. This combination of incentives and power leads us to believe that parents are a primary source of gender stereotypes and we model that process in Sections IV and V.

Section IV presents our core model on the parental formation of beliefs for female children. Our model follows standard economic assumptions and links the persuasion process to a Bayesian signaling model. Parents can send costly signals, including altering the education of their children or their own workplace behavior, which may shape children's beliefs, either about their own ability or about the ability of women as a whole. While the model uses the word "ability" to describe the source of uncertainty, it could equivalently be interpreted as the psychic returns from working and child-rearing, so the model can be interpreted as describing the perpetuation of traditional values.

We focus on differential education choices by gender. If young women believe that parents' have access to private information about their children's ability, then choices about educational investment will be seen as a meaningful signal about their own ability. We first focus on women's beliefs about their own ability in Section IV, but then discuss how the model would shape societal beliefs in Section V.

In the model, parents are altruistic towards their children but they have an independent desire to have more grandchildren. This desire creates an incentive for them to try to generate beliefs that lead to more child-bearing. If education increases the returns from working in the labor force relative to child-bearing, this will generate lower levels of women's education, even if women know their ability levels with certainty. The under-provision of education effect becomes more pronounced if parents, but not their daughters, have private information about the ability of their own daughters or of women generally. Parents of skilled daughters may have an incentive to try to imitate parents of less able children by giving them less education, which may persuade daughters that their own time is best spent in child-bearing. If daughters have rational beliefs, this will cause more able women to think that they are merely average, but will not lead to any aggregate misperception about women's ability.

In Section V, we turn to three extensions of the model. We first discuss the ways in which parental choices may alter societal beliefs, of both men and women. In a rational model that leads to a separating or semi-pooling equilibrium, mistaken stereotypes get quickly undone. If the outcome in the model is pooling, then these stereotypes persist. The impact of parental persuasion will be particularly strong, and have more extreme consequences, if children are credulous Bayesians who make the understandable error of overestimating their parents' altruism (Glaeser and Sunstein, 2009). Trusting their parents too much leads daughters to underestimate their parents' incentive to act strategically. This tendency will heighten the parents' incentive to behave in a strategic manner, by under-investing in education.

At the end of Section V, we discuss the timing of work and childrearing, drawing on Goldin and Katz (2000). In this model, women (of varied educational attainment) choose when to schedule a continuous term of home production for child rearing, either early or late in their life cycle. The critical implication is that parental investment in misinformation makes sense when women have children early but not late. This fact implies that the shifts in the timing of women's child-bearing should have had a major effect on the supply of gender stereotypes. Over the long run, technologies such as the pill, which allows women to control the timing of fertility, may have reduced the incentive to persuade daughters that their time is better spent bearing children.

Section VI concludes and discusses the interplay between sources of incorrect information and real world experience. Working before child-bearing means that there is enough information to counteract persuasion. In a similar fashion, gender-related quotas that limit the number of women on the job seem unlikely to persist in the same way as glass-ceilings that prevent women from rising above a certain level. Gender related quotas should be unstable, if they are sustained with incorrect beliefs, because the few women hired for the job end up providing information that counteracts false beliefs. Glass ceilings, by contrast, provide no such evidence, which allow false beliefs to persist and maintain the incentives to perpetuate such beliefs.

II.   Discrimination and the Social Formation of Beliefs

We have a great deal of information about the relative productivity of men and women in the household, the availability of market-provided household services, and perceived workplace discrimination against women (e.g. Goldin, 1990, Blau et al. forthcoming). We have less evidence on beliefs about female competence. Perhaps this dearth of information is understandable. In the very recent past in the United States we would hardly expect many respondents to honestly admit to thinking that women are less capable. Nonetheless, the relative absence of polling data about female competence makes it difficult to fully document shifts in beliefs about women and their capacities.

There is however a great deal of anecdotal evidence suggesting that women have often faced strong belief-related barriers to employment. Men have often held strong opinions that women

were just not up to certain jobs. Often, these beliefs have crumbled in the face of reality, but certainly some of these beliefs persist even today.

Attitudes towards Women and Work

In this subsection, we briefly review the polling data that are available about gender stereotypes from the General Social Survey (GSS) and other sources. The General Social Survey, and other surveyors, has been asking questions about traditional gender roles since the early 1970s. Unfortunately, these gender-role related questions do not map neatly into any particular taste or belief. A patriarchal viewpoint can reflect a higher opinion of female productivity in the household sector, or a belief that employers discriminate unfairly against women.

Figure 1, for example, shows the average responses to the question "It is much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family" by birth year for men and women separately. The graph shows a strong downward pattern for both men and women. For cohorts born at the start of the twentieth century, almost all men and women thought that traditional gender roles were best.   The share of respondents sharing that view declines to about 30 percent by 1950 and then levels off.   There are some odd positive upticks in the responses to the question in the most recent cohorts, but this may reflect measurement error. The basic pattern documents a profound change across cohorts born in the first half of the twentieth century, and this pattern presents itself during every year in which the survey question was asked. [3]

The second figure shows a similar response to the GSS question asking whether mothers' working outside the home is harmful or harmless for young children. Again, cohorts born at the start of the twentieth century almost uniformly believed that children were hurt by women working outside the home.   By 1960, almost half of respondents did not state this belief. Even though an overwhelming majority of respondents are comfortable with women working as a general matter, some still say that working while children are young is harmful.

There are far fewer questions that seem to directly capture assessments of female competence, and most that are relevant concern very particular tasks or occupations. The General Social Survey asks two highly specialized questions for particular years that would seem to relate to female competence.

In 1974 and 1982, the survey asked if men make better political leaders. The cohort pattern, shown in Figure 3, displays a clear change over time. About 40 percent of people born earlier in

---

[3]There is no survey evidence for cohorts born in the nineteenth century United States. Olivetti (this volume) documents that some, but not all countries, experienced a U-shaped female labor force participation pattern. It would be interesting to know whether attitudes in these countries towards women participating in the workforce roughly track the time series of female labor force participation.

the twentieth century think that men make better political leaders. By the latter decades of the century, this belief is down to 20 percent. We cannot generalize from political competence to competence in the workplace, but the effects are still quite striking.

A second question that is potentially related to ability was asked in 1996.   Men and women were both asked if women earn less than men because they work less hard.   This question about female work effort shows a striking non-linearity (shown in Figure 4), where beliefs about greater male effort decline with year of birth during the first half of the twentieth century and then a rise after that date.   We have no satisfactory explanation for this pattern, but it does suggest that cohort does have an impact on these beliefs.

The Social Formation of Beliefs

Why do discriminatory beliefs differ radically over groups and across time?   The economics of discrimination began when Gary Becker (1957) presented a model based on the preferences of employers, customers and fellow workers.   Becker's approach posits that some members of one group dislike working with or buying from members of another group.   The Becker model describes the reality of the mid-1950s, and provides many keen insights, like the negative impact on profits generated by an employer's discriminatory tastes.[4]

Even if whites had no innate dislike of blacks and men were willing to work with women, members of one group might still benefit if they were able to coordinate to expropriate the rights of another group (Krueger, 1963, Thurow, 1969), or if there was a society-wide equilibrium that restricts the choices of a disadvantaged group (Akerlof, 1976).[5]   The South's Jim Crow system was not merely the decentralized preferences or beliefs of ordinary people. It was socially and legally organized, and seems in many contexts to have generated transfers from blacks to whites. Those transfers were perhaps most obvious in the case of segregated schools, which allowed tax dollars to be spent far more heavily on white, rather than black children, especially when blacks were particularly immobile (Margo, 1991).

These models certainly fit many aspects of the Jim Crow south, and they may also reflect some forms of gender-based discrimination as well. As Myrdal (1944) discussed in his classic study of American segregation, integration-oriented whites were no more allowed to travel in black railcars than blacks were allowed to travel white cars.   Firms proudly trumpeted their whites only policies, and the system only changed with massive legal intervention from the Federal government, which can be seen as breaking the old equilibrium with outside force.   Margo

---

[4] Lazear's (1999) model of culture and language provides a complementary communication-based explanation for some forms of discrimination in the labor market. Different cultures, or ways of speaking, can make coordination difficult and lead to lower productivity.

[5] Akerlof (1976) presents a model where a caste system, such as the Jim Crow south, was an unfortunate but stable equilibrium that reflected a society wide rule where members of one clique are punished for interacting with members of a second clique.

(1991) predicts that centralized discriminatory behavior would start to change as blacks could move north and indeed that seems to have happened.

It is less clear that there was an organized conspiracy against women in the mid-twentieth century, that was similar to the Jim Crow system in the south, or that the legal pressure exerted by the Equal Pay Act of 1963 or the Civil Rights Act of 1964 had the same cathartic impact for women that it did for African-Americans. Moreover, neither centralized discrimination models nor the Becker taste-based discrimination model can explain the changing nature of views towards African-Americans and women, because they were not intended to make beliefs or preferences endogenous.[6] In centralized discrimination models, members of the ruling clique rationally respond to incentives, and have neither negative opinions nor disproportionate ill will towards either women or minorities.

Arrow's statistical discrimination model (1973) provides an alternative model that can explain discriminatory hiring practices and beliefs. The model suggests that employers and ordinary people have a low opinion of certain groups and these low opinions lead to discriminatory behavior. Certainly, it appears to be the case that at various times employers have held a low opinion of the competence of both blacks and women. Indeed, Goldin (this volume) is closely related to the statistical discrimination theory suggesting that opposition to women in particular jobs is based on an assessment of female ability generally, which may be lower than the ability level in an occupation at the time.

However, the great challenge of statistical discrimination models is that they typically also assume that people are fairly rational in their belief formation. This implies that attitudes need to be tethered to reality. Yet it is difficult to accept that there was much evidence to suggest that either women or blacks were as inept as many mid-century employers appear to have thought. Previous work (Glaeser, 2005) focusing on beliefs about malevolence (rather than competence) emphasized that while Southern voters a century ago seem to have been convinced that African-Americans were a great threat to their safety, it was whites, not blacks, who had systematically enslaved, brutalized, sexually assaulted and even killed members of the other group. It is harder to document the error in beliefs about competence, but it seems quite likely that many people had beliefs about women and minorities that were not based on any real evidence and that bore little resemblance to the truth.

