



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Women's career choices, social norms and child care policies

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

F. Barigozzi, H. Cremer, K. Roeder (2018). Women's career choices, social norms and child care policies. JOURNAL OF PUBLIC ECONOMICS, 168, 162-173 [10.1016/j.jpubeco.2018.10.004].

Availability:

This version is available at: <https://hdl.handle.net/11585/650256> since: 2022-05-03

Published:

DOI: <http://doi.org/10.1016/j.jpubeco.2018.10.004>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Barigozzi, F., Cremer, H., & Roeder, K. (2018). Women's career choices, social norms and child care policies. Journal of Public Economics, 168, 162-173.

The final published version is available online at:

<https://doi.org/10.1016/j.jpubeco.2018.10.004>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Women's career choices, social norms and child care policies*

Francesca Barigozzi[†] Helmuth Cremer[‡] Kerstin Roeder[§]

October 2016, revised September 2018

Abstract

Our model explains the observed gender-specific patterns of career and child-care choices through endogenous social norms. We study how these norms contribute to the emergence of a gender wage gap. We show that via the social norm a couple's child-care and career choices impose an externality on other couples, so that the *laissez-faire* is inefficient. We use our model to study the design and effectiveness of three commonly used policies. We find that child-care subsidies and women quotas can be effective tools to mitigate or eliminate the externality. Parental leave, however, may even intensify the externality and decrease welfare.

JEL-Classification: D13, H23, J16, J22

Keywords: Social norms, child-care, women's career choices, child care subsidies, women quotas, parental leave

*We thank Claudia Olivetti, Maria Racionero, two reviewers and the coeditor, Johannes Spinnewijn, for their constructive and helpful comments. Francesca Barigozzi dedicates this paper to her mother, the first female Professor of Electrical Engineering in Italy.

[†]University of Bologna, Italy, Email: francesca.barigozzi@unibo.it.

[‡]Toulouse School of Economics, France, Email: helmuth.cremer@tse-fr.eu.

[§]University of Augsburg, Germany, Email: kerstin.roeder@wiwi.uni-augsburg.de.

1 Introduction

While the participation of women in the labor force has increased steadily over the last decades, gender inequalities in the labor market remain substantial; see Blau and Kahn (2017) and Goldin (2006). Significant gender differences in wages, hours of work and occupational choices continue to exist in all OECD countries, where women with a median wage earn on average 15% percent less than their male counterparts. This differential cannot be explained solely by gender differences in schooling, number of hours worked, experience and job characteristics; see, *e.g.*, Blau and Kahn (2017) and Kleven *et al.* (2018).¹ Many studies suggest that the effect of having children on gender gaps, the so called “child penalty”, remains significant, even in developed countries where fertility rates have declined; see Bertrand *et al.* (2010), Goldin (2014b), Kleven and Landais (2017), and Kleven *et al.* (2018). Gender differences are particularly striking when it comes to positions of leadership.²

These career outcomes are mirrored by gender data on hours of work and child-care provisions; the share of couples where both parents work full time is well below 50% in most European countries.³ It is predominantly the women who work part time, while at the same time, are the main providers of child care within the family; see Paull (2008) and Ciccia and Verloo (2012).

Recent lines of research emphasize the role of social norms in shaping women’s choices and outcomes; first within the marriage and the labor markets, and then, within the family (see Bertrand (2011) for a general overview and Bursztyn *et al.* (2017) for a recent study on NBA female students trying to improve their marriage options by “acting wife”). In spite of increased labor market opportunities for skilled women in all countries, negative social attitudes toward working mothers continue to exist. This social norm may cause mothers who work full time to feel guilt when delegating the care of their children to others; see Guendouzi (2006), Rotkirch and Janhunen (2010) and Rose (2017). In the psychology literature this is called “mother’s guilt”. Thus, social norms may provoke the differential sorting of men and women across occupations with women entering low pay occupations with more flexible working hours; see, *e.g.* Card *et al.* (2016).

In this paper, we present a simple model which explains the observed different gender patterns of career and child-care decisions through endogenously determined social norms. Our model shows how these norms contribute to the emergence of a gender wage gap (GWG). We show

¹The factors explaining the gender wage gap have been decomposed into observable differences between men and women and a residual. The residual is often attributed to discrimination: unequal pay for equally qualified work. Recently, Gallen (2018) used a different approach: she measures the gender *productivity gap* by estimating the efficiency units lost in a firm-level production function if a worker is female, holding other explanatory covariates such as age, education, experience, occupation, and hours worked constant.

²See European Commission’s report “More women in senior positions”, 2010.

³See the OECD Family Database.

that through social norms, an individual couple’s child-care and career decisions may impose an externality on other couples so that the (female) labor market sorting observed in the *laissez-faire* (*LF*) equilibrium may be inefficient. Our model provides a theoretical underpinning for the empirical results on gender-specific labor market outcomes, and brings together the various effects which have been documented, showing how they interact and explaining the persistence of gender differences in child-care and career decisions. In particular, our model provides a comprehensive and consistent explanation of the empirical evidence presented by Kleven *et al.* (2018). They show that even though the gender gap in education is now mostly closed in developed countries, gender segregation in the labor market persists, with child penalties explaining an increasing part of gender inequality (up from about 40% in 1980 to about 80% in 2013). They also show that these are related to behaviors and norms inherited from previous generations, and specifically from parents to daughters.

At the same time, our model presents a framework which provides guidance for gender-oriented labor market and child-care policies. In particular, we study the design and effectiveness of three commonly used policy instruments: child-care subsidies, women quotas and parental leave.

We consider a population of parents who choose their career path, child-care arrangements and consumption. There are two possible career paths; a full-time high-career path, and a more flexible low-career path. Individuals who take up the high-career path must work full time and thus cannot provide child care. The low-career path offers more flexible working hours and individuals can freely allocate their time between market labor and child care. The wage rate in the low-career path is the same for all spouses within each couple. The high-career path, by contrast, comes with additional (future) career and earning possibilities which differ across couples and gender. They are perfectly correlated within couples, but the level that can be achieved by the mother may be lower than that available to the father. This income difference applies to persons who have the same abilities and career histories; we shall refer to it as “wage discrimination”. This discrimination (if any) is exogenous in our setting. In line with the empirical evidence, it represents only part of the overall GWG. The rest of this GWG is endogenous and due to child-care and career choices.

Couples are confronted with a social norm concerning child-care activities. The norm derives from the previous generation’s behavior. More precisely, it is determined by child-care decisions made by the median couple of the preceding generation. Deviations from the social norm may impose a cost on the mother who may feel guilt about taking up a high-career path and buying full-time child care on the market, when the majority of mothers in the previous generation personally took care of their children.

With two career paths available to each spouse we have four scenarios to consider, but only

two are relevant in equilibrium. The first is the “traditional couple”, where the mother chooses the low-career path and provides some child care, while the father opts for the high-career path. The second is a couple where both parents take up the high-career path, work full time and rely entirely on market child care. We determine the marginal couple who is indifferent between these two choices, and show that those with lower future high-career earning opportunities opt for the traditional couple arrangement. The sorting pattern of mothers into different career paths thus generates the equilibrium GWG.

The most interesting case to consider is a LF steady state in which the norm is binding, that is, in which the median couple is traditional. Roughly speaking this may correspond with two types of first-best (FB) solutions. In the first scenario, a no-norm allocation would be the efficient solution (without externality), but the economy is “stuck” in the wrong steady state where the norm binds. In the second scenario, the norm is binding both in the efficient solution and in the LF . In both cases, the LF then involves mothers within traditional couples who spend “too much” time on child care, because of the negative externality their choice imposes on high-career couples. For the same reason, the share of traditional couples is “too large”. Furthermore, the GWG in the LF is always inefficiently high.

We show that a uniform child-care subsidy financed by a uniform lump-sum tax is welfare improving in both cases. When there is a no-norm FB, the policy can implement this outcome. In the other case, it cannot reestablish the FB, but it is an efficiency enhancing second-best policy.

Women quotas, are also effective in achieving the efficient no-norm steady state. However, a women quota is ineffective in the case where the norm binds in the efficient solution. While women quotas can affect the share of women in the high-career path, they have no leverage on the level of child care provided by the traditional couple. Consequently, they do not reduce the externality.

Finally, parental leave is essentially a subsidy for informal care, which has the adverse effect of exacerbating the negative externality. However, it has the beneficial effect of allowing the high-career couples who opt for parental leave to freely choose their child care. In other words, it breaks the complementarity between the high-career path and full-time market child care, which otherwise characterizes our setting. This positive effect may or may not outweigh the increase in norm cost.

2 Other related literature

The Introduction has presented evidence on the extent of the GWG and the impact of norms and social attitudes on womens’ career choices. Most of the literature on gender roles and policies

is empirical. We shall now briefly mention some theoretical papers on norms as well as some of the literature dealing with gender policies.

There is some theoretical literature on norms and identity, starting with the pioneering work of Young. Young (1993; 1998 and 2008) interprets social norms as a means to coordinate people's expectations in strategic interactions that possess multiple equilibria. Using evolutionary game theory he shows how small variations in behavior at the individual level can trigger major norm shifts at the societal level. In Akerlof and Kranton (2000; 2010), individuals may suffer a disutility by deviating from the social categories that are associated with their identity (that is, an individual's sense of self), which causes behavior to conform toward those norms. Bénabou and Tirole (2011) endogenize the identity prescriptions, payoffs and cognitive costs discussed by Akerlof and Kranton (2000; 2010) and develop a theory of moral behavior in which people care about "who they are".

The notion of norm that we use here is rather different, and would effectively not qualify as the "norm" in Young's view. In his model, the norm coordinates behavior in a context with multiple equilibria. In our model, it is the norm which effectively *creates* a possible multiplicity of equilibria via the negative externality. Without the externality there would be a single and efficient equilibrium in our setting. Specifically, while norms (such as being polite or having good table manners) are often viewed as a mechanism to avoid negative externalities, they may in fact have the opposite effect; in particular, when it comes to gender relations and equality. This is precisely the view adopted in our paper. While it may be "non-conformist" with respect to much of the theoretical literature on norms, we consider that it may explain, at least in part, many of the findings recurrently documented in the empirical literature such as the persistence of the child penalty. Additionally, as mentioned in the Introduction, "mother's guilt" is a well-documented phenomenon in the psychology literature.