If beliefs about blacks and women systematically differed from reality, it becomes necessary to focus on theories that can generate widespread divergence between the truth and beliefs. There are at least two well-known systematic biases that can potentially generate such beliefs internally, without any external persuasion: the fundamental attribution error and self-serving

---

[6] Subsequent work by Becker and Murphy (2000) makes preferences endogenous and this paper is strongly indebted to their work. Our decision to focus on belief rather than preference formation reflects our own preference for the greater discipline created by belief-formation models, as in Section V, that require at least a partial Bayesian framework. In the case of the model in Section IV, results would be identical if we allowed preference formation.

biases. If the fundamental attribution error leads observers to associate the negative outcomes of others with intrinsic personal characteristics, rather than external constraints, then individuals could readily believe that poor labor market outcomes for either blacks or women represent low levels of innate ability rather than discrimination. Self-serving biases, which lead people to prefer views that make them see themselves in a positive light, could also lead white men to have negative views of blacks and women, because such views prop up white self-esteem.

While these behavioral quirks may have contributed to negative assessment of blacks and women, there are limits to the power of these theories. For example, women's own belief in gender stereotypes, discussed in the previous subsection, cannot be the result of self-serving biases, since the beliefs do not seem to be self-serving.   Moreover, the fundamental attribution error suggests that adverse outcomes for others are attributed to intrinsic factors, but that personal disappointments are blamed on external constraints. Yet women themselves often seem to share patriarchal beliefs.

Here, we focus on the social formation of error, and our critical assumption is that human beings are sensitive to social persuasion. In the discussion and two models that follow, individuals will be reasonably rational, but they will not totally discount falsely generated signals about the characteristics of out-groups.

On one level, the social formation of error runs against a long-standing tendency of economists to assume a high level of rationality and even accuracy in beliefs. Yet if we accept that mid-twentieth century white males had erroneous opinions of the ability levels of blacks and women, we must consider at least the possibility that some beliefs have little basis in reality.   While our approach runs against the economist's predilection for hyper-rationality, it fully embraces the role that incentives can play in the generation of all sorts of outcomes, including incorrect beliefs.

Naturally, those incentives must battle against the incentives of listeners to learn the truth. In the political context, those incentives may be quite weak. After all, no individual voter has a strong incentive to ascertain the truth about any particular story, if the truth will only serve to make his or her vote a bit wiser.   In the labor force context, those incentives may be quite stronger.

Moreover, we will assume that widely spread falsehoods will not persist if there is obvious evidence to the contrary.   In any sensible learning model, this fact will suggest that racial or sex-based quotas are not typically stable, while glass ceilings may be.   The existence of a glass ceiling towards women (or perhaps a low dark roof for blacks in the Jim Crow South) ensures that there is no hard evidence on how women or blacks can perform in higher positions. The absence of information allows incorrect beliefs to persist.

Discrimination vs. Hatred

These models also help us to distinguish discrimination from hatred. Hatred is modeled as a belief that an out-group is malevolent, and prone to engage in harmful behavior if they are empowered.   Discrimination is a belief that an out-group is different and perhaps less capable, but not necessarily harmful or malign.   Hatred leads to policies such as segregation and genocide, as in-groups attempt to shield themselves from the perceived threat. Discrimination will lead to different hiring practices and perhaps even exclusion from political decision-making. Yet policies based on beliefs about lesser ability levels will not attempt to explicitly harm the out-group, because the out-group is not perceived as dangerous.   While we might wish to harm people who are perceived to be malevolent, before they harm us, we have little incentive to attack people who are perceived as less able. .

Historically, African-Americans have suffered from both discrimination and hatred. They have been perceived as being less competent, and they have also been perceived as being a threat. These beliefs were able to persist, arguably, because blacks were excluded from positions where they might do (perceived) harm and kept out of jobs where they could have demonstrated ability.

Women have suffered from discrimination but not typically from hatred.   The primary experience of extraordinary altruism in the lives of most men is the self-sacrificing behavior of their own mothers, which would make it hard to accept that women are somehow naturally malevolent.   Indeed, many of the most profound opponents of women in the workplace or in politics, who certainly subscribe and even promulgate views about female competence, have also held up women as the fairer sex that is more generous and good-hearted than men.   When Senator Vest of Missouri opposed women's suffrage in 1887, he said "I believe that [women] are better than men, but I do not believe they are adapted to the political work of this world."

It is historically rare for out-groups to be simultaneously depicted as malign and incompetent. Indeed, such views would be counter-productive if a hate-producer is looking to generate support for policies that are harmful to the out-group.   If a group is incompetent, then it is less threatening and that would mean less need to engage in defensive mechanisms. Jews, for example, have historically been depicted as both malign and powerful, which together justified the use of extreme anti-Semitic policies.   The Soviet Union was depicted as an Evil Empire, which called for massive U.S. military spending at the time. If the Soviet Union was merely an evil bumbling bureaucracy (arguably an accurate description during the Reagan era) then there would have been far less need for military spending.

In the case of patriarchic beliefs, it is possible to conceptually distinguish beliefs about ability and societal norms. A woman, for example, might stay in the home because she believes that her workplace productivity is relatively low in comparison to her productivity in the household. Alternatively, she may believe that staying home is just the "right" thing to do.   But while these two notions may differ in some deep sense as a practical matter they are indistinguishable.. There is a conceptual distinction between believing that women are less able in the workplace or more

able in the household sector, and surely both beliefs have existed, but when it comes to time allocation decisions the beliefs are interchangeable.

III.   The Entrepreneurs of Error: Sources of Sexism

If common beliefs are socially formed, then they are unlikely to be produced by accidents. Instead, interested individuals must have incentives to spread falsehood.   In this section, we discuss various potential sources of misinformation about female ability levels, and explain our decision to focus on one particular source, parents. We focus on cases where spreading misinformation is intentional and instrumental.   There have certainly been countless instances where politicians, for example, have uttered gender stereotypes, but most of the time, this seems more likely to reflect pre-existing norms, rather than any conscious political strategy.   We therefore look for a setting where someone with the power to persuade also has the motive to depict women as either less capable or more suited for work outside the labor market.

Political Entrepreneurs

In Glaeser (2005), political entrepreneurs spread hatred against an out-group because hatred complemented the policies proposed by those politicians.   The model suggests that a steady supply of erroneous beliefs requires low costs of widespread persuasion, persistent policy differences between parties that disproportionately impacted an out-group, political weakness of the out-group and the relative segregation of that out-group to reduce alternative sources of information.

These conditions are far less likely to hold for women, than for Blacks or Jews, and they certainly do not hold in the more distant past, when we believe gender-related beliefs were already strong. Cheap political persuasion, outside of cities, requires both the printing press and voter literacy.   For this reason, politically-induced hatred of groups, as opposed to religiously induced hatred, appears to have been a largely nineteenth century innovation.

Two prominent gender-related issues emerged in U.S. politics during the nineteenth century: female suffrage and temperance (eventually, prohibition).   Prominent leaders in female suffrage, like Elizabeth Cady Stanton, also led temperance organizations and prominent temperance leaders, like Frances Willard, were also suffragists.   Prohibition was partially justified as a policy that would protect wives and children from abusive, drunken husbands, and suffrage was justified as the means of passing prohibition.   The rise of gender-related issues made it possible at least that they opponents of these issues would have turned to sexism, just as the opponent of policies that granted modest aid to Blacks or Jews turned to racism and anti-Semitism. Beliefs about female competence would be particularly relevant to the issue of female suffrage, and arguments about female incapacity were routinely made by the opponents of suffrage.

Yet even in these cases, the political language was limited perhaps because the parties never split decisively on suffrage, and politicians had far less chance of changing male beliefs about women, than they did of conjuring the fear of a race riot. The early connection between these issues and abolitionism (Fogel, 2000) may have made them a more natural fit for the Republican Party, and Republicans were stronger supporters of the bills that eventually led to the nineteenth Amendment, but neither issue became a major party plank until 1916, when both platforms supported extending voting rights to women.[7] Neither party endorsed prohibition before the passage of the eighteenth Amendment. By 1916, a large number of states already allowed female suffrage, especially in Presidential elections, making it politically unwise to insult a large voting bloc.

Moreover, since the nineteenth Amendment was passed, women have gone from being politically absent to the second largest and now the largest voting bloc. While there has been plenty of vilification on both sides of the debate on abortion rights, the suggestion that abortion limitations are justified by broad limits on female decision-making ability has been fairly rare (suggestions that teenage girls are incapable of making wise decisions are more common), presumably because telling a majority of voters that they are stupid (or evil) would seem to be immense electoral folly.[8]

Opponents of the Equal Rights Amendment (ERA) were more likely voice their opposition to unnecessary Federal regulation, rather than to say that discrimination was broadly justified on ability related grounds. The Republican Platform of 1980 affirmed "our Party's historic commitment to equal rights and equality for women," and supported "equal opportunities for women," but also claimed that "states have a constitutional right to accept or reject a constitutional amendment without federal interference or pressure," and that Federal "pressure against states which refused to ratify ERA" must cease. Phyllis Schlafly was the most prominent political entrepreneur opposed to the amendment, and she based her opposition both on a defense of traditional family structure and by claiming that the amendment would strip women of traditional privileges, such as avoiding the draft. While there have been instances where politicians do seem to have actively promoted gender stereotypes, particularly around the issue of female suffrage, this seems to have been a relatively minor phenomenon, at least relative to the spread of stereotypes by other actors.

Market Entrepreneurs

---

[7] In 1916, the Republican Platform "favors the extension of the suffrage to women, but recognizes the right of each state to settle this question for itself," while the Democrats "recommend the extension of the franchise to the women of the country by the States upon the same terms as to men." The Republicans are endorsing suffrage, but not an amendment to force it on unwilling states, while it is unclear if the Democrats are supporting such an amendment or not. In 1872, the Republican Platform provided the amorphous words "The Republican party is mindful of its obligations to the loyal women of America for their noble devotion to the cause of freedom " and "the honest demand of any class of citizens for additional rights should be treated with respectful consideration."

[8] Democrats do, of course, assert that Republicans are waging a "war on women," a charge that Republicans hotly deny.

A belief that women are less capable in the market place has one obvious beneficiary: competing male co-workers, as suggested by Goldin (this volume). This would suggest that men should have the incentive to spread the idea that women are less competent. Within a corporate hierarchy, presumably the sensible strategy would be to emphasize the limits of a particular woman. In other settings, where no single female competitor exists, then it may make more sense to disparage women more broadly.