The assumption that the norm is determined by the behavior of a majority of individuals in the previous generation is specific to our paper. However, the main underlying idea is close to that of Lindbeck *et al.* (1999; 2003), where the social norm is also endogenous and increasing in the *share* of the population adhering to it.⁴

There is also an extensive, mostly empirical, literature studying the policies we consider; see Olivetti and Petrongolo (2017) for a recent overview. The most widely discussed policy is women quotas. Bertrand, *et al.* (2018) and Matsa and Miller (2013) show that the GWG decreased for women in leadership positions after the implementation of a women quota in Norway. Child care

⁴Bisin and Verdier (2000; 2001) concentrate on the *optimal* transmission of norms and preferences within the family and assume that children's preferences are acquired through an adaptation and imitation process, which depends not only on the social environment but also on the parents' socialization efforts. Building on these papers, some authors have examined the intergenerational transmission of cultural traits when children are exposed to parental (vertical) and non-parental (oblique) socializations; see Panebianco (2014), and Büchel *et al.* (2014).

subsidies have also been extensively studied. The evidence that they increase the participation of women in the labor force is indisputable; see Averett *et al.* (1997); Kimmel (1998); or Gelbach (2002). Our theoretical paper predicts that women quotas and child-care subsidies not only improve outcomes for women in the labor market, but at the same time help to dissolve costly social norms.

The literature on parental leave is mainly empirical and yields mixed results; see Ruhm (1998), Lalive and Zweimüller (2009), and Felfe and Lalive (2018). Two exceptions are the theoretical papers by Bastani *et al.* (2016) and Del Rey *et al.* (2017). The former shows that, because anti-discrimination legislation requires identical contracts for both genders, firms may use the length of parental leave as an instrument to screen for workers' family-orientation. Del Rey *et al.* (2017) study the effect of leave duration on unemployment and wages in a search and matching model. We identify another, formerly neglected, channel through which parental leave affects women's welfare. Specifically, although parental leave increases the participation of women in the high-career path, it exacerbates the negative externality generated by the social norm.

3 Economic environment

Consider a population of couples with children, the size of which is normalized to one. Each couple consists of a female partner 'f', a male partner 'm', and a given number of children. Couples choose their career path, the mode of child care, and their consumption.

Labor market. There exist two types of career paths (indexed by j). First, a fully engaging high-career path, $j = h$, where individuals who take up this career path have to work an entire day, which we normalize to one. Second, a less demanding low-career path, $j = \ell$, offering flexible working hours, where individuals can freely choose how much time to spend in the labor market. The time not spent at work can be used for child care c_i , where $i = f, m$. Both jobs pay the wage rate y , but the high-career path comes with additional future earning possibilities q_i .⁵ We let $q_m \in [0, \bar{q}]$ and $q_f = \alpha q_m \in [0, \alpha \bar{q}]$, with $\alpha \in (0, 1]$. A level of $\alpha < 1$ generates the unexplained component in the Oaxaca–Blinder decomposition of the GWG; see Blinder (1973) and Oaxaca (1973). It reflects plain discrimination: unequal pay for equally qualified workers, which though declining, continues to be documented in nearly all developed countries. While $\alpha < 1$ adds a measure of realism to the descriptive part of our model, it is not essential for our results, which all continue to hold when $\alpha = 1$. Future revenue q_m is distributed according to

⁵As a real world example, consider the low-career path as a job such that conciliating working-time and family duties is relatively easy but comes at costs of no career opportunities as, for instance, a school or nursery teacher. On the contrary, in the high-career path, promotions are possible if the worker shows to be fully committed and dedicated to the job (also in terms of time physically spent in the company); for instance, a manager can later become chair or CEO of the company.

G with a density g . The median q is such that $G(q^M) = 0.5$. Future earning opportunities are perfectly correlated in a couple.⁶ Consequently, there is a single level of q_f associated with each level of q_m .⁷

Child care. Care for children provided by the spouse(s) is denoted by c_i , while that bought in the private market is denoted by c_p . The latter costs p per unit of time. We let $p = y$, meaning that the current salary of one member in the couple exactly covers the costs of buying full-time child care on the private market.⁸ The children must be taken care of for the entire day, implying $c_m + c_f + c_p = 1$. Couples in which both parents choose the high-career path thus have to fully rely on private child care. When parents enter a flexible job their salary decreases proportionally to the time devoted to care. Informal and private care constitute a family public good and its value to the parents is given by:

$$V(c_m, c_f, c_p) = v(c_m + c_f) + \beta v(c_p),$$

where $v' > 0, v'' < 0$ and $v(0) = 0$. Care provided by the father and mother are thus perfect substitutes while informal and private care are imperfect substitutes, with private care being (weakly) less welfare-enhancing than informal care, $\beta \in (0, 1]$.⁹ Apart from child care, each parent derives utility from the consumption of their labor income.

Social norm. Couples are confronted with a social norm concerning child-care activities. The norm derives from the previous generation's behavior. In particular, if in the previous generation the median mother chose the low-career path and (partly) looked after the children, then this choice represents the social norm for the current parents.¹⁰ Deviations from the social norm are costly. Mothers may feel guilt about taking up the high-career path and buying full-time child care on the private market if the majority of mothers in the previous generation personally took care of their children.

Formally, we represent the social norm for mothers belonging to generation t as a cost of the

⁶Assortative mating is commonly observed and has been increasing over the last decades; see Schwartz and Mare (2005).

⁷There is a literature studying how career choices affect the outcome in the marriage market; see Fernandez *et al.* (2004), Bertrand *et al.* (2016) and Bursztyn *et al.* (2017). We do not consider this issue. Career decisions are not made once and for all, and the choices made earlier in life can be adjusted as conditions change and couples are already formed; career choices and child-care decisions are then intertwined. In other words, the career choices we study are likely to be taken “downstream” from the marriage market. While it is statistically true that a significant proportion of “our” young parents will eventually get divorced, it seems to us that the decisions we consider are likely to be affected only marginally by the mothers' future position in the marriage market.

⁸This assumption is simply a normalization that has no relevance for our results. Without it we would obtain a term proportional to $(p - y)$ in the first-order conditions with respect to child care. This would affect the equilibrium levels of child care but otherwise all other results are not affected.

⁹See, for instance, Gregg *et al.* (2005), Bernal (2008), and Huerta *et al.* (2011).

¹⁰In all countries, the vast majority of both men and women hold the view that a woman should not work full time as long as she has very young children living at home. Kleven *et al.* (2018) present evidence that such a norm is transmitted through generations, from parents to their daughters. On the transmission of social norms between generations see also Fernandez (2007), Fernandez and Fogli (2009), and Fernandez *et al.* (2004).

full-time job given by $\gamma(\max\{0; c_f^{M,t-1} - c_f^t\})$, where $c_f^{M,t-1}$ is time spent with children by the median woman in the previous generation. The parameter $\gamma \in [0, 1]$ reflects the costs of norm deviations.¹¹

In the *LF*, we concentrate on the case where the median mother in the previous generation entered the flexible job market and took care of the children.

Assumption 1 (Social norm active in period t) *In the LF steady state a majority of couples are traditional, implying $c_f^{M,t-1} > 0$.*

We will focus on decisions made in the steady state. Thus, we omit the period superscript for all variables except for child care provided by the median mother in the previous generation, which is indexed by “ $M, t - 1$ ”.

Timing. The timing of couples’ decisions is as follows: first, parents choose their career path and then, in the second stage, they choose consumption and the amount of child care (be it formal or informal). Parents act cooperatively and maximize the sum of their utilities.

4 Couple’s optimization

We first analyze the choice of child-care activities given the career path and then, by proceeding backward, we consider the couple’s choice. The two-stage approach of the couple’s decision making process is adopted for the sake of presentation. Because no new information is revealed between the two stages and nothing else changes, it yields the same outcome as a simultaneous choice.

4.1 Second Stage: Child-care activities

Since each spouse chooses between two career paths, there are four potential types of couples. However, only two of these are relevant in equilibrium, the traditional couple, where only the father enters the high-career path while the mother enters the flexible job market, and the high-career couple, where both parents take up the high-career path.

4.1.1 Only the father enters the high-career path

This scenario exactly replicates the social norm and the mother does not suffer a norm cost, that is $\gamma(\max\{0; c_f^{M,t-1} - c_f^t\}) = 0$. Since the father took up the high-career path he is not able to take care of the children, and $c_m^* = 0$. Welfare of this couple is denoted by W_ℓ (because the

¹¹The assumption that the cost of the social norm is linear in the size of the deviation is without generality loss.

female partner enters the low-career path). Noting that $c_f + c_p = 1$ the couple chooses child care provided by the female partner, c_f , to maximize:¹²

$$\max_{c_f} W_\ell = y + (1 - c_f)y + q - p(1 - c_f) + v(c_f) + \beta v(1 - c_f).$$

Given that $p = y$, the first-order condition with respect to c_f can be written as:

$$c_\ell^* \equiv c_f^* : \quad v'(c_\ell^*) = \beta v'(1 - c_\ell^*), \quad (1)$$

where c_ℓ^* is the total amount of child care provided by couple ℓ . From (1), marginal utility from informal child care equals the marginal benefit from private care.

The traditional couple's welfare is given by:

$$W_\ell^* = y + q + v(c_\ell^*) + \beta v(1 - c_\ell^*), \quad (2)$$

where its optimal consumption is given by $y + q$.

4.1.2 Both parents enter the high-career path

When both parents enter the high-career path their common earnings amount to $2y + q_m + q_f = 2y + q(1 + \alpha)$. Neither of the parents is able to provide child-care services so that $c_p^* = 1$. The mother suffers psychological costs equal to $\gamma(\max\{0; c_f^{M,t-1} - 0\}) = \gamma c_f^{M,t-1}$. A high-career couple simply consumes its income. Welfare is denoted by W_h (because the female partner enters the high-career path). Noting that $p = y$, welfare can be written as:

$$W_h^* = y + q(1 + \alpha) + \beta v(1) - \gamma c_f^{M,t-1}, \quad (3)$$

where the couple's optimal consumption is given by $y + q(1 + \alpha)$.

4.1.3 Other scenarios

The other two scenarios are clearly not relevant.¹³ Reversing the norm is never optimal. In other words, a scenario in which only the mother enters the high-career path is always dominated by that in which only the father enters the high-career path, which involves no norm costs and higher future benefits. Similarly, having both parents entering the flexible job market can never be optimal since the couple then forgoes future benefits q_m .

¹²The assumption that families' utility functions are linear in consumption does not affect the qualitative results for the *LF* and the *FB*. However, its implementation would then require a system of personalized lump-sum transfers and taxes, so that our simple instruments would no longer be sufficient to implement a *FB*. Furthermore, the regressivity of the considered instruments would affect their optimal second-best level; see the Conclusion for more discussion.

¹³For a detailed analysis see the online Appendix.

4.2 First Stage: Job market decision

At the first stage, the couple compares its welfare levels and chooses its career path such that the couple's welfare is maximized. Formally, we must compare the traditional couples' welfare with the welfare of couples in which both parents enter the high-career path:

$$W_h^* = y + q(1 + \alpha) + \beta v(1) - \gamma c_f^{M,t-1} \stackrel{\leq}{\geq} W_\ell^* = y + q + v(c_\ell^*) + \beta v(1 - c_\ell^*). \quad (4)$$

High-career couples enjoy a larger utility from (future) consumption because $y + q(1 + \alpha) > y + q$, but a lower utility from child care than type- ℓ couples. The mothers' career choices thus depend on their labor market opportunities αq .

The condition $W_h^* = W_\ell^*$ implicitly defines the marginal couple in period t , \hat{q}_t , as a function of that in period $t - 1$, \hat{q}_{t-1} . Using Equation (4) and imposing the tie-breaking rule that when a couple is indifferent between the career paths the mother chooses the high-career one, yields:

$$\hat{q}_t = \begin{cases} A \equiv \frac{1}{\alpha} [v(c_\ell^*) + \beta [v(1 - c_\ell^*) - v(1)] + \gamma c_\ell^*] & \text{if } \hat{q}_{t-1} > q^M, \\ B \equiv \frac{1}{\alpha} [v(c_\ell^*) + \beta [v(1 - c_\ell^*) - v(1)]] & \text{if } \hat{q}_{t-1} \leq q^M. \end{cases} \quad (5)$$

The expression for A gives the marginal couple in a given period, when the marginal couple in the previous period was to the right of q^M , so that the social norm is binding (and determined by c_ℓ^*). The expression for B specifies the marginal couple when the social norm is not binding. Obviously, $A > B$.

The dynamics of the marginal couple for the cases compatible with Assumption 1 are illustrated by Figure 1. It assumes that $q^M = \bar{q}/2$, but this is just for the sake of illustration and is of no relevance for the arguments. Panel (a) obtains when $A > B > q^M$; in this case there is a single steady state with a binding norm, which is globally stable. In Panel (b), with $A > q^M > B$, there are two steady states. The larger one implies a binding norm while the norm does not bind at the smaller steady state equilibrium. Both steady states are locally stable; which one is achieved depends on the initial level of \hat{q} .¹⁴

The main results obtained so far are summarized in the following proposition.

Proposition 1 (The marginal couple in the LF steady state) *In the LF steady state:*

- (i) *the marginal couple (that is, the couple where the mother is indifferent between the high- and the low-career path) is defined by the following value of future job market opportunities*

\hat{q}^* :

$$\hat{q}^* = \frac{1}{\alpha} [v(c_\ell^*) + \beta [v(1 - c_\ell^*) - v(1)] + \gamma c_\ell^*], \quad (6)$$

¹⁴Since the dynamics are simple, we do not provide the phase diagrams. A simple inspection of the diagram shows that in Panel (b) when we start from $\hat{q}_0 > q^M$, the larger steady state is achieved within a period, and the situation for $\hat{q}_0 < q^M$ is exactly symmetric. In Panel (a) $\hat{q}_0 > q^M$ brings us again to the steady state in a single period, while 2 periods are necessary when $\hat{q}_0 < q^M$.

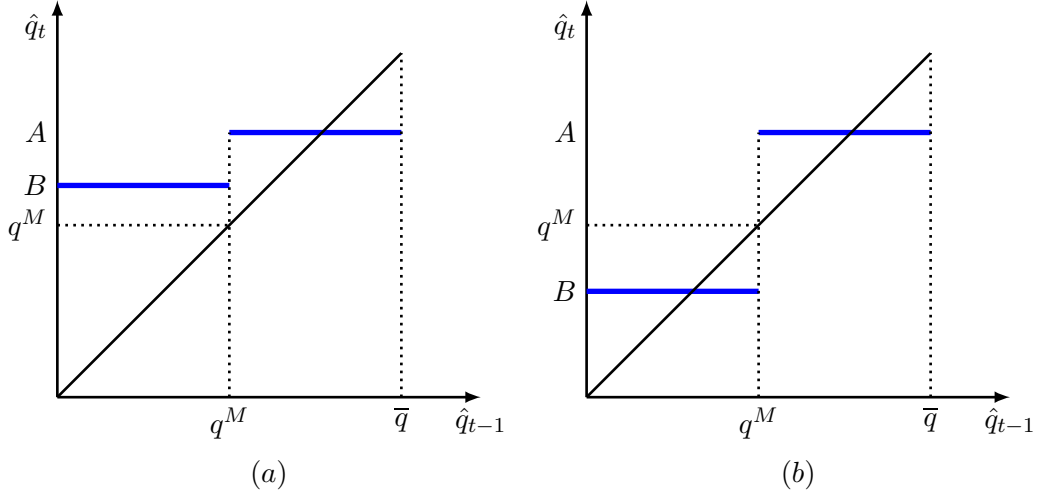


Figure 1: *Dynamics of the marginal couple.* The horizontal segment at $A \equiv \frac{1}{\alpha} [v(c_\ell^*) + \beta [v(1 - c_\ell^*) - v(1)] + \gamma c_\ell^*]$ represents the marginal couple in period t when the norm is binding ($\hat{q}_{t-1} > q^M$) while $B \equiv \frac{1}{\alpha} [v(c_\ell^*) + \beta [v(1 - c_\ell^*) - v(1)]]$ represents the marginal couple in period t when the norm is not binding. The steady states are at the intersection(s) of the horizontal segment(s) and the 45 degree line. In Panel (a) there is a single (globally stable) steady state (with a binding norm). In Panel (b) there are two locally stable steady states. Assumption 1 implies that the higher one is achieved.

where $q^M < \hat{q}^*$ and c_ℓ^* is determined by equation (1). In couples with $q \geq \hat{q}^*$, the mother chooses the high-career path and in couples with $q < \hat{q}^*$, the mother chooses the low-career path;

(ii) female participation in the high-career path is higher (\hat{q}^* is lower), the higher α (less discrimination) or β (better market care) and the lower γ (norm cost).

Part (ii) is obtained by differentiating (6) and by taking (1), which determines c_ℓ^* , into account. Differentiating the latter equation shows that c_ℓ^* does not depend on α or γ , but decreases with β . Consequently, the effects of α and γ on \hat{q}^* are simply given by the partial derivatives, while for β the indirect effect via c_ℓ^* reinforces the direct effect.¹⁵

The GWG is defined as the *difference in total income earned by women and men* in the steady state equilibrium and is given by:

$$GWG = \int_0^{\bar{q}} [y + q]g(q)dq - \left[G(\hat{q}^*)[y(1 - c_\ell^*)] + \int_{\hat{q}^*}^{\bar{q}} [y + \alpha q]g(q)dq \right].$$

¹⁵The envelope theorem implies that:

$$\frac{\partial \hat{q}^*}{\partial c_\ell^*} = \gamma > 0.$$

With $\partial c_\ell^*/\partial \beta < 0$ and $\partial \hat{q}^*/\partial \beta < 0$ we then have that the total effect is given by $d\hat{q}^*/d\beta < 0$.

The following proposition summarizes the results obtained in this section and provides a decomposition of the GWG.

Proposition 2 (Characterization of the LF) *When a social norm affecting those mothers who do not provide child care exists, $\gamma > 0$, and/or the job market suffers from unequal opportunities, $\alpha < 1$, then:*

- (i) *it is never optimal for the father to take up the low-career path;*
- (ii) *couples with job opportunities higher or equal to the threshold \hat{q}^* choose the high-career path for both parents;*
- (iii) *the GWG decomposes as follows*

$$GWG = \underbrace{G(\hat{q}^*)yc_\ell^*}_{\text{child penalty}} + \underbrace{\int_0^{\hat{q}^*} qg(q) dq}_{\text{sorting}} + \underbrace{\int_{\hat{q}^*}^{\bar{q}} (1 - \alpha)qg(q) dq}_{\text{plain discrimination}}. \quad (7)$$

Recall that, in our model, all women are active in the labor market and there is no heterogeneity in education. Consequently, the GWG consists of gaps in hours worked because of family duties and in return to labor supplied in sectors where men and women are employed. The first term in (7) thus represents the pure “child penalty”: mothers in traditional couples do not work full time because they provide some child care; see Bertrand *et al.* (2010), Goldin (2014a), Kleven *et al.* (2018).¹⁶ The second term of the GWG accounts for the extra revenue earned by men because they all choose the high-career path. Interestingly, this second term is affected by social norms and child-care decisions through \hat{q}^* , hence our model offers a consistent explanation of how social pressure and/or persisting inequality in the labor market explain the small share of women in leading positions together with lower wages. Finally, the last term in (7) captures the unexplained component of the GWG in the Oaxaca–Blinder decomposition, or the plain discrimination part; it vanishes when $\alpha = 1$.

Before turning to policy design, we determine the efficient share of female participation in the high-career path. This benchmark will turn out to be crucial when determining the effectiveness of the considered policies.