Spreading false beliefs will be more common when women really are a potential threat, and this means that we can make sense of the rise and female of discrimination in certain jobs that is discussed by Goldin (2000). During the early twentieth century, the threat of a female competitor was small and this meant that men spent little effort on persuading prospective bosses not to hire women. During the mid- twentieth century, the threat became more obvious and men began to persuade more assiduously. At the end of the twentieth century, there were enough positive examples of women working that misinformation had much less effect.

Several factors would be necessary for this persuasion to represent a dominant force. First, people making hiring decisions would need to be susceptible to persuasion from the subordinates who will compete with the new employee. This is not inconceivable—deans, for example, are quite reliant on faculty members when hiring—and junior faculty members are often allowed to weigh in on junior faculty hires. This process does suggest that persuasion would be occupation specific. It may be possible to persuade a superior that one's particular task (mathematics, construction work) requires male attributes, but it is unlikely to be as easy or as sensible to try to persuade the superior that women are less capable at all workplace tasks. However, if women are accepted as being less able in enough occupations, presumably the natural inference is that there is something more general at play..

Second, the persuaders would need to solve the free rider problem. No single worker has much of an incentive to persuade. The propagation of these beliefs would be therefore be more likely in small firm settings, or in cases where other organizations exist to collectively represent the interests of male workers. For example, in 1941, the United Auto Workers (UAW) filed a strike against the Kelsey-Hayes wheel plant, demanding "the removal of girl employees from machine work" (Milkman, 1982). But while the UAW might demand segregation-by-job in particular plants, and would regularly fight for equal pay provisions that reduced the possibility of men losing jobs to lower cost female employees, the union was far more interested in representing female employees than disparaging them.

Third, if beliefs have some connection to evidence and Bayesian reasoning, then discriminatory beliefs in the workforce can only persist when there is no evidence to the contrary, which is true even if beliefs come from other sources. Hard discriminatory barriers, justified by these beliefs, may be able to persist, while quotas, based on incorrect beliefs, seem unlikely to be stable. Many have argued that women working at typically male jobs during World War II helped dispel the idea that they were incapable of doing these typically male activities. The relative durability

of glass ceilings may be connected with the formation of beliefs, because they ensure that there is no direct evidence on upper level administrators in one particular company, and advocates of discrimination can more plausibly argue that upper level jobs are more heterogeneous across firms than lower level jobs. That heterogeneity makes it easier to deny the relevance of female achievements in other firms.

Individual workers might disparage women, and unions might occasionally strike against female employment, but overall discriminatory beliefs spread by co-workers do not appear to have been a major force, presumably because of the relatively weak incentives and limited ability for workers to spread discriminatory beliefs to employers. Industrialists have every incentive to see through male claims about female incompetence and look for low cost labor, as Lowell did when he started his textile mills almost two centuries ago. While co-workers may have served as an occasional source of discriminatory beliefs, they are unlikely to be a significant force, especially in more traditional societies.

The alternative market entrepreneur who has an incentive to promulgate gender stereotypes is the consumer goods company. Friedan (1963) is the primary proponent of this point of view.. For the Friedan argument to be persuasive the industry must be oligopolistic, consumer goods must strongly complement not substitute for women's time at home, and the costs of persuasion must be low. It is possible that these conditions existed when Friedan (1963), although they seem unlikely to hold today. Many important home products—the dishwasher, pre-made meals— substitute rather than complement time spent in the home, suggesting that their sellers should have been advocates of women working, not the opposite. There is little doubt that magazines and advertisements provided many examples of the joys of home-making, but the instrumental aim of those examples seems far more likely to generate positive associations for using a particular product. Even a washing machine company has the incentive to show a happy woman at home with her washing machine, not because the company wants to her to stay at home, but rather because it wants her to think about how wonderful having a washing machine can be.

Family Entrepreneurs

The long history of patriarchal attitudes, before mass media, before widespread democracy, before even the possibility of significant female integration into the workforce, suggests that these attitudes ultimately have a deeper source. Perhaps the deepest source of all is the family or clan itself, and ancient institutions, such as the church, that are often allied with adults in the family. If patriarchic views are common, if not ubiquitous, then it seems reasonable to believe that they are delivered for deep reasons and there is no deeper motivation than the perpetuation of the gene pool.

A particularly natural reason for supplying patriarchic beliefs is that these beliefs increase child-bearing. Fertility is typically seen as a complement towards being in the home and substitute with being away from home. Children typically need child-care and that is typically most

cheaply provided at home. Multiple pregnancies are often more difficult for working mothers to fit into their schedules. Given that fathers always bear far less of the costs of pregnancies and often been far less of the cost of child-rearing, empowering men within the household may also lead to higher levels of fertility, especially in cultures that lack inexpensive, reliable birth control.

For basic biological reasons, grand-parents will often want more children than their own children will independently desire, because the grandparents receive a direct benefit from grandchildren, over and above the indirect impact that grandchildren have through their children's own welfare. If children have maximized their own welfare with respect to their own progeny, the envelope theorem implies that grand-parents will desire a higher number of grandchildren. There are multiple means of prodding children to be fertile, including bribes and verbal haranguing, but investing in beliefs may be a reasonable tool.

Parents have both a strong motive and abundant means of influencing children's beliefs, such as exposing children to gender stereotypes in childhood literature. Weitzman et al. (1972) examines children's storybooks in the US, and finds pervasive differences in the ways that genders are depicted, with boys being adventurous and girls being pretty and passive. Bereaud (1975) examines French children's books and similarly finds that they portray girls as "timid, passive and dependent" and women "in the traditional housewife role or in low-paid, unskilled occupations." Children's books are bought by parents, so it is reasonable to believe that parents want such images broadcast to their own children.

In the pluralistic United States today, parents can also choose other influences, such as religion. If parents want to encourage childbearing, then they can take their children's to religious institutions that encourage child-bearing. Some of the most extreme examples of pro-natalist religious entities are the Mormon Church and various ultra-orthodox groups. These institutions and the traditional Catholic Church also encouraged large families and traditional female lifestyles. Religious support for child-bearing may reflect both a desire to cater to parents who want grandchildren, but also a desire to fill the pews in decades to come. Religious groups that did not support child-bearing, such as the Shakers, tend to disappear over time.

As we will model, parents can also engage in more costly signals to children about their abilities. A mother may herself adopt a traditional lifestyle to convince her daughters to do the same and her sons to marry someone who acts similarly. Providing little education for daughters is another means of suggesting that her possibilities in the workplaces are limited, and that she should focus more on producing grandchildren. We will formally model under-education of women.

We will focus on the signaling choices of individuals, which will inevitably lead to some heterogeneity in the population. That heterogeneity may be smoothed out by institutions, such as churches, which will lead to a more ubiquitous set of attitudes. A state may also embrace

traditional lifestyle choices for pro-natalist reasons, which may in turn be motivated by the desire for a large army. Hitler's Germany for example, pushed a strong ideology of motherhood and traditional female roles (Rossy, 2011).

Empirically, demographers have documented that parental preferences do affect children's preferences and decisions on marriage and childbearing. Axinn, Clarkberg and Thornton (1994) show that mothers' preference for the size of their children's family is significantly positively correlated with the children's family size preferences when the children are young adults. Barber (2000) shows that both sons and daughters whose mothers prefer early marriage, large families and low minimum education for their children end up entering parenthood earlier. This effect is significant controlling for family income, parental education, the mother's work choice and other family background variables. Such evidence corroborates our idea that parental influence is possibly quite powerful.

IV. Gender Stereotypes, Education and Daughters

The critical assumption in our model is that the parents care about the welfare of their children <u>and</u> their grandchildren. Parents would like to prod their children to have more children themselves, thereby increasing their total number of descendants. Evolutionary theory would seem to suggest such preferences, as would ordinary observation. In particular, popular culture is replete with examples of parents wanting their children to get married and have at least one child of their own.

This assumption about parental preferences then influences parental investments in their children, especially when those preferences shape the beliefs of those children.   Parents have many tools for influencing beliefs about female competence in the workforce, including telling stories, attending religious services, maternal behavior and so forth.   We will focus on the provision of education for daughters. Female education is a particularly important signal that parents can send daughters about their productivity outside the home. .

We will focus on beliefs about female competence, but this is only one possible interpretation of the model.   The "competence" parameter can also be interpreted as the psychic benefit of working.   An alternative interpretation, therefore, is that parents attempt to persuade their daughters that there are lower returns from working in the formal sector than from bearing children.   In a sense, the model can be interpreted as suggesting that to increase the size of the third generation parents are trying to persuade their daughters of the virtues of "traditional values."[9]

---

[9] This interpretation relates to Boustan and Collins (this volume) that documents that non-working mothers are more likely to have non-working daughters. If non-working mothers have stronger preferences for both their own children and for grandchildren, then they will indeed be willing to invest more in changing the preferences of their progeny

Our model assumes three generations (grandparents, parents and children). The grandparents act first in period one. They select the investment in human capital for a specific child in the second generation. We assume that we are looking at the decision of grandparents after their own fertility decisions have been made. In period two, the parents' generation then decides on the number of children that they have and the human capital of those children., The children make no choices in the model and are assumed to be homogeneous. Table 1 discusses the core assumptions of the model and the implication of eliminating those assumptions.

The Period Two Decisions by the Second Generation

In period two, daughters in the second generation (parents) choose fertility levels, N, human capital levels for their boy children (H_{CM}) and human capital levels for their girl children (H_{CF}) to maximize:

(1) Consumption $+ \alpha\big(V(N) + .5N\varphi(H_F)g_M(H_{CM}) + .5N\varphi(H_F)g_F(H_{CF})\big)$,

where $V(.), \varphi(.)$ $g_M(.)$ and $g_F(.)$ are all increasing, concave functions. The function $\varphi(H_F)$ captures the extent to which maternal human capital translates into the well-being of the second generation. We assume that ½ of all children are male and that the benefits of skill may be different between boys and girls.

Consumption is assumed to equal $\delta W(H_F)A_F(1 - T_{HF}) + Y_0 - .5NH_{CM} - .5NH_{CF}$, where $\delta W(H_F)A_F(1 - T_{HF})$ reflects the wife's earnings, which equals $\delta W(H_F)A_F$ (the wife's wage rate) times $1 - T_{HF}$ (the time spent working) and Y_0, reflects any other income including husbands' earnings. Household time is proportional to the number of children, so $T_{HF} = N\underline{t}_C$. The wage equals a discrimination level, denoted δ, times a wage function that is increasing in the women's human capital ($W(H_F)$) times an ability level A_F. We further assume that W(.) is increasing, concave and that $lim_{x \to 0} W(x) = \infty$. We first assume that women make fertility decisions before observing their workplace productivity, and make work timing decisions based on an estimate: $\hat{A}_F$. We relax that assumption in Section V.