5 First-best and inefficiency of *laissez-faire*

Not only the cost of the social norm but also the norm itself is endogenous. It disappears if, in the previous generation, the majority of women entered the high-career path. Hence, a suitably

¹⁶One exception to the “child penalty” evidence is Gallen (2018). Using Danish data she finds that both mothers and non-mothers are paid less than men, but the (low) relative pay of mothers is completely explained by productivity differences. In contrast, women without children are estimated to be as productive as men but are paid less.

designed policy might be able to eradicate the norm, but we show below this is not necessarily optimal.

In this section, we characterize the FB in order to describe the inefficiencies created by the social norm in the *LF*. Two types of benchmark cases exist: the one in which the social norm is binding in the steady state; and the one where it is not binding. One might at first think that with unrestricted instruments and information it is always optimal to eradicate the norm. However, intuitively, this implies that more than half of the couples choose a high-career path which may not be optimal, for instance, when domestically provided child care is valued sufficiently higher than market care (β very small), or when discrimination is particularly high (α very small). The nature of the FB that obtains is important: when the FB is norm-free, even the restricted instruments we consider will be sufficient to restore efficiency. This is no longer true when the norm binds in the FB in which case we have to consider a second-best policy.

We consider a utilitarian social welfare function which is given by the (unweighted) sum of steady state utilities of all households.¹⁷ Recall that a job market allocation specifies the amount of child care provided by traditional couples and the identity of the marginal couple, which, in turn, determines whether or not the social norm is binding in the steady state. Hence, we have to derive c_ℓ and \hat{q} that maximize the following social welfare function:

$$\begin{aligned} \max_{c_\ell, \hat{q}} SW = & \int_0^{\hat{q}} [y + q]g(q) dq + G(\hat{q}) [v(c_\ell) + \beta v(1 - c_\ell)] \\ & + \int_{\hat{q}}^{\bar{q}} [y + q(1 + \alpha)]g(q) dq + (1 - G(\hat{q})) [\beta v(1) - \gamma c_f^M], \end{aligned} \quad (8)$$

where $c_f^M = c_\ell$ if $\hat{q} > q^M$, that is if the social norm is binding. If, instead $\hat{q} \leq q^M$, the social norm is not binding and $c_f^M = 0$. We denote the solution to the first scenario as (c_ℓ^n, \hat{q}^n) and a norm-free solution as (c_ℓ^o, \hat{q}^o) .

Let us first consider the case where the norm is not binding. The solution to (8) is then characterized by the following two first order conditions:

$$v'(c_\ell^o) = \beta v'(1 - c_\ell^o), \quad (9)$$

$$\hat{q}^o = \frac{1}{\alpha} [v(c_\ell^o) + \beta [v(1 - c_\ell^o) - v(1)]]. \quad (10)$$

This solution is only meaningful when (10) implies $\hat{q}^o \leq q^M$. While Figure 1 represents the *LF* it is also useful for understanding the FB. Comparing (1) and (5) to (9)–(10) shows that this outcome corresponds exactly to the lower steady state represented in Panel (b). This is not surprising: when a no norm *LF* steady state exists and is attained there is no reason why it should not be efficient. The problem is that when we start from a situation where the norm

¹⁷This is for ease of exposition only. With quasi-linear preferences, the properties of the FB we derive are effectively valid for *all* Pareto-efficient allocations.

binds, which is what Assumption 1 states, without policy intervention, the economy gets “stuck” in the wrong, inefficient steady state with a binding norm.

We now turn to the case where the norm is binding in the FB. The solution to (8) is then characterized by the following two first-order conditions:

$$v'(c_\ell^n) = \beta v'(1 - c_\ell^n) + \gamma \frac{1 - G(\hat{q}^n)}{G(\hat{q}^n)}, \quad (11)$$

$$\hat{q}^n = \frac{1}{\alpha} [v(c_\ell^n) + \beta [v(1 - c_\ell^n) - v(1)] + \gamma c_\ell^n]. \quad (12)$$

Since here the social norm is active in the steady state, we must have a higher share of traditional couples in the population so that $q^M < \hat{q}^n$.

Comparing the *LF* (Equation 1) with the FB level of child care (Equation 11) shows that the marginal costs of informal care provision (the RHS of Equations 1 and 11) are higher in the FB than in the *LF*. This implies that traditional couples provide less child care in the FB than in the *LF*: $c_\ell^n < c_\ell^*$. Intuitively, child care provided by traditional couples imposes a negative externality, measured by the term $\gamma(1 - G(\hat{q}^n))/G(\hat{q}^n)$, on all high-career couples.¹⁸

This difference in child care also has a bearing on the marginal couple. Equation (6) coincides, for a given c_ℓ , with the condition determining the marginal couple in the FB. However, since:

$$c_\ell^* = \arg \max \{v(c_\ell) + \beta v(1 - c_\ell)\}, \quad (13)$$

and v is concave, we necessarily have that the RHS of (12) is smaller than the RHS of (6) implying $\hat{q}^n < \hat{q}^*$.

The two scenarios, (c_ℓ^n, \hat{q}^n) and (c_ℓ^o, \hat{q}^o) describe interior solutions in the sense that $\partial SW/\partial c_\ell = \partial SW/\partial \hat{q} = 0$ for a given level of c_f^M . Specifically, we have $c_f^M = c_\ell^n$ in the binding norm case and $c_f^M = 0$ in the non-binding norm case. This implies that couple \hat{q}^n is effectively indifferent between the two career paths when $c_f^M = c_\ell^n$, while couple \hat{q}^o is indifferent for $c_f^M = 0$.

However, we cannot rule out the case where the distribution of q 's is such that $q^M < \hat{q}^o$ at c_ℓ^o and $c_f^M = 0$. In other words, maximizing (8) with respect to \hat{q} for $c_f^M = 0$ may yield a solution which is larger than q^M . This, in turn, is inconsistent with $c_f^M = 0$. To have consistency, we then have to consider a constrained solution where we impose $\hat{q} = q^M$.¹⁹ This amounts to assigning all couples with $q \in [q^M, \hat{q}^o)$ to the high-career path so that the norm is indeed not binding.

¹⁸The no-norm optimum and the binding-norm *LF* yield the same level of child care; see (9) and (1). This is because in the first case the externality is not relevant while in the second case it is ignored (which is inefficient given that the norm is binding). While this comparison across regimes is not in itself interesting (and does *not* mean that c_ℓ^* is efficient), we mention this property for future reference because it will be relevant to assess welfare in the transition generation below.

¹⁹Formally, this can be achieved by stating the optimization for $c_f^M = 0$ as a Kuhn-Tucker problem imposing the constraint that $q^M \geq \hat{q}$ associated with a multiplier $\lambda \geq 0$, so that the FOC of the Kuhn-Tucker expression with respect to \hat{q} is given by $\partial SW/\partial \hat{q} = \lambda$. This yields the unconstrained solution with $\hat{q} = \hat{q}^o$ when $\lambda = 0$, while the constrained solution with $\hat{q} = q^M$ (and $\partial SW/\partial \hat{q} > 0$) obtains when $\lambda > 0$.

Couples with $q \in [q^M, \hat{q}^o)$ would prefer the low-career path if they were free to choose. However, in a FB world they can be assigned to a different path. Such a scenario is optimal when it yields a higher welfare than that achieved with a binding norm, that is with (c_ℓ^n, \hat{q}^n) .²⁰ For the remainder of the paper we shall concentrate on the two other cases. Like the binding norm FB, the constrained FB cannot be implemented by the instruments we consider. Consequently, the important distinction is between the (unconstrained) no norm FB, which can be implemented at least by some of the instruments, and the other FBs (the binding norm FB and constrained no-norm one) which imply that we remain in a second-best setting so that the exact nature of the FB is no longer relevant.

The main results obtained in this section are summarized in the following proposition.

Proposition 3 (Welfare analysis of the LF with a binding norm) *Comparing the LF equilibrium described in Proposition 1 with the FB:*

- (i) *informal child care in LF, c_ℓ^* , is inefficiently high because of the negative externality it exerts on high-career mothers through the social norm;*
- (ii) *female participation in the high-career path is always inefficiently low in LF;*
- (iii) *the GWG is inefficiently high in the LF.*

Point (iii) requires some explanations. For any given q , in the LF, the female spouse's earnings are less than or equal to her FB earnings. Specifically, the component of the GWG due to child penalties is reduced in the binding norm FB because women's labor income increases due to the fall in informal child care ($c_\ell^n < c_\ell^*$). In addition, no matter the nature of the FB allocation, the GWG from adverse sorting decreases because more women now enter the high-career path and benefit from future prospects (\hat{q}^o and \hat{q}^n are strictly lower than \hat{q}^*).

The following proposition describes how the various parameters of our model affect the efficient solution. Noting that a no-norm optimum requires $\hat{q}^o < q^M$, and the q^M depends solely on the distribution of q but not on the other parameters, the comparative statics properties of (9)–(10) yield the following Proposition:²¹

Proposition 4 (No-norm vs binding norm in the FB) *A no-norm FB is more likely to prevail than a binding norm optimum, the larger (i) q^M , (ii) α , or (iii) β .*

²⁰The counterpart to this case with a constrained binding norm is when the maximization of (8) yields $q^M > \hat{q}^n$ for $c_f^M = c_\ell^n > 0$ which is not possible. It would require setting $\hat{q} > q^M$, forcing some couples into their less-preferred career path in order to create a binding norm and thus a negative externality. This solution is clearly not optimal; it is necessarily dominated by the constrained no-norm allocation.

²¹As in Proposition 1 (ii), we have to account for the indirect effect of β via c_ℓ , but from the envelope theorem this is not a first-order effect here.

When q^M is large, then $\hat{q}^o \leq q^M$ will hold for a larger set of the other parameters (namely α and β) so that the non-binding norm solution becomes more likely. Additionally, the binding norm solution becomes more costly, and thus the no-norm solution becomes more appealing, when more mothers tend to choose the high-career path. This is also the case when the labor market is close to equality, or formal child care arrangements are high quality (α or β are large). Finally, the norm cost γ has no impact on the existence of an (unconstrained) no-norm solution. However, since social welfare with a binding norm decreases as γ increases, a higher norm cost does make a binding norm optimum less likely. These two observations, may at first appear to be inconsistent, but they are simply due to the fact that as γ becomes sufficiently large (while $\hat{q}^o > q^M$) the constrained no-norm solution will emerge.

Our model thus suggests that in economies with low gender discrimination (large α) and a high-quality child-care system (large β), the eradication of the social norm is likely to be optimal; all of these factors concur to increase social welfare when the norm disappears.

The next section shows that the nature of the FB affects the effectiveness of the considered child-care and gender policies in a rather striking way. Specifically, when the FB is no-norm, two of the three simple instruments we consider are sufficient to achieve (decentralize) this solution. Consequently, Proposition 4 and the subsequent discussion are important in order to assess the practical policy implications of our model. In Section 7, we illustrate these predictions by showing how existing empirical results and stylized facts can be used to determine which type of FB can be expected to be relevant in specific countries or types of countries.

6 Welfare improving policies

We now analyze how the government can improve efficiency by focusing on policy design. Recall that as per Assumption 1, we concentrate on situations where the initial steady state involves a binding norm and is thus inefficient. Specifically, we study the effects on women's career choices and child care provision of three policies currently used in the real world; namely, (*i*) uniform child care subsidies, (*ii*) women quotas and (*iii*) parental leave.

In analyzing these policies, we must differentiate between situations where the FB can be implemented with the considered instrument and those where it cannot be achieved so that we are in a second-best setting. We concentrate on the steady state, but when the FB can be achieved we also calculate the transition utility in order to determine which of the eligible policies yields the larger transition utility.

6.1 Uniform subsidy on formal child care

We first consider a uniform subsidy s on market child care reducing its price to $p - s$.²² Assume that the subsidy is financed by a uniform lump-sum tax T levied on all couples. The consumption level for high-career couples is then given by $y + q(1 + \alpha) + s - T$, while it is $y + q + s(1 - c_\ell^S) - T$ for low-career couples. With a subsidy on private care, optimal informal child care, denoted by c_ℓ^S , by type- ℓ couples is implicitly determined by:

$$v'(c_\ell^S) = \beta v'(1 - c_\ell^S) + s. \quad (14)$$

A subsidy on market care increases the marginal costs of informal care (RHS of Equation 14) and informal care will be lower than in LF : $c_\ell^S < c_\ell^*$.

6.1.1 Social norm is non-binding in the FB

This corresponds to the case represented in Panel (b) of Figure 1. There are two potential steady states, one of which is efficient. However, the economy is stuck in the inefficient state because historically, the majority of couples has been of the traditional type where the mother chooses the flexible career path.

In this case, a transitional policy implemented for a single period (generation) is sufficient to achieve the efficient steady state. More precisely, it is sufficient to determine s so that the marginal couple is given by q^M . Hence, s solves:

$$q^M = \frac{1}{\alpha} \left[v(c_\ell^S) + \beta [v(1 - c_\ell^S) - v(1)] + \gamma c_f^{M,t-1} - s c_\ell^S \right]. \quad (15)$$

Budget balance requires:

$$T = s [1 - G(q^M) c_\ell^S]. \quad (16)$$

Condition (15) ensures that given the costs of the social norm $\gamma c_f^{M,t-1} = \gamma c_\ell^*$, the marginal couple goes down to q^M . Consequently, in the next period $c_f^{M,t} = 0$, and we are in the FB steady state described by (9)–(10) and no further intervention is necessary, because convergence is obtained with one period delay (or after one transitory generation).

Note, however, that the solution during the transition period is inefficient. In particular, child-care provision c_ℓ^S is inefficiently low. The norm is still binding, but because it will no longer be binding in the subsequent period its efficient level would be $c_\ell^* = c_\ell^o > c_\ell^S$; see Footnote 18.

It can be easily checked that welfare within the transition period is given by:

$$SW^S = \alpha [1 - G(q^M)] E_h[q] + E[q] + y + G(q^M) [\beta v(1) + v(c_\ell^S) + \beta v(1 - c_\ell^S) - \gamma c_\ell^*], \quad (17)$$

²²The case where $s = p$ can be interpreted as free (possibly public) provision of child care.

where $E[q]$ and $E_h[q]$ are defined as:

$$E[q] \equiv \int_0^{\bar{q}} qg(q)dq \quad \text{and} \quad E_h[q] \equiv \frac{\int_{q^M}^{\bar{q}} qg(q)dq}{1 - G(q^M)}. \quad (18)$$

6.1.2 Social norm is binding in the FB

In this case the uniform subsidy is no longer sufficient to implement the FB. We then study the second-best policy assuming an interior solution and concentrate on the steady state. With a binding norm we have $c_f^M = c_\ell^M = c_\ell^S$, and the welfare function is given by:

$$\begin{aligned} \max_{s,T} SW^S = & \int_0^{\hat{q}^S} [y + q + s(1 - c_\ell^S) - T]g(q)dq + G(\hat{q}^S)[v(c_\ell^S) + \beta v(1 - c_\ell^S)] \\ & + \int_{\hat{q}^S}^{\bar{q}} [y + q(1 + \alpha) + s - T]g(q)dq + (1 - G(\hat{q}^S)) [\beta v(1) - \gamma c_\ell^S], \end{aligned} \quad (19)$$

subject to the budget constraint which is now given by:

$$T = s [1 - G(q^S)c_\ell^S]. \quad (20)$$

Child-care provided by traditional couples continues to be determined by Expression (14) and \hat{q}^S is defined by:

$$\hat{q}^S = \frac{1}{\alpha} [v(c_\ell^S) + \beta [v(1 - c_\ell^S) - v(1)] + (\gamma - s)c_\ell^S]. \quad (21)$$

Substituting (20) into the welfare function, the first-order condition of (19) with respect to s can be written as:

$$s(\hat{q}^S) = \gamma \frac{1 - G(\hat{q}^S)}{G(\hat{q}^S) + g(\hat{q}^S)c_\ell^S \frac{\partial \hat{q}^S / \partial s}{\partial c_\ell^S / \partial s}}. \quad (22)$$

Since $\partial c_\ell^S / \partial s < 0$ and $\partial \hat{q}^S / \partial s = (1/\alpha)[-c_\ell^S + \gamma \partial c_\ell^S / \partial s] < 0$, Equation (22) implies:

$$s(\hat{q}^S) < s^P(\hat{q}^S) \equiv \gamma \frac{1 - G(\hat{q}^S)}{G(\hat{q}^S)}. \quad (23)$$

In words, the optimal s is smaller than the Pigouvian subsidy s^P that restores efficiency of informal child care for a given level of \hat{q}^S . The Pigouvian tax rule, $s^P(\hat{q})$, is obtained by equating (11) and (14), and it depends on \hat{q} since the costs of the externality depend on the number of high-career couples. From (23) we see that, given \hat{q}^S , the subsidy on market care is set at a lower level than the Pigouvian subsidy. This is because a uniform subsidy benefits high-career couples more than traditional ones. For high-career couples, market care is given and equal to one so that the subsidy represents a windfall gain. Consequently, the policy will distort \hat{q} downwards which was otherwise optimally chosen in *LF* for any given level of the traditional couples' child care. Observe that this comparison is based on tax *rules* (that is, given q^S), the

first- and second-best *levels* of the subsidy, $s^P(\hat{q}^n)$ and $s(\hat{q}^S)$, cannot be compared since \hat{q}^S and \hat{q}^n differ.

The previous subsection has shown that when a no-norm first best exists, it can be implemented by a subsidy that applies during the transition period only. The second-best policy that is relevant in the binding-norm case, on the other hand, must be applied on a permanent basis. The dynamics are again simple and the new steady state is achieved after a single period. However, if the policy were abandoned, the economy would return to the initial steady state.

Recall that here and throughout the paper we focus on efficiency only; with quasi linear preferences redistribution does not matter. However, it is clear that the child-care subsidy we considered is regressive and this effect may be confined to the transition generation, or be permanent; see the Conclusion for further discussion.

The following proposition summarizes our main results.

Proposition 5 (Uniform subsidy on formal child care) *Consider a uniform subsidy on child care financed by a uniform lump-sum tax:*

- (i) *when the social norm is not binding in the efficient allocation, this FB can be achieved by a uniform subsidy which is implemented for one period only and is set to make the marginal couple coincide with the median one. In the transition period in which the subsidy is imposed, informal child care is inefficiently low. In the subsequent periods, efficiency is fully restored;*
- (ii) *when the social norm is binding in the efficient allocation, we have a second-best solution. The uniform subsidy is implemented indefinitely. It mitigates the norm costs by reducing the median couple's informal child care provision. Efficiency is only partially restored; informal child care is lower than in the LF but (given \hat{q}^S) it is larger than efficient.*

6.2 Women quotas

We now consider a different policy, namely a women quota (WQ), which requires that the number of women in the high-career path, $1 - G(\hat{q}^{WQ})$, out of the total number of high-career employees, $1 + 1 - F(\hat{q}^{WQ})$, is not lower than r , that is:

$$r \leq \frac{1 - F(\hat{q}^{WQ})}{2 - F(\hat{q}^{WQ})}, \quad (24)$$

where \hat{q}^{WQ} denotes the marginal couple under a WQ policy. We argue that employers who are confronted with a WQ reduce gender inequalities to make it more attractive for women to enter the high-career path. Specifically, they increase the premium to their female high-career employees by s_f financed by a reduction in the salary of their high-career males by t_m so

that their profits remain unchanged.²³ Hence, implementing a WQ in practice translates into imposing firms to reduce wage inequalities.

6.2.1 Social norm is non-binding in the FB

Again, we start with the case represented in Panel (b) of Figure 1 where the FB solution is such that the social norm is not binding but where the “wrong” steady state is achieved in the *LF*. In that case, a transitory policy (implemented during a single period) is sufficient and it must be designed to make q^M the marginal couple. That way the norm disappears in subsequent periods.