The first order conditions that determine human capital level investments are $g_M'(H_{CM}) = g_F'(H_{CF}) = 1$. We let G_T denote $.5\big(g_M(H_{CM}) + g_F(H_{CF})\big)$, and H_T denote $.5H_{CM} + .5H_{CF}$ evaluated at the welfare maximizing levels of human capital investment. We assume that $G_T \geq H_T$.

Three assumptions together ensure that the investment in children's human capital is independent of the number of children: quasi-linear preferences, the benefits from investing in children scales up linearly with the number of children and the costs of human capital investment similarly scale up linearly with the number of children. Quality and quantity of children are not completely independent, however, as the net benefit from investing in quality will impact the incentive to have more children.

The optimal fertility choice is characterized by the first order condition $\delta W(H_F)\hat{A}_F \underline{t}_C = \alpha(V'(N) + G_T) - H_T$. Differentiating this equation implies that the number of children is increasing with α and decreasing with $\delta, H_F, \underline{t}_C$ and $\hat{A}_F$. Other than α, all of these parameters effectively increase the opportunity cost of having more children.

We use this equation to implicitly define a function $N(W(H_F)\hat{A}_F)$, which represents the number of children that a women will have depending on her level of human capital and beliefs about her workplace ability. The other elements that determine utility have been suppressed because they are fixed. Holding $\hat{A}_F$ and other parameters constant, the derivative of N with respect to H_F is $\frac{\delta W'(H_F)\hat{A}_F \underline{t}_C}{\alpha V''(N)} < 0$. The second derivative of N with respect to H_F is negative as long as $V'''(N)$ isn't too negative, as it won't be if V(.) has a standard form such as vN^σ, with $\sigma < 1$.

The total welfare of a female child equals $Y_0 + \delta W(H_F)A_F(1 - N\underline{t}_C) + \alpha(V(N) + NG_T) - NH_T$, where N will be optimally chosen in response to the other parameters. This welfare level and the choice function N are then ingredients into the decision-making of the first generation.

The Period One Decision by the Grandparents' Generation

We now turn to the grandparents' generation, and focus on their choice of investment in human capital for a single, female child in the second generation. The grandparents choose only the level of human capital, which carries a cost H_F, just like the human capital by the next generation. The grandparents' welfare will equal α_1 times the daughters' direct welfare, $Y_0 + \delta W(H_F)A_F(1 - N\underline{t}_C) + \alpha(V(N) + NG_T) - NH_T$, plus the utility that grandparents get directly from the third generation, which equals α_2 (a second altruism parameter) times $V(N) + NG_T$, which is also the welfare that the second generation receives from the third generation. .

Throughout the model, we will assume that the first generation accurately assesses the ability of their children in the second generation. This assumption can be relaxed, as long as the grandparents retail some private information. For example, the grandparents could have some private information about the state of the labor market or a private, imperfect signal about the daughters' ability level. As long as the grandparents' have some private information than a signaling game will still occur, but if the grandparents were known to know nothing more than the second generation, then they would have no ability in a rational model to influence the beliefs of their children.

Assuming that parents accurately assess the daughter's ability level A_F, the welfare of parents (that is related to a specific child) equals:

$$(2) \quad \alpha_1\left(Y_0 + \delta W(H_F)A_F(1 - N\underline{t}_C) + \alpha(V(N) + NG_T) - NH_T\right) + \alpha_2(V(N) + NG_T) - H_F$$

where H_F refers to the investment of human capital in the parent. The parameter α_1 reflects the direct impact of the second generation's welfare on the welfare of the first generation. The

parameter α_2 reflects the impact of the third generation's welfare on the welfare of the first generation. If the second generation chooses their fertility level to maximize their own welfare, then the derivative of grandparents' welfare with respect to N, the number of grandchildren, equals $\alpha_2(V'(N) + G_T)$, which is strictly positive. Given these preferences, grandparents will always want their children to have more progeny than they will naturally choose on their own.

Standard evolutionary preferences suggest that animals act as if they care about reproducing their gene pool, not just for a single generation but for generations to come. One approach that grandparents' might have is to provide cash assistance that is tied to the number of children produced. Many grandparents do explicitly subsidize grandchildren, if they have the resources, by providing funds for education or even buying a house in a neighborhood with a good school district.

Grandparents will also want parents to spend a bit more investing in their children's human capital, and this might reverse the results of the model. We have structured preferences and production functions so that there is no tradeoff between quantity and quality, and where the daughters' human capital does not increase the human capital of the next generation. If daughters' human capital did contribute directly to the human capital of their own children, then grandparents' would have a stronger incentive to invest in their daughters. This effect would tend to mute the implications of the model.

We have chosen a stark and simple case to highlight how a desire for own grandchildren may lead to lower human capital investment in girls, and the generation of beliefs about female inability in the workplace, but we are well aware that reasonable perturbations of the model could generate alternative predictions. For example, grandparents could conceivably care so much about grandchildren's quality that they might actually not want higher fertility levels. We have also assumed that maternal human capital only impacts child-bearing by increasing opportunity costs. If maternal skills help generate human capital in the next generation, then this would create a grandchild—related incentive for investment in daughters, as described above.

We first focus on investments in a daughter's human capital, assuming that A_F is known at every point. We then turn to the possible scenario in which the parents, but not the daughter has received a private signal about the daughter's ability, in which case investing in education can serve as a costly signal to the daughter of her skills. Finally, we address sexist indoctrination of sons.

When the future mother's ability level is known to all, then the first order condition for the grandparents is:

$$(3) \quad \alpha_1 \delta W'(H_F) A_F \left(1 - N\underline{t}_C\right) + \alpha_2 (V'(N) + G_T) \frac{\delta W\prime(H_F) A_F \underline{t}_C}{\alpha V''(N)} = 1$$

We assume that second order conditions, described in the Appendix, hold for this to be a maximum.

Given our assumption on the second order condition, Proposition 1 follows[10]:

Proposition 1: Parents will invest a positive amount in daughters' education if $-\frac{N_0 V''(N_0)}{V'(N_0)+G_T} >$ $\frac{\alpha_2}{\alpha\alpha_1}\frac{N_0 \underline{t}_C}{1-N_0 \underline{t}_C}$ where N_0 represents the number of children chosen by a daughter with no education. If this condition holds, and parents do invest in a positive amount of education, then the level of education is declining with α_2 and increasing with α_1. The level of education will increase with δ and A_F if and only $\frac{(-\alpha V''(N))}{(\delta A_F \underline{t}_C)^2 W'(H_F)W(H_F)} > \left(\frac{\alpha_2}{\alpha} - \alpha_1 - \frac{\alpha_2 V'''(N)(V'(N)+G_T)}{\alpha(V''(N)^2}\right)$

Proposition 1 implies that parents will always invest a positive amount in their daughter's education if α_2 is sufficiently small, and that the amount of education that their daughter receives is decreasing as α_2 rises. The incentive to under-invest in daughters is directly a function of the altruism towards grandchildren, but of course, this would diminish if daughters' human capital were an input into the human capital of the next generation. By contrast, as the grandparents care more about their daughters relative to their grandchildren, investment in the daughters' education will rise.

The parameters δ and A_F are complements to daughters' education, and they will typically cause the investment in the daughters' education to rise, as long as α_2 is relatively small, so the dominant effect of these parameters is to increase the payoff to daughters' education. A somewhat less intuitive possibility is that if α_1 is sufficiently low, higher values of δ and A_F which increase the returns to work may actually reduce the tendency to invest in daughters' education. If α_1 is low enough, then the grandparent only cares about investing in human capital because it impacts the supply of eventual grandchildren. As higher values of δ and A_F reduces the number of grandchildren directly, this may sufficiently increase the grandparents' demand for more grandchildren that they may offset these higher labor market returns with less investment in human capital.

In sum, in the model with perfect information there is an incentive to under-invest in daughters (in order to induce them to have more children), but there is no attempt to shape the beliefs of daughters about themselves or about women in general. We now turn to the situation in which the parent has some private information. When modeling the investment in daughters, we assume that this information is about the young woman herself. When discussing investment in sons, we will assume that this information is about women more generally.

A crucial assumption of the model is that the second generation's human capital does not influence the "quality" of the third generation. To briefly illustrate how drastically results can

---

[10] Proofs may be found in the working paper version of this chapter (Glaeser and Ma 2013).

change when this assumption is relaxed, assume that $G_T = G_0 + g(H_F - H_F^0)$, where G_0, g and H_F^0 are constants. The parental first condition is now $\alpha_1\left(\delta W'(H_F)A_F(1 - N\underline{t}_C) + \alpha g N\right) +$
$\alpha_2\left((V'(N) + G_T)\frac{\delta W'(H_F)A_F\underline{t}_C}{\alpha V''(N)} + gN\right) = 1$. The first generation will still have an incentive to underinvest in daughters' education to increase fertility, but the comparative static on α_2 is reversed $(V'(N) + G_T)\frac{\delta W'(H_F)A_F\underline{t}_C}{\alpha V''(N)} < gN$, so the grandparents who care about their grandchildren invest more—not less—in their daughters. We will drop this assumption now, but return to it in our later section on credulous Bayesians.

Belief Formation

We now turn to the core of the model --the formation of daughters' beliefs. The first generation parents know their daughters' ability, but daughters themselves only infer their talents from parental investment in their human capital. In equilibrium, a daughter whose parents invest heavily, both personally and through external investments, will typically infer that she has abundant raw skill, since we assume that such skill is a complement with investment in the model. If parents ignore a daughter's education, then she will naturally infer that she has little innate talent. At this point, we focus on the formation of beliefs by a single individual, but in the next section, we discuss the implication of this for beliefs by sons and by society as a whole.

The timing of the model, preferences and production functions are just as before, but we now assume that A_F can take on two values 1 and 1-a, and daughters are more able with probability p. Parents learn their daughter's skill, make an investment and then the daughter infers her skill from their investment levels and then makes her own fertility choice. We define the investment level chosen under perfect information as H_F^{Skill} for skilled daughters and $H_F^{Unskill}$ for unskilled daughters.[11] These are benchmark quantities that would be chosen if the daughters knew their ability level. We also define $H_{Skill}^{Unskill}$ as the level of human capital that would be chosen by parents of unskilled daughters if their daughters believe erroneously that they are skilled. We let π denote daughter's assessment that she is high skilled, based on the human capital level that she has received, and $N(H_F, \pi)$, denote her fertility level, which is decreasing in both H_F and π. First generation welfare can then be denoted $U_i(H_F, N(H_F, \pi))$ for i=u and s, depending on whether the daughter is skilled or unskilled.