Evaluating Expression (24) at $\hat{q}^* = q^M$ shows that this requires a WQ of $r = 1/3 = (1 - 1/2)/(2 - 1/2)$. In words, at least on third of workers in the high-career path must be women. To achieve this, s_f and t_m have to be chosen so that the median couple is indifferent between the high- and low-career path for the female spouse, that is:²⁴

$$\begin{aligned} y + q^M[1 - t_m + \alpha(1 + s_f)] + \beta v(1) - \gamma c_\ell^{M,t-1} = \\ y + q^M[1 - t_m] + v(c_\ell^*) + \beta v(1 - c_\ell^*), \end{aligned} \quad (25)$$

where $c_\ell^{M,t-1} = c_\ell^*$ and c_ℓ^* is implicitly determined by Equation (1). Profit neutrality for the firms requires:

$$t_m E[q] = s_f [1 - G(q^M)] \alpha E_h[q] \quad \Rightarrow \quad s_f \alpha = \frac{E[q]}{[1 - G(q^M)] E_h[q]} t_m \equiv A t_m, \quad (26)$$

where, given Definition (18), $A > 1$. Observe that this policy reduces the wage inequality, and when s_f and t_m are sufficiently large it could even be reversed with $q_m = q(1 - t_m) < q_f = \alpha q(1 + s_f)$. The policy remains effective even in that case as long as it does not reverse the ranking of career choices within couples, that is when the norm cost for fathers is sufficiently large to prevent them from choosing the low-career path.

Welfare within the transition period when WQs are in place is given by:

$$SW^{WQ} = \alpha[1 - G(q^M)] E_h[q] + E[q] + y + G(q^M)[\beta v(1) + v(c_\ell^*) + \beta v(1 - c_\ell^*) - \gamma c_\ell^*]. \quad (27)$$

Comparing (17) and (27) shows that $SW^{WQ} > SW^S$. Unlike the uniform subsidy, which distorts c_ℓ , the WQ policy achieves $\hat{q}^* = q^M$ together with an efficient child-care provision, even in the

²³This modeling strategy of a WQ appears the most coherent in our setting where firms have no active role, and it is also justified by the literature mentioned in Section 2. In particular, Matsa and Miller (2013) and Bertrand *et al.* (2018) show that, after the implementation of a WQ in Norway, the GWG decreased for those women on boards.

²⁴As a referee correctly pointed out, in our setting a WQ is equivalent to a gender-based tax with a progressive income tax for men and a regressive income tax for women. However, the two policies differ in their informational requirements. A gender-based tax would require that workers' future career prospects q are publicly observable, while the implementation of a WQ requires only that q is observed by the employer.

transition period, that is $c_\ell^* = c_\ell^o$. Recall that while c_ℓ^* is inefficient when the norm is binding *in the following period*, it is efficient because the norm will no longer be binding after the transition period; see Footnote 18. In other words, because by definition $c_\ell^* = \arg \max_c [v(c_\ell) + \beta v(1 - c_\ell)]$ the last term in brackets in (27) is larger than its counterpart in Equation (17). Consequently, welfare in the transition generation is larger under a WQ than with a subsidy on market child care. Since both policies yield the same steady state (the FB) this pleads in favor of a WQ.

6.2.2 Social norm is binding in the FB

Once again we show that this instrument is not sufficient to implement a FB optimum with a binding norm so that, like for the subsidy, we are in a second-best setting. It is plain that the WQ won't affect the traditional couples' level of child care so that we have $c_f^M = c_\ell^M = c_\ell^*$, defined by Equation (1) no matter what.

The second-best solution is then obtained by choosing the WQ (or the transfers t_m and s_f necessary to reduce the wage inequality) so as to maximize the following welfare function:

$$\begin{aligned} \max_{t_m, s_f} SW^{WQ} &= \int_0^{\hat{q}^{WQ}} [y + q[1 - t_m] + v(c_\ell^*) + \beta v(1 - c_\ell^*)] g(q) dq \\ &+ \int_{\hat{q}^{WQ}}^{\bar{q}} [y + q[1 - t_m + \alpha(1 + s_f)] + \beta v(1) - \gamma c_f^*] g(q) dq, \end{aligned} \quad (28)$$

subject to the profit neutrality constraint:

$$t_m E[q] = s_f \alpha \int_{\hat{q}^{WQ}}^{\bar{q}} qg(q) dq, \quad (29)$$

where \hat{q}^{WQ} , which specifies the marginal couple under the WQ policy, is given by:

$$\hat{q}^{WQ} = \frac{1}{\alpha(1 + s_f)} [v(c_\ell^*) + \beta [v(1 - c_\ell^*) - v(1)] + \gamma c_f^*]. \quad (30)$$

Equation (30) shows that \hat{q}^{WQ} is a decreasing function of s_f . As s_f increases the effective level of α increases and the participation of women in the high-career path increases.

Differentiating the Lagrangian expression, denoted by \mathcal{L}^{WQ} , with respect to t_m yields:

$$\frac{\partial \mathcal{L}^{WQ}}{\partial t_m} = -E[q] + \mu E[q] = 0. \quad (31)$$

where, μ is the multiplier associated with the profit constraint (29). Equation (31) shows that $\mu = 1$. Because all fathers are in a high-career path, t_m is effectively a lump sum tax (so that the social marginal utility of income is equal to that of the couples, namely 1). Consequently, the marginal social benefit is equal to the marginal social cost; there is no deadweight loss.

The derivative of \mathcal{L}^{WQ} with respect to s_f is given by:

$$\frac{\partial \mathcal{L}^{WQ}}{\partial s_f} = \alpha \int_{\hat{q}^{WQ}}^{\bar{q}} qg(q) dq - \mu \alpha \int_{\hat{q}^{WQ}}^{\bar{q}} qg(q) dq + \mu s_f \alpha \hat{q}^{WQ} g(\hat{q}^{WQ}) \frac{\partial \hat{q}^{WQ}}{\partial s_f} \leq 0. \quad (32)$$

Since $\mu = 1$ and $\partial \hat{q}^{WQ} / \partial s_f < 0$, Equation (32) is equal to zero for $s_f = 0$ and it is negative for $s_f > 0$. This means that the optimal policy implies $s_f = 0$, and with (29), also $t_m = 0$. In other words, no WQ should be imposed. Intuitively, as long as the norm is binding, marginal changes in \hat{q} do not reduce the norm cost and the traditional couples' level of child care remains at c_ℓ^* . Consequently, WQs create no benefits but they involve a cost by distorting the marginal couple. To see this, observe that while \hat{q}^* differs from the FB level, it is efficient given the level of child care c_ℓ^* . Formally, this can be seen by assuming that \hat{q}^{WQ} can be directly controlled, rather than being determined by (30) in which case the solution implies $\hat{q}^{WQ} = \hat{q}^*$. Intuitively, as long as the norm is binding, individual couples' career decisions do not *create* any externality and are therefore efficient.

The main results obtained for the WQ policy are summarized in the following proposition.

Proposition 6 (WQ) *Consider a WQ requiring a minimum share of women in the high-career path and being implemented by a premium s_f to female high-career employees, that is financed by a reduction t_m in the salary of high-career males so that profits are unchanged, then:*

- (i) *when the social norm is non-binding in the efficient steady state, a WQ set to make the marginal couple coincide with the median one implements the efficient solution. A transitional (one period) policy is sufficient, and the efficient steady state is attained after a single transition period. Informal child care by traditional couples is not affected by the policy. Consequently, a WQ policy dominates the uniform subsidy because it yields a larger welfare in the transition period but the same welfare in steady state;*
- (ii) *when the social norm is binding in the efficient steady state, the policy is ineffective and reduces welfare.*

6.3 Parental leave

So far, we have assumed that there is a perfect complementarity between the high-career path and full private child care.²⁵ The two policies we studied before do not affect this complementarity which instead vanishes with parental leave (PL), as we explain below.

Assume that PL entitles a parent (mother or father) to keep the salary y during a given period while taking leave from work to care for the (newborn) child. In the case of low-career workers, PL comes as a “free lunch” because informal care no longer has an opportunity cost in terms of salary. In the case of high-career workers, PL implies that they obtain the same flexibility as low-career workers and are free to decide how much of their time to devote to informal child care. However, taking PL and being absent from work comes at a cost in terms of

²⁵We have even imposed it on the efficient allocations described in Section 5.

future earning opportunities.²⁶ We denote $k \in (0, 1)$ the share of future earning opportunities that are maintained by workers in the high-career path when they request PL. We assume that PL is financed by a lump sum tax T imposed on all couples. The formal analysis of this policy is available in the online Appendix. Here, we restrict ourselves to stating the main results and to presenting the proof in such a way so as to highlight the intuition behind the results.

PL differs from the other policies in that it is an option: couples must decide whether or not to benefit from PL.²⁷ We show that the share of couples where mothers work full-time and provide no informal care *decreases* when PL is introduced. Furthermore, informal child care provided by mothers who work part time *increases* when PL is available.

To understand how these results arise, note first that when PL is available *all* couples will choose the high-career path. This involves no cost but only benefits. Former low-career couples can provide the child care they prefer at a lower cost and enjoy some of the future earnings opportunities. The crucial question for the determination of the norm cost is then to know which of the mothers opt into PL (and work part time) and which ones do not and work full time. Obviously, PL is more attractive the higher is q so that couples below the marginal couple \tilde{q}^{PL} will opt in for PL while those with $q > \tilde{q}^{PL}$ will work full time. So far, this is mainly a definition but the interesting property is that we must have $\hat{q}^* < \tilde{q}^{PL}$. This is because the *LF* marginal couple \hat{q}^* who were indifferent between the low- and high-career paths (part-time or full-time work for the mother) now strictly prefers to work part time (and opt into PL).²⁸ This is because this option has become more attractive (lower cost of child care and preservation of some future earning opportunities) while utility associated with full-time work does not change.

This establishes our first result, namely that the share of couples where mothers work full time and provide no informal care *decreases* when PL is introduced. The second result stating that informal child care provided by mothers who work part time *increases* when PL is available follows directly from the fact that PL decreases the cost of informal care (there is no longer an opportunity cost of lost earnings).

Consequently, PL cannot eradicate the norm because by Assumption 1 we have $q^M < \hat{q}^*$ so that we must have $q^M < \tilde{q}^{PL}$. Furthermore, it cannot even mitigate the norm because the level of child care provided by part-time working mothers *increases* (so that the median level increases).

PL can be welfare improving *only* if the benefit from future earning opportunities accruing to new couples entering the high-career path (those with $q < \hat{q}^*$) more than compensates for the

²⁶This is in line with the findings by Kleven *et al.* (2018) who show that even after 10 or 20 years the “child penalty” persists.

²⁷PL does not affect the ranking of incomes within a couple. Consequently, it will be the mother (if any parent) who will opt in PL.

²⁸Which means that \hat{q}^* must be in the range where couples opt into PL, that is below \tilde{q}^{PL} .

foregone career prospects, $\alpha(1-k)q$, of career mothers opting for PL (those with $\hat{q}^* \leq q < \tilde{q}^{PL}$) and the additional costs of the social norm affecting full-time working mothers (those with $q \geq \tilde{q}^{PL}$). Formally, this is the case if:²⁹

$$\alpha k \int_0^{\hat{q}^*} qg(q) dq - \alpha(1-k) \int_{\hat{q}^*}^{\tilde{q}^{PL}} qg(q) dq - \gamma [1 - G(\tilde{q}^{PL})] (c_h^{PL} - c_\ell^*) > 0. \quad (33)$$

To sum up, PL presents the advantage of breaking the complementarity between the high-career path and market child care but is otherwise costly. Indeed, not only does PL not allow the norm to be eradicated, it even exacerbates it because informal child care provided by couples opting into PL is higher than in *LF*.³⁰

7 Discussion

Complementarity We assume complementarity between private child care and the high-career path (except when PL becomes an option), which jointly determine the norm cost. Either policies making high-careers or private child care more attractive can thus, in principle, eradicate the norm by switching enough couples to the high-career path to make the share of traditional couples drop below the critical level. Child-care subsidies and WQs can achieve this but PL cannot; while it makes the high-career path more attractive, it also fosters part-time work and informal child care which reinforces the norm.

Social norm Let us examine how our results are robust to a general specification of the norm. We abstract from the dynamics to concentrate on second-best policies with binding norms. Assume that the norm cost is given by $\gamma(\max\{0; K[G(\hat{q})]c_\ell - c_\ell\})$. In words, it depends on the *share* of couples who provide informal care and on the *level* of this care. When $K(G) = 0$ for $G \leq 1/2$ and $K(G) = 1$ for $G > 1/2$ we return to the median norm, while $K(G) = G$ yields a norm that depends on the *average* provision of informal care (see, for example, Aronsson and Granlund 2015). More generally, one could think about K as a logistic function translating the idea that when the share of traditional couples is small, the norm becomes insignificant. The counterpart to our binding norm assumption would be that we are in the upper part of the curve. Our median specification is a stylized representation of this case.

When $K'(G) = K(G)/G$, that is when the norm depends on the average level of informal care, the Pigouvian level of the subsidy implements the FB: it yields the appropriate levels of

²⁹In case of unpaid leave the only effect a PL policy has is an increase in flexibility for those who enter the high-career path. We then have $\hat{q}^* = \hat{q}^{PL}$ and $c_\ell^* = c_h^{PL}$ so that the last two terms in Expression (33) drop and a PL policy is always welfare improving.

³⁰This cost could be mitigated if the policy were to be combined, for instance, with a subsidy on formal child care. Thus, from a practical perspective the best way to go may be to implement a mix of different policies.

both c_ℓ and \hat{q} .³¹ However, when $K'(G) < K(G)/G$, as is the case with the median norm (where $K'(G) = 0$), and with the upper part of any logistic curve, the Pigouvian subsidy that restores efficiency of informal child care for a given level of \hat{q} is too large, and yields too many traditional couples. We then return to our result such that the second-best rule requires a level of s set below the Pigouvian rule; see Inequality (23). A WQ, on the other hand, can *never* achieve the (binding norm) FB because it only affects the share of traditional couples but has no leverage on their informal care. While this policy may no longer be *totally* ineffective in the general case, it is less effective than the subsidy. This is most obvious with the average norm where the subsidy implements the FB while the WQ does not. In the general case, one can easily show that any \hat{q} that can be achieved through a WQ can also be obtained with a uniform subsidy but with a smaller level of informal care and thus a lower norm cost. Finally, the results concerning PL would not change under this more general specification of the norm.

To sum up, except for some minor nuances our results are robust with regard to the specification of the norm cost.³²

Welfare function We have considered a *welfarist* social welfare function. Alternatively, a *paternalistic* approach, such that the norm-based part of individuals' preferences is not welfare relevant, could have been used. The government would then try to induce each household to behave as if the norm were absent and only the FB allocation without the norm would be relevant; see Equations (11) and (12). Informal child care under LF would be optimal because the externality does not matter, but the share of career couples would still be too low in LF since career couples are wrongly discouraged by the norm.

Model's predictions and empirical evidence The first part of the paper offers an explanation of the various market outcomes we currently observe in different countries. As an example, consider the Mediterranean (Spain, Italy and Greece) and Nordic countries (Denmark, Sweden and Finland). Compared to Nordic countries, Mediterranean countries are characterized by a more significant gender discrimination (α is small) and by child-care structures offering only partial support to working parents, suggesting a relatively low β .³³ Based on these stylized

³¹Computations are available in the online Appendix.

³²Another alternative is to assume that mothers complying with the norm experience a utility *gain*. In the LF the low-career couples would enjoy a positive externality because they are a majority. The behavioral predictions from the "norm as a benefit" are similar, while the policy conclusions may or may not differ. For instance solutions such as those described in Section 6.1.1 and 6.2.1, where the FB with a majority of high-couples can be achieved, would continue to apply, albeit for different reasons. The high-career couples would then be the conformists enjoying the positive externality. In other cases (such as the counterpart to Section 6.1.2) policy recommendations would be the opposite. The positive externality view would indeed imply that policies should encourage the traditional view of gender roles and would aim at increasing the GWG, which appears to be a rather repugnant conclusion. In addition, this positive externality approach appears empirically less relevant.

³³See Bettio and Plantenga (2004), O'Neill (2003) and Blau and Kahn (2006).

observations the situation observed in these countries is consistent with the predictions of Proposition 1. One indeed observes that these two types of countries currently show a large disparity in the time that mothers and fathers devote to informal child care, with the greatest inequality in child care provisions appearing in Spain and the smallest in Denmark. Furthermore, the share of career mothers is currently already much higher in Nordic countries; see Garcia *et al.* (2011).³⁴

In its normative part, our model shows that a given policy is likely to have a different impact according to the type of FB steady state that is relevant in the considered country. This, again, depends on the country's cultural and historical traditions, and on its economic fundamentals. In particular, it is possible that a social norm is so pervasive and wide-spread that it optimally persists in the efficient steady state. This is more likely to be the case in countries with significant gender discrimination. On the contrary, if a society is relatively closer to gender equality of opportunity then overcoming the norm might be beneficial. From Proposition 4 the observed profile of α 's and β 's makes it more likely in Nordic countries than in Mediterranean ones that the FB solution involves an eradication of the norm. Hence, we expect a benevolent Nordic government to design policies to reach a modern norm-free society, while a Mediterranean government may implement policies aimed at reducing the costs of the existing social norm.

One testable prediction of our model is that WQs might perform well in Nordic countries, but be a less effective policy in Mediterranean ones. In this regard, our findings are in line with the political economy literature showing that, while WQs for electoral candidates have been significant in many European countries, they performed relatively better in Scandinavian countries where they have helped women to achieve the political gains they enjoy today (see, among others, Rosen 2017).

A second clear prediction of our model is that child-care subsidies represent the most effective second-best instrument. This is again in line with the existing empirical evidence that shows that child care availability and, to a lower extent, family allowances have the strongest and least ambiguous effect on female labor supply compared to other policies (see, among others, Del Boca *et al.* 2009 and Olivetti and Petrongolo 2017). In addition, our model suggests that child-care subsidies should be relatively more effective in Mediterranean than in Nordic European countries, because the former are characterized by a larger gender gap. For an empirical test of this result one could compare the effects of child-care subsidies on the participation of women in top positions in Mediterranean and in Nordic countries. However, a main difficulty is that the structure and generosity of child-care subsidies vary substantially, with policies being much

³⁴Our *LF* equilibrium can also explain why inequality in the German-speaking part of the Swiss labor market is higher than in the French-speaking part, even though the institutional environment is similar (see Steinhauer 2013). There is a high degree of tolerance towards working mothers with preschool-aged children in the French-speaking region, but this is much less the case in the German-speaking region. In our model, this is reflected by a costs of norm deviations for mothers, which is larger in the German than in the French-speaking region.

less generous in Mediterranean than in Nordic countries.

8 Conclusion

This paper presents a simple model to explain observed gender patterns of labor market and child care decisions through (endogenously determined) social norms. We concentrate on the steady state and characterize the LF solution when the norm is binding, that is when the median couple is traditional. This solution implies that informal child care and the share of traditional couples are larger than efficient.

We show that a linear subsidy on market child care is always welfare-improving, albeit to a differing degree depending on the efficient steady state. When this is of the no-norm type, a transitional (one period) policy implements the efficient steady state, otherwise we have a second-best solution since the externality will only partly be mitigated. WQs imposed during a single transition period are effective in achieving a FB no-norm steady state. In this situation, a WQ yields a higher transition welfare than child-care subsidies. When, however, the efficient steady state implies a binding norm, then the policy is ineffective; it has no impact on the traditional couples' child care and thus cannot mitigate the externality. A PL policy tends to foster informal care and exacerbates norm costs. However, it can break the complementarity between high-career and full market care which otherwise characterizes our model. It can be second-best efficient when the benefit from future earning opportunities accruing to new couples entering the high-career path outweigh the additional costs of the social norm affecting the couples who do not take up PL.

Throughout the paper we concentrate on efficiency issues. With quasi-linear preferences and a utilitarian welfare function, redistribution and equity considerations are of no relevance.³⁵ This is important to keep in mind, because the uniform subsidy we consider is clearly regressive; it provides larger benefits to high-career couples who also have a higher income than traditional couples. When the policy is transitory, the regressive effect will be relevant but only within a single period. However, it will continue when the policy is permanent. Consequently, one can expect the optimal second-best subsidy to be lower when redistribution is accounted for, either because individual preferences are concave or because social welfare applies a concave transformation to individual utilities. The redistributive impact of WQ, on the other hand, is more complicated to assess. Because WQs imply only transfers across high-career couples, they are obviously less regressive than the subsidy. Of the three policies considered in the paper, PL is the only one with clear progressive effects, given that all mothers are able to enjoy future earning prospects from the high-career path.