We have assumed that there are only two groups in the population and that the parents of the more skilled wish their daughters thought that they were less skilled. This feels particularly harsh, but that harshness can be reduced if we instead assumed that there were a variety of subgroups in the population and daughters knew their subgroup. Then the assumed desire to

---

[11] The values of H_F^{Skill} and N_{Skill} satisfy $\delta W(H_F^{Skill})\underline{t}_C + H_T = \alpha(V'(N_{Skill}) + G_T)$, and $1 = W'(H_F^{Skill})\delta\left(\alpha_1(1 - N_{Skill}\underline{t}_C) + \frac{\alpha_2(V'(N_{Skill})+G_T)t_C}{\alpha V''(N_{Skill})}\right)$, and other values are defined similarly.

push the daughter's assessment of her ability downward only means that parents wish that the daughter thinks she is less able relative to her subgroup, not relative to the entire population. In this setup, the parents of a talented daughter would be happy to have her realize that she is more able than most women or prevalent gender stereotypes, but would still want her to shade her self-confidence downward slightly. As discussed above, the ability parameter can also be interpreted as reflecting the value of working inside or outside the home, and can therefore be seen as capturing values rather than innate ability.

Locally, the welfare of the grandparents is decreasing with π—the daughter's belief in her own competence—but we go further and assume that this derivative holds globally as well, so that parents would always prefer their daughters' to think that they have a lower probability of being able. This assumption implies that skilled parents would like to imitate unskilled parents. If that assumption does not hold, then there will be little incentive to manipulate beliefs. This assumption follows automatically if "a" is sufficiently small, so that the fertility choices by less skilled children are only slightly greater than the fertility choices of more skilled children.

We further assume a minimum level of investment that parents are legally required to make, which is denoted $\underline{H}_F$, and that this is less than $H_F^{Unskill}$. Thus in a world with perfect information, this lower bound will not bind.

In a Bayesian separating equilibrium where skilled and unskilled daughters receive different levels of education, then daughters learn their "type." If both types of parents choose a level of investment in equilibrium, then daughters will believe that they are skilled with some probability weakly between zero and one. We will use an equilibrium refinement to suggest which equilibrium seems most likely to exist.

What are possible outcomes if the parents of skilled and unskilled daughters choose different levels of education? In these outcomes, the parents of skilled daughters must choose H_F^{Skill}, since that maximizes $U_S(H_F, N(H_F, 1))$. Any other investment level will not change daughters' self-assessments, and will only reduce parental welfare. Moreover, any alternative candidate separating equilibrium investment level will generate a deviation to this point, since the parents of the skilled cannot change the daughters' beliefs in an adverse way—as long as there is separating—and they can better the match between their daughters' skill and her human capital level.

Since the welfare of the parents of skilled, in any separating equilibrium, is determined by $U_S\left(H_F^{Skill}, N\left(H_F^{Skill}, 1\right)\right)$, it is helpful to determine the range of values of investment for the parents of the less skilled, which would induce the parents of the more skilled to imitate them:

Lemma 1: There exists one value of H, denoted $\widehat{H}$, at which $U_S\left(H_F^{Skill}, N\left(H_F^{Skill}, 1\right)\right) = U_S\left(\widehat{H}, N(\widehat{H}, 0)\right)$, and for all values of H between $\widehat{H}$ and H_F^{Skill}, $U_S\left(H_F^{Skill}, N\left(H_F^{Skill}, 1\right)\right) <$

$U_S(H, N(H, 0))$, but if $H < \widehat{H}$, $U_S\left(H_F^{Skill}, N\left(H_F^{Skill}, 1\right)\right) > U_S(H, N(H, 0))$. If $H > \widehat{H}$, then $U_U(H, N(H, 1)) MU_U\left(\widehat{H}, N(\widehat{H}, 0)\right)$. At $H = \widehat{H}$, holding beliefs constant, the welfare of parents of skilled daughters is strictly increasing in H, and $\widehat{H}$ is rising with δ and α_1 and falling with α_2.

Lemma 1 helps determine the structure of a separating equilibrium. The value of H, denoted $\widehat{H}$, if the highest investment of human capital by parents of the less skilled that will not induce the parents of the more skilled to attempt to imitate them. If the less skilled invest more than $\widehat{H}$, then the parents of the more skilled will choose to imitate them, for $U_S\left(H_F^{Skill}, N\left(H_F^{Skill}, 1\right)\right) < U_S(H, N(H, 0))$, but if the parents of the less skilled invest less than that amount then the more skilled will not benefit by imitating them. If the parents of the less skilled are investing $\widehat{H}$, then they will not benefit by mimicking the parents of the more skilled.

In a separating equilibrium, the unskilled parents however, may well end up choosing an investment level other than $H_F^{Unskill}$. While that skill level is perfectly matched to their daughters' ability and their preferences, it may not be an equilibrium since if $H_F^{Unskill} > \widehat{H}$, it may lead the parents of the skilled to want to imitate them. Moreover an alternative investment level will not necessarily generate a deviation, since a deviation towards an alternative investment level may cause beliefs to change in a way that hurts the welfare of the parents of the less skilled.

A pooling equilibrium is also possible, but generically, it will only be possible for the two types to pool at more than one human capital level. In an equilibrium where parents of the same type choose two different levels of human capital, these parents must be indifferent between the two levels of investment. The beliefs that make one type of parent indifferent between two levels of investment will not make the other type of parent indifferent between two types of investment. Formally if $U_S\left(H_F^1, N(H_F^1, \pi_1)\right) = U_S\left(H_F^2, N(H_F^2, \pi_2)\right)$, where H_F^1 and H_F^2 represent the two different investment levels and π_1 and π_2 represent the beliefs at the two investment levels, then generically $U_U\left(H_F^1, N(H_F^1, \pi_1)\right) \neq U_U\left(H_F^2, N(H_F^2, \pi_2)\right)$.

If we place no further restrictions on off-the-equilibrium-path beliefs, then multiple equilibrium are possible. For example, it is possible for there to be a continuum of pure separating equilibrium, where parents of skilled daughters choose H_F^{Skill} and parents of unskilled daughters choose any value of H below $\widehat{H}$, as long as the parents of unskilled daughters prefer that value of H and being known to have an unskilled daughter to choosing $H_{Skill}^{Unskill}$, which is the best that they can do if their daughters believe that they are skilled. It is also possible for the unskilled parents to choose two levels of H that yield equal utility levels, as long as one is above $H_F^{Unskill}$ and one is below H_F^{Skill} and both yield equal utility. It is also possible for there to be a pooling equilibrium, where both parents of skilled and unskilled daughters choose a common level of H, as long as the payoff for the parents of skilled daughters is better off than if they choose H_F^{Skill}

and the parents of unskilled daughters are better off than if they chose $H_{Skill.}^{Unskill}$. Semi-pooling equilibria are also possible, where some fraction of both groups mix and choose a common equilibrium, as well also choosing some separate investment level.

However, many of these seem like unlikely outcomes since if $H_F^{Unskill} < \hat{H}$, daughters' who observed a deviation to $H_F^{Unskill}$, which would yield higher welfare for the parents, would surely still infer that they were less able, since choosing $H_F^{Unskill}$ would produce less welfare for parents of the more skilled than they are already receiving in equilibrium. Separating equilibrium, therefore, seem most likely to yield exactly two investment levels, one for each skill level.

To formalize this intuition and generate a unique equilibrium when $\hat{H} > H_F^{Unskill}$, it is sufficient to assume a perfect Bayesian equilibrium where daughters believe that if one type of parents would never deviate to a human capital level H, given any rational fertility response, then the deviation must come from the other group. This assumption leads to proposition 2a:

Proposition 2a: If $\hat{H} > H_F^{Unskill}$, then there is no pooling equilibrium, skilled parents choose to invest H_F^{Skill} and unskilled parents choose $H_F^{Unskill}$.

This proposition suggests that there is one most likely outcome if $H_F^{Unskill}$ is sufficiently low. Since the parents of the skilled would not choose $H_F^{Unskill}$ even if their daughters would change their beliefs, then daughters reasonably believe that such an investment level can't come from parents of the more skilled. Since they believe that an investment level of $H_F^{Unskill}$ must come from the parents of the unskilled, then parents of the unskilled will choose that investment level and a unique equilibrium results.

Proposition 2a describes the equilibrium for a range of parameter values, when the parents of the skilled have little incentive to imitate the parents of the weak. In this case, the outcomes with imperfect information are identical to the outcomes with perfect information. There is some incentive to reduce the education of daughters, of both skill levels, but little actual deceit or misinformation.

But what about situations in which $\hat{H} < H_F^{Unskill}$, and the parents of skilled daughters would like to pretend to be parents of unskilled daughters? In these cases, there is the potential for misinformation, and no possibility that parents will behave exactly as they did in the full information case.

All sorts of possible outcomes seem to coexist. For example, skilled and unskilled might pool at some relatively high level of schooling. This equilibrium would be maintained if daughters interpreted any deviation as coming from parents of the more skilled. Alternatively, the less skilled might choose some extremely low level of schooling (less than $\hat{H}$) and this would be

maintained by if daughters believed that any deviation came again from parents of the more skilled.

To select a single equilibrium prediction for a wider range of parameter values, we now assume a variant of the D1 refinement (Cho and Kreps, 1987): if an off-the-equilibrium path investment level is more attractive for one type of parent, given any set of beliefs by children, then children assume that this type of parent has generated this deviation with probability one.   In this model, the children's response to the parents' human capital investment is their fertility level. If $N_S^*(H)$ and $N_U^*(H)$ makes the parents of skilled and unskilled children respectively indifferent between their equilibrium payoff and any deviation H, then if $N_S^*(H) > N_U^*(H)$ the deviation seems more likely to have come from a parent of a skilled child.   If $N_S^*(H) < N_U^*(H)$ then the deviation seems more likely to have come from a parent of an unskilled child.   The D1 refinement requires to think that the deviation comes, with probability one, from the parent of an unskilled child if and only if $N_S^*(H) < N_U^*(H)$.   This is a strong assumption that produces a single equilibrium for all parameter values:

Proposition 2b: If $H_F^{Unskill} > \widehat{H} > \underline{H}_F$, then skilled parents choose H_F^{Skill} and unskilled parents choose $\widehat{H}$. If $\widehat{H} < \underline{H}_F$, then all unskilled and some skilled parents choose $\underline{H}_F$. In that case, the number of parents of skilled daughters choosing $\underline{H}_F$ will decrease with α_1 and increase with α_2.

The equilibrium then follows the value of $\widehat{H}$, which as discussed in Lemma 1, is a function of the labor market discrimination against women. When the value of δ is very low and women have weak opportunities in the labor market, then $\widehat{H} < \underline{H}_F$. In this case, it is impossible for parents of the unskilled to choose a level of education that separates themselves completely from the parents of the skilled. In this case, the parents of the skilled and unskilled both choose the minimum level of girls' education, and skilled and unskilled daughters alike are both less educated than they would be under complete information. This may represent the setting in highly traditional societies where parents radically under-invest in their daughters.

Ultimately, there can be a complete pooling equilibrium where all parents end up providing girls with only the legal minimum of education.   Any deviation upwards will be seen as an indication that the girl is skilled, and will generate lower fertility levels. This force essentially traps society in a world where women are less educated and unable to distinguish among the more or less skilled.

For higher levels of δ, where labor market discrimination is less severe, the desire to distort views influences the education choices of the parents of the less skilled, but not the parents of the more skilled. The parents of the less skilled provide less education for their daughters in order to distinguish themselves from the parents of the more skilled. Their daughters end up having more children both because their opportunity cost of time is less and also because they know themselves are less able in the workforce.

When δ is high, and there is little labor market discrimination, then $\hat{H} > H_F^{Unskill}$, and a separating equilibrium will exist with no distortion of parental incentives as discussed in Proposition 2b. The parents of the less skilled educate less—the parents of the more skilled educate more.

Parental altruism works throughout this model. As parents care about their daughters more, relative to their grandchildren, pooling is less likely and skilled daughters receive more education. But if parents are particularly focused on their long-run genetic legacy, then daughters pay the cost in lower educational outcomes.

The model has several implications. When labor market discrimination is strong, then parents of skilled and unskilled daughters alike choose to provide them with minimal education. The skilled daughters may particularly suffer, because their parents are trying to ensure that they do not realize their skills.

As women are less discriminated against, this leads to more investment in the skilled daughters, and there can be a discrete jump in educational investment for this group. Previously, some members of the skilled group will be treated like less skilled children, and a lucky few will receive more schooling. Afterwards, all members of the skilled group will get more schooling. We think of this as capturing the gradual rise in women's college education in the U.S. during the early twentieth century.

Eventually, signaling concerns lose power in a pure separating equilibrium, and the skills essentially serve to maximize grandparents' welfare. Of course, all daughters will still be undereducated, because parents are trying to engender more fertility, but they will at least become informed about their talents. The underinvestment in female education may vanish altogether, if parents lose control over the educational investment of their daughters, if female education does little to reduce fertility, or if daughters' education leads to more investment in the quality of grandchildren, which grandparents value.

This model suggests that the population would have only two levels of education for women, but that would not be the case if there were visible differences in parameters across the population. In that case, different parameters will lead to different equilibria, although for any given set of observable parameters, parents will still use education to influence their daughters' beliefs.

If taken literally, then in the parameter space when pooling occurs, skilled daughters do not know that they are skilled, but at least they, and everyone else, correctly infers the share of women in the population who are skilled. Yet children may be unable to actually know the true share of skilled daughters, since they do not observe any daughters being well education. Since there is little hard evidence on skills, parents may be able to persuade sons and daughters alike that skilled daughters are rare even if they are common. Such stories would not be falsified by anything in the children's experience.

Once separation occurred, then such stories would fall apart.   We believe that this signaling model therefore may be connected to broader societal beliefs, even though the model itself contains no improper updating.   In a regime in which women are not differentiated by skill, it would seem possible for parents and their allies to argue that the share of skilled women is low, which would encourage both sons and daughters to produce more grandchildren.[12]

V.   Extensions: Systemic Beliefs, Credulous Bayesians and Work Timing

The previous section described our core model, in which parents deceive daughters about their ability level by under-investing in education.   In that setting, we assumed that daughters knew the population's propensity to be unskilled, although not whether they are skilled themselves. As such, there will be some daughters who underestimate their own skills in the pooling equilibrium, but no society-wide tendency to diminish women.   We now drop the assumption that children know the society level distribution of skills, and assume that parents but not children have this knowledge.

Society-Wide Discrimination

The previous discussion focused on the formation of beliefs of daughters about their own ability, but we know perturb the model to focus on systemic beliefs about woman's ability.   In this case, we assume that children are unsure about "p" – the probability that women are less ability and specifically that they initially assume that p is uniformly distributed on between p and 1. Parents are assumed to have no more knowledge of p than children, and children don't observe the society wide level of skills among daughters until after they make fertility choice.   To simplify matters, we assume that the actually value of p is arbitrarily close to one, so that there are only an arbitrarily small number of parents with less able girls. The actual ability of the daughters will only be revealed after they make their fertility and work choices.

If children correctly understand their parents' preferences, then the model is essentially unchanged.   There are ranges of values at which there is separating, semi-pooling and pooling. we are in a situation quite similar to the one discussed above. Assuming our version of D1, Proposition 3 follows:

---

[12] The core assumption of our model is that parents want more grandchildren than their sons and daughters naturally will give to them. The same parental preferences should also generate incentives to engage in other forms of belief investment, most notably inculcating opposition to homosexual lifestyles. If homosexuality leads to less own grandchildren, then parents who value own grandchildren will invest in their children's beliefs to that end. They will attempt to convince them that homosexuality will lead to unhappiness and perhaps worse. In this setting as well, religious organization may offer parents a means of perpetuating beliefs that serve their biological interests. If the church supports traditional lifestyles and opposes homosexuality, then parents may have an incentive to take their children to church despite their own private religious beliefs.

Proposition 3: If the second generation initially does not know the share of daughters who are capable they will learn that share if they observe the education decisions of the first generation when $\frac{\alpha_2}{\alpha_1}$ is sufficiently low, so that there is either a separating or semi-pooling equilibrium, but not if there is a pooling equilibrium.

The intuition of this proposition is that observing the ex post distribution of skills can typically enable observers to deduce the true distribution of skills in female population, unless there is full pooling. If there is a separating equilibrium, then there will obviously be complete revelation. If there is semi-pooling, then the proportion of parents of skilled daughter who under-educate them can be inferred from the parameter values. The actual share of under-educated women then would allow observers to infer the actual distribution of skills.

Only in the case of full pooling, where there is literally no information about the underlying distribution of skills does ignorance persist. In that case, a belief that women are less able can persist, because parental knowledge never gets transformed into action.

There are reasons, however, that we might doubt this version of the world. This suggests a very stark difference between female and male education. It suggests that as discrimination decreases, there should be sharp societal jump in beliefs as soon as some women receive more education and this education is observed. The proposition does require widespread information and a fair amount of rationality, neither of which may exist in reality; that said, it does suggest that discriminatory beliefs can disappear in a generation, even if parents persist in under-educating their daughters.

Credulous Bayesians

An alternative approach, which is somewhat less attractive given strict assumptions about hyper-rationality but which perhaps lies closer to the truth, is that children mis-perceive parental preferences, and believe that $a_2 = 0$. As such, they believe that parents make decisions only to improve children's welfare, and not to manipulate their children's beliefs. This a version of Glaeser and Sunstein's (2009) "Credulous Bayesianism" where agents use Bayes' rule to make inferences but they underestimate the incentives of people around them to persuade. Given the pervasive altruism that exists in parent-child relationships it would be particularly natural for children to think that parents are particularly benign.

If a_2 is thought to be zero exactly, then children will look at parents' investment in daughters and believe that these investments maximize: $\alpha_1\left(Y_0 + \delta W(H_F)\hat{A}_F\left(1 - N\underline{t}_C\right) + \alpha(V(N) + NG_T) - NH_T\right) - H_F$, and hence $\hat{A}_F(H_F) = \frac{1}{\alpha_1 \delta W'(H_F)\left(1 - N\left(W(H_F)\hat{A}_F(H_F)\right)\underline{t}_C\right)}$ and $\hat{A}_F{'}(H_F)$ is positive as long as children believe that their parents' maximization problem is concave. Proposition 4 follows:

Proposition 4: If children believe that $\alpha_2 = 0$, then parents will choose to invest in daughters' education so that both sons and daughters believe that women are less talented than men. Higher values of α_2 will cause human capital investment to fall, the belief in women's ability to decline, and fertility to rise.

Since children may believe that parents lack ulterior motives, underinvestment in daughters is interpreted as meaning that daughters are expected to be less able.   This belief will occur in both daughters and sons, and it will become more extreme if parents have stronger preferences towards grandchildren.

In a more complicated version of the model considered in the working paper version of this chapter (Glaeser and Ma, 2013), with both sons and daughters, we found that the underinvestment in daughters becomes more extreme when there are more boys in the household and less extreme when there are more girls.   The logic of this effect is not that boys take up girls' resources (there are no income effects in this model), but rather that boys present an added target for indoctrination. While the logic of this argument appears clear, it runs counter to the finding of Butcher and Case (1994) that women with only female siblings receive less, not more, education.   Of course, these findings could reflect forces outside of the model, such as spillovers from boys to girls during the middle twentieth century.

To illustrate the importance of our assumption that daughters' human capital doesn't influence the "quality of her own children," we return to the assumption that $G_T = G_0 + g(H_F - H_F^0)$. In this case, parents who care about the next generation would have competing incentives.   They would want to invest more in their daughters, so that their grandchildren end up being more skilled, and invest less in their daughters, so that their grandchildren are more abundant.   If the two forces are perfectly balanced, then the credulous Bayesians will actually be right.

This possible extension offers one suggestion about why gender-based discrimination may have faded. If daughters receive more investment because the returns to human capital in the next generation increase, then they and the men around them will infer that they have a higher ability level themselves, assuming that they incorrectly underestimate the altruism towards the next generation.   If this hypothesis is correct, then the rise in returns to skill might have the added impact of reducing gender stereotypes.

Timing of Work and Persuasion

The models of persuasion that we have discussed ultimately assume that women are choosing their fertility levels with little direct knowledge of their workplace ability, but that would seem to depend on the timing of work and childbearing. If the mother works initially, she will surely have a better assessment of her talents from that direct source than from anything she may have inferred from either refrigerator advertisements or even her parents' investment in her human capital.   That knowledge will then essentially eliminate the incentive to persuade initially.

For simplicity, we continue to assume that each child requires a fixed time investment of $\underline{t}_c$, although we ignore investment in children's human capital. We assume that mothers maximize the expected value of $y_0 + \int_{t=0}^{1} y(t)dt + \alpha V(N)$, where y(t) is the earnings at each t, so there are no discounting issues. Women end up being paid their expected or realized productivity level multiplied by $\delta W(H_F)$.

We only consider two options in childbearing. First, the mother has children immediately, basing her fertility decision on her expected workplace earnings. Second, the mother delays child-bearing to the point where she has learned her actually productivity in the workplace and then decides on fertility knowing her actual ability level. We ignore more complicated strategies, and assume that the primary costs of delay are health or time related, so that the expected time cost of N children, when child-bearing begins at time t_0, will be $g(t_0)N\underline{t}_C$ where $g(0) = 1$ and $g'(t_0) > 0$. These assumptions capture both the added difficulty of having children when older and that the ability to produce children has historically been impossible for women beyond some age. We ignore other benefits of later child-bearing (more experience in life) and other costs (more human capital may depreciate during the child-bearing period).

Moreover we continue to assume A_F equals either 1 or 1-a; hence the probability that ability equals 1 is given by $1 - \frac{1-\hat{A}_F}{a}$, where $\hat{A}_F$ is the women's expected ability level. The expected payoff from having children immediately will be $y_0 + \delta W(H_F)\hat{A}_F(1 - N\underline{t}_C) + \alpha V(N)$, meaning that as before N satisfies $\delta W(H_F)\hat{A}_F\underline{t}_C = \alpha V'(N)$. With delay, the woman learns her true ability level and eventually chooses fertility to maximize $g_\psi \delta W(H_F)A_F\underline{t}_C = \alpha V'(N)$, where we let g_ψ denote the health cost of delaying fertility until the point of knowledge. This leads to Proposition 5.

Proposition 5: There exists a value of g_ψ at which women are indifferent between postponing work or postponing childbearing, and for higher values of g_ψ women will postpone childbearing and for lower values of g_ψ, women will postpone work. If women postpone childbearing, then changes in the initial beliefs about workplace will have no impact on their fertility decision.

Proposition 5 suggests that medical advances that permit delayed childbearing may have far-reaching impacts on social beliefs. If women are making fertility decisions early in life, then those decisions will be based not on their actual workplace productivity, but rather on the information that they have gleaned about the relative pleasure of working and child-bearing. That position of ignorance creates a possible role for persuasion for grandparents interested in encouraging fertility, or anyone else interested in persuading men and women.

But if women obtain substantial work experience before having children, then the impact of any such persuasion is highly muted. The knowledge gained in the labor force will surely swamp the knowledge inferred from parental education decisions or the persuasion of consumer goods

companies. As such, the delay in child-bearing can powerfully change the incentives for persuasion. This effect connects the time series of female labor force participation with the time series of opinions about female competence at work. As Goldin (2006) describes, women were initially prone to work after marriage and then the pattern switched and more women worked earlier. That switch should, if the model's assumptions are correct, act to reduce the incentive to invest in gender-related beliefs and stereotypes. If women are waiting to learn their type before having children, then they are likely to be less responsive to parental misinformation about their ability level or likelihood of enjoying work.

Explaining the History

The model suggests that parents have an incentive to persuade their daughters that the returns to the workplace are lower, in order to increase the number of grandchildren that they in turn produce. This is, without doubt, a partial story. As in Goldin (this volume) there is also workplace discrimination that occurs when male incumbents suggest that women are bad at particular jobs. But discrimination in the family seems likely to be more powerful, and could potentially shape women's beliefs about their deepest abilities, not just their talents at a particular task.

Can the model help us explain the decline in gender-related stereotypes that seems to have occurred over the course of the twentieth century? We have discussed two potential causes for changing beliefs. One possibility is rising returns to skill also increases the returns for investing in daughters even if those daughters never work, because their skills will translate into more capable grandchildren. This is a possibility, but it does have some trouble with the timing of the changes. We typically think of rising returns to skill as a post-1975 phenomenon, and clearly gender stereotypes changed significantly before that date. The women's movement preceded rather than followed that great widening of wages in the U.S. economy.

The alternative hypothesis is that the timing of women's work and child-bearing changed, which ensured that stereotypes were less important, because women could base their decisions on harder facts, and consequently the incentive to inculcate such stereotypes also declined. Indeed, in almost any sensible model of error, delaying child-bearing would have decrease incorrect beliefs about workplace productivity and decrease the incentives to generate those incorrect beliefs. This suggests that the power of the pill may have been both to increase women's options and help change wider beliefs about women.

VI.   Conclusion

This chapter has discussed different sources of gender stereotypes. We argued that Frieden's model of belief formation by consumer goods producers required assumptions that seemed unlikely to hold during most time periods. Parental persuasion seems a more likely force driving the perpetuation of gender-related beliefs.

We recognize that this model will run counter to the experience of many daughters, who experienced parents who pushed them to succeed, and gave them nothing but positive affirmation of their own talents.   While such occurrences do run counter to the literal structure of the model, we do not believe that they are incompatible with a somewhat richer view of the world.

In many cases, parents may have been more interested in grandchildren quality than in grandchildren quantity. If daughters' human capital, and even workplace success, ended up leading towards more investment in grandchildren, then grandparents would indeed have an incentive to push their daughters towards education and success in the workforce.

Moreover, we have treated parents as the only source of information available to children. Consider a world in which there are a variety of social institutions which broadcast messages about women's ability in work. We may even assume that these institutions exist to cater to parents who want allies in prodding children towards child-bearing. Churches, for example, often seem to have served that role.

If parents believe that their daughters or sons are already exposed to information depicting women as less competent, and they also believe that their daughters will invest too little in themselves if they adopt those social beliefs (even given the parents' pro-grandchildren preferences), then the parents may work against those social beliefs.   For example, assume that the prevalent social belief is that women have ability level $\underline{A}$ and that parents know that their own daughter has ability level $\overline{A}$.   Those parents may not want the daughter to behave as if she knew her full ability level, but they may still want to think that she as an ability level higher than $\underline{A}$.   They will then tell their daughter to disregard the negative stereotype, even if they would prefer it if she thought her own ability was slightly less than $\overline{A}$.

Gender-related beliefs do seem to have had an impact on labor markets and family choices. Those beliefs do not seem to have always been based on reality. We have adopted an economic approach to error that emphasizes the incentives to mislead. We hope that future work develops further models along this line, and does more to subject our model and related theoretical work to serious empirical tests.

References

Arrow, Kenneth J. 1973. "The Theory of Discrimination." In Orley Ashenfelter and Albert Rees, eds. *Discrimination in Labor Markets.* Princeton, N.J.: Princeton University Press, 3-33.

Akerlof , George. 1976. "The Economics of Caste and of the Rat Race and Other Woeful Tales." *The Quarterly Journal of Economics*, 90(4): 599-617.

Axinn, William G., Marin Clarkberg, and Arland Thornton. 1994. "Family Influences on Family Size Preferences." *Demography* 31(1):65-79.

Barber, Nigel. 2000. *Why parents matter: parental investment and child outcomes*. Westport, CT: Bergin & Garvey.

Becker, Gary. 1957. *The Economics of Discrimination.* Chicago: University of Chicago Press.

Becker, Gary S., and Kevin M. Murphy. 2000. *Social Economics: Market Behavior in a Social Environment*. Cambridge, MA: Harvard University Press.

Bereaud, *Susan R.* 1975. "Sex role images in French children's books." *Journal of Marriage and the Family*, 37:194-207.

Blau, Francine, Peter Brummund and Albert Yung-Hsu Liu. "Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding System," *Demography* (forthcoming).

Boustan, Leah P. and William J. Collins. This volume. "The Origin and Persistence of Black-White Differences in Labor Force Participation".

Butcher, Kristin F., and Anne Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *Quarterly Journal of Economics* 109(3):531-563.

Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 102(2):179-221.

Fogel, Robert William.2000. *The Fourth Great Awakening and the Future of Egalitarianism.* Chicago: University of Chicago Press.

Friedan, Betty. *The Feminine Mystique.* New York: W.W. Norton. 1963.

Glaeser , Edward L. 2005. "The Political Economy of Hatred," *Quarterly Journal of Economics*, 120(1): 45-86.

Glaeser, Edward L. and Yueran Ma. 2013. "The Supply of Gender Stereotypes and Discriminatory Beliefs," National Bureau of Economic Research Working Paper No. 19109, June.

Glaeser, Edward L. and Cass Sunstein. 2009. "Extremism and Social Learning," *Journal of Legal Analysis*, 1(1): 263-324.

Goldin, Claudia. 1990. *Understanding the gender gap: An economic history of American women.* Oxford: Oxford University Press.

Goldin, Claudia. 2000. "Labor Markets in the Twentieth Century." In *Cambridge Economic History of the United States, Vol. III, The Twentieth Century*, edited by Stanley J. Engerman and Robert E. Gallman. New York: Cambridge University Press, 549-625.

Goldin Claudia. 2006. "The quiet revolution that transformed women's employment, education, and family." *American Economic Review* 96:1–21

Goldin, Claudia. This volume. "A Pollution Theory of Discrimination: Male and Female Differences in Occupations and Earnings".

Goldin, Claudia, and Lawrence F. Katz. 2000. "Career and Marriage in the Age of the Pill." *American Economic Review* 90(2):461-465.

Gunderson, Elizabeth, Gerardo Ramirez, Susan Levine, and Sian Beilock. 2012. "The Role of Parents and Teachers in the Development of Gender-Related Math Attitudes." *Sex Roles* 66(3/4): 153.

Krueger, Anne O. 1963. "The Economics of Discrimination." *Journal of Political Economy*, 71(5): 481-486

Lazear, Edward P. 1999. "Culture and Language." *Journal of Political Economy*. 107(6): 95-126.

Lerner, Gerda. 1987. *The Creation of Patriarchy*. Oxford: OUP.

Margo, Robert A, 1991. "Segregated Schools and the Mobility Hypothesis: A Model of Local Government Discrimination," *The Quarterly Journal of Economics*, 106(1): 61-73.

Milkman, Ruth. 1982. "Redefining 'Women's Work': The Sexual Division of Labor in the Auto Industry during World War II" *Feminist Studies* 8(2) 336-372.

Myrdal, Gunnar. 1944. *An American Dilemma*: *The Negro Problem and Modern Democracy.* New York: Harper & Brothers.

Olivetti, Claudia. This volume. "The Female Labor Force and Long Run Development: The American Experience in Comparative Perspective".

Rossy, Katherine M. 2011."Politicizing Pronatalism: Exploring the Nazi Ideology of Women through the Lens of Visual Propaganda, 1933-1939." *The Graduate History Review* 3(1).

Simon, Rita J., and Jean M. Landis. 1989. "A Report: Women's and Men's Attitudes about a Woman's Place and Role." *The Public Opinion Quarterly* 53(2):265-276.

Thurow, Lester C. 1969. *Poverty and Discrimination*, Washington D.C.: Brookings Institution.

Weitzman, L. J., Eiffel, D., Hokada, E., & Ross, C. 1972. "Sex-role socialization in picture books for preschool children." *American Journal of Sociology*, *77*:1125-1150.

Table 1: Assumptions and Implications	
Assumption	Implication if Eliminated
# 1: Grandparents desire more grandchildren than the middle generation.	The model's results disappear and grandparents no longer try to persuade their daughters. This is the critical assumption.
# 2: Ability increases the returns from market work more than child-bearing.	If ability complements child-bearing more than work, then assumption # 1 implies that parents would want daughters to think that they are more, not less, able.
# 3: Cash expenditures cannot eliminate the time costs of child-bearing.	Women who are more productive in the workforce might have more children rather than less, this would similarly eliminate the incentive to perpetuate stereotypes.
# 4: Binary Ability Level	Multiple ability types would complicate the model, but not eliminate the basic result that parents want to shade daughters' assessment of their workplace productivity downward.
# 5: Children infer the off-the-equilibrium path assumptions come from parents that would benefit from such a deviation given a narrower range of response (the D1 assumption).	There are multiple equilibria which include the one on which we choose to focus.

Figure 1:
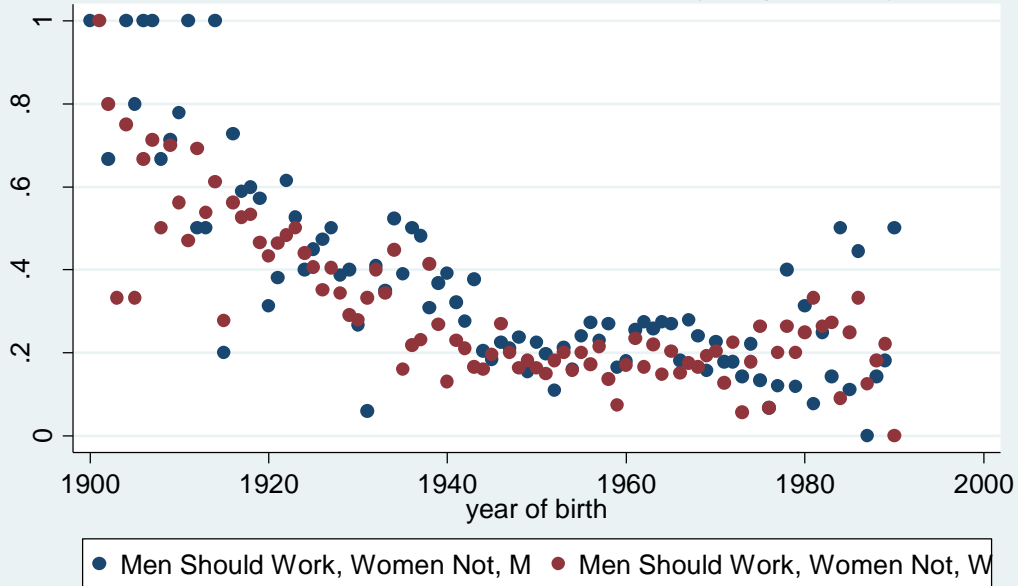**Men Should Work, Women Should Not (Multiple Years)**

- Men Should Work, Women Not, M   • Men Should Work, Women Not, W

Source: General Social Survey
Text: A husband's job is to earn money; a wife's job is to look after the home and family.

Figure 2:
Women Working Does Not Harm Children (Multiple Years)

- Mother's Work OK for Kids, M   • Mother's Work OK for Kids, W

Source: General Social Survey
Text: A pre-school child is likely to suffer if his or her mother works.
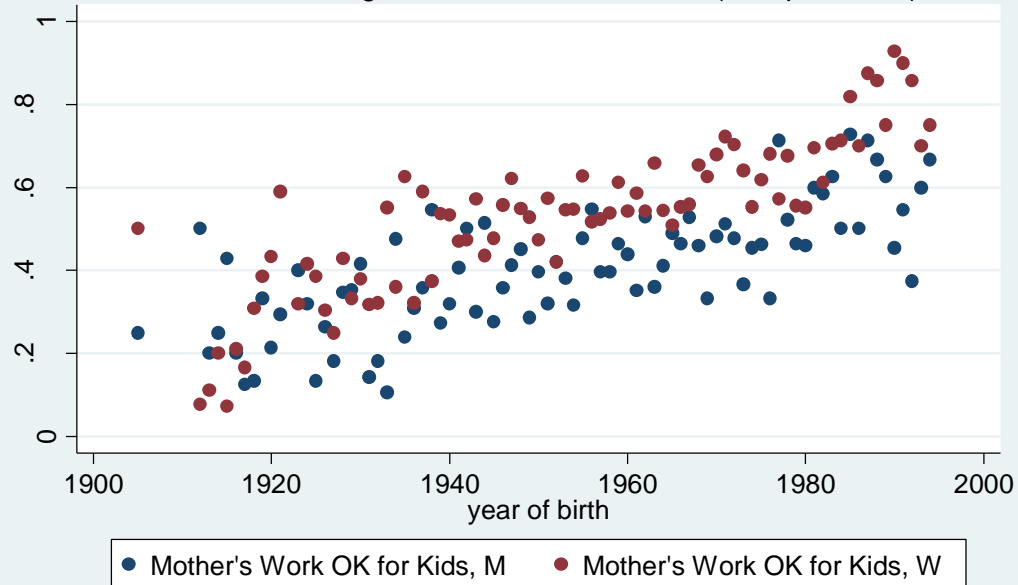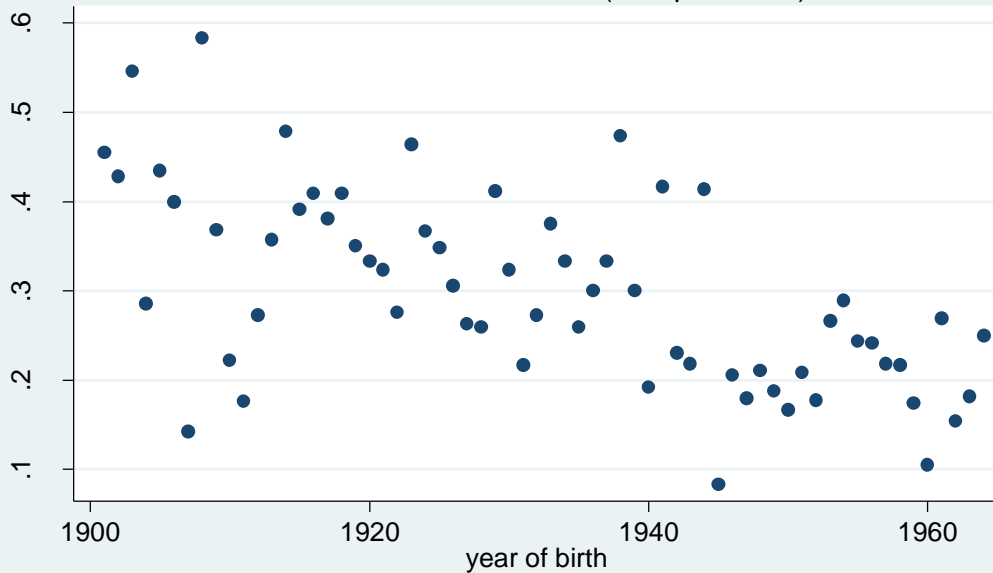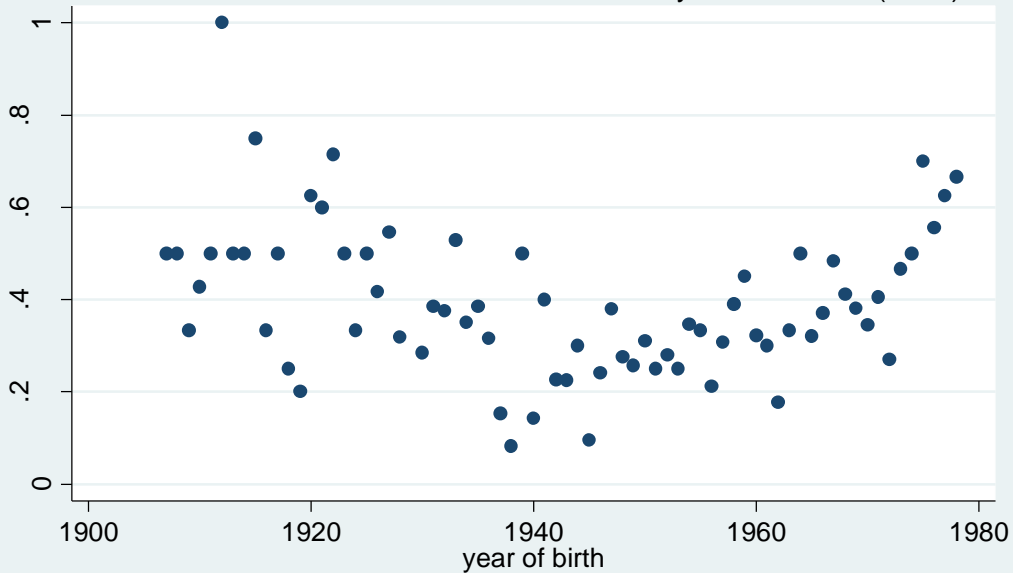
Figure 3:
Men are Better at Politics (Multiple Years)

Source: General Social Survey
Text: Would you say that most men are better suited emotionally for politics than are most women, that men and women are equally suited, or that women are better suited than men in this area?

Figure 4:
Men Earn More Than Women Because They Work Harder (1996)

Source: General Social Survey
Text: Men work harder on the job than women do.
How important do you think this reason is for explaining why women earn less?