³⁵For a model in which redistributive aspects are considered when social norms are in place see Aronsson and Granlund (2015).

References

- [1] **Akerlof, G.A., and R.E. Kranton**, “Economics and identity,” *The Quarterly Journal of Economics*, 2000, 715–753.
- [2] **Akerlof, G.A., and R.E. Kranton**, “Identity economics,” *The Economists’ Voice*, 2010, 7 (2).
- [3] **Aronsson, T., and D. Granlund**, “Gender norms, work hours, and corrective taxation,” *Journal of Behavioral and Experimental Economics*, 2015, 56, 33–39.
- [4] **Averett, S.L., Peters, E.H., and D.M. Waldman**, “Tax credits, labor supply, and child care,” *Review of Economics and Statistics*, 1997, 79 (1), 125–135.
- [5] **Bastani, S., T. Blumkin and L. Micheletto**, “Anti-discrimination legislation and the efficiency-enhancing role of mandatory parental leave”, 2016, *CESifo WP 5917*.
- [6] **Bénabou, R., and J. Tirole**, “Identity, morals and taboos: beliefs as assets,” *Quarterly Journal of Economics*, 2011, 126, 805–855.
- [7] **Bernal, R.**, “The effect of maternal employment and child care on children’s cognitive development,” *International Economic Review*, 2008, 49 (4).
- [8] **Bertrand, M., Goldin C., and L.F. Katz**, “Dynamics of the gender gap for young professionals in the financial and corporate sectors.” *American Economic Journal: Applied Economics*, 2010, 2 (3), 228–55.
- [9] **Bertrand, M.**, “New perspectives on gender,” *Handbook of Labor Economics*, 2011, 4, 1543–1590.
- [10] **Bertrand, M., Black, S.E., and A. Lleras-Muney**, “Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway,” *forthcoming, Review of Economic Studies*, 2018.
- [11] **Bertrand, M., P. Cortes, C. Olivetti and J. Pan**, “Social norms, labor market opportunities, and the marriage gap for skilled women”, 2016, *NBER WP 22015*.
- [12] **Bettio, F., and J. Plantenga**, “Comparing care regimes in Europe.” *Feminist Economics*, 2004, 10 (1), 85–113.
- [13] **Bisin, A., and T. Verdier**, “Cultural transmission, marriage, and the evolution of ethnic and religious traits,” *The Quarterly Journal of Economics*, 2000, 115 (3), 955–988.
- [14] **Bisin, A., and T. Verdier**, “The economics of cultural transmission and the dynamics of preferences,” *Journal of Economic Theory*, 2001, 97 (2), 298–319.
- [15] **Blau, F.D., and L.M. Kahn**, “The US gender pay gap in the 1990s: slowing convergence,” *Industrial and Labor Relations Review*, 2006, 60 (1), 45–66.
- [16] **Blau, F.D., and L.M. Kahn**, “The gender wage gap: extent, trends, and explanations,” *Journal of Economic Literature*, 2017, 55 (3), 789–865.

- [17] **Blinder, A.**, “Wage discrimination: reduced form and structural estimates,” *Journal of Human Resources*, 1973, 8, 436–55.
- [18] **Buechel B., T. Hellmann, M.M. Pichler**, “The dynamics of continuous cultural traits in social networks”, *Journal of Economic Theory*, 2014, 154, 274–309.
- [19] **Bursztyń L., Fujiwara T. and Pallais A.**, “‘Acting wife’: marriage market incentives and labor market investments”, *American Economic Review*, 2017, 107 (11), 3288–3319.
- [20] **Card, D., A.R. Cardoso, and P. Kline**, “Bargaining, sorting, and the gender wage gap: quantifying the impact of firms on the relative pay of women.” *Quarterly Journal of Economics*, 2016, 131 (2), 633–86.
- [21] **Ciccia, R., and M. Verloo**, “Parental leave regulations and the persistence of the male breadwinner model: using fuzzy-set ideal type analysis to assess gender equality in an enlarged Europe,” *Journal of European Social Policy*, 2012, 22 (5), 507–528.
- [22] **Del Boca, D., S. Pasqua and C. Pronzato**, “Motherhood and market work decisions in institutional context: a European perspective,” *Oxford Economic Papers*, 2009, 61(1), 147–171.
- [23] **Del Rey, E., M. Racionero and J.I. Silva**, “On the effect of parental leave duration on unemployment and wages”, *Economics Letters*, 2017, 158 14–17.
- [24] **European Commission**, “More women in senior positions”, Directorate-General for Employment, Social Affairs and Equal Opportunities, January 2010, available at http://ec.europa.eu/danmark/documents/alle_emner/beskaeftigelse/more_women_in_senior_positions.pdf
- [25] **Felfe, C., and R. Lalive**, “Does early child care help or hurt children’s development,” *Journal of Public Economics*, Forthcoming.
- [26] **Fernandez, R.**, “Alfred Marshall lecture: women, work, and culture,” *Journal of the European Economic Association*, 2007, 5 (2–3), 305–332.
- [27] **Fernandez, R., and A. Fogli**, “Culture: an empirical investigation of beliefs, work, and fertility,” *American Economic Journal: Macroeconomics*, 2009, 1 (1), 146–177.
- [28] **Fernandez, R., Fogli, A., and C. Olivetti**, “Mothers and sons: preference formation and female labor force dynamics,” *The Quarterly Journal of Economics*, 2004, 1249–1299.
- [29] **Fortin, N.M.**, “Gender role attitudes and the labour-market outcomes of women across OECD countries,” *Oxford Review of Economic Policy*, 2005, 21 (3), 416–438.
- [30] **Gallen, Y.**, “Motherhood and the gender productivity gap,” The University of Chicago Harris School of Public Policy, Working Paper, January 26, 2018.
- [31] **Garcia, I., J.A. Molina, and V.M. Montuenga**, “Gender differences in childcare: time allocation in five European countries”, *Feminist Economics*, 17 (1), January 2011, 119–150.
- [32] **Gelbach, J.B.**, “Public schooling for young children and maternal labor supply,” *The American Economic Review*, 2002, 92 (1), 307–322.

- [33] **Gregg, P., E. Washbrook, C. Propper, and S. Burgess**, “The effects of a mother’s return to work decision on child development in the UK,” *The Economic Journal*, 2005, 115, 48–80.
- [34] **Goldin, C.**, “The quiet revolution that transformed women’s employment, education, and family,” *The American Economic Review*, 2006, 96 (2), 1–21.
- [35] **Goldin, C.**, “A pollution theory of discrimination: male and female differences in occupations and earnings human capital in history”, in *The American Record*, Leah Platt Boustan, Carola Frydman, and Robert A. Margo, Editors, University of Chicago Press, 2014a, 313–348.
- [36] **Goldin, C.** “A grand gender convergence: its last chapter.” *American Economic Review*, 2014b, 104 (4), 1091–1119.
- [37] **Guendouzi, J.**, “‘The guilt thing’: balancing domestic and professional roles,” *Journal of Marriage and Family*, 2006, 68 (4), 901–909.
- [38] **Huerta, M.d.C., W. Adema, J. Baxter, M. Corak, M. Deding, M.C. Gray, W.-J. Han, and J. Waldfogel**, “Early maternal employment and child development in five OECD countries,” *OECD Social, Employment and Migration Working Papers*, 2011, 118.
- [39] **Kimmel, J.**, “Child care costs as a barrier to employment for single and married mothers,” *Review of Economics and Statistics*, 1998, 80 (2), 287–299.
- [40] **Kleven, H., and C. Landais**, “Gender inequality and economic development: fertility, education and norms”, *Economica*, 2017, 84, 180–209
- [41] **Kleven, H., C. Landais, and J.E. Søgaard**, “Children and gender inequality: evidence from Denmark”, *NBER WP 24219*, 2018.
- [42] **Lalive, R., and J. Zweimüller**, “How does parental leave affect fertility and return to work? Evidence from two natural experiments” *The Quarterly Journal of Economics*, 2009, 1363–1402.
- [43] **Lindbeck, A., S. Nyberg and J. Weibull**, “Social norms and economic incentives in the welfare state,” *The Quarterly Journal of Economics*, 1999, 114 (1), 1–35.
- [44] **Lindbeck, A., S. Nyberg and J. Weibull**, “Social norms and welfare state dynamics”, *Journal of the European Economic Association*, 2003, 1 (2–3), 533–542.
- [45] **Matsa, D.A., and A.R. Miller**, “A female style in corporate leadership? Evidence from quotas,” *American Economic Journal: Applied Economics*, 2013, 3 (3), 136–69.
- [46] **OECD**, “Doing better for families,” *OECD Employment Database*, 2010.
- [47] **Olivetti, C., E. Patacchini, and I. Zenou**, “Mothers, peers and gender identity,” *Mimeo*, 2016.
- [48] **Olivetti C. and B. Petrongolo**, “The economic consequences of family policies: lessons from a century of legislation in high-income countries” *Journal of Economic Perspective*, 2017, 31 (1), 205–230.

- [49] **O’Neill, J.**, “The gender gap in wages, circa 2000,” *The American Economic Review*, 2003, 93 (2), 309–314.
- [50] **Oaxaca, R.**, “Male–female wage differences in urban labor markets”, *International Economic Review*, 1973, 14, 693–709.
- [51] **Panebianco, F.**, “Socialization networks and the transmission of interethnic attitudes”, *Journal of Economic Theory* 2014, 150, 583–610
- [52] **Paull, G.**, “Children and women’s hours of work,” *The Economic Journal*, 2008, 118, F8–F27.
- [53] **Rose, J.**, “Never enough hours in the day: employed mothers’ perceptions of time pressure,” *Australian Journal of Social Issues*, 2017, 52, 116–130.
- [54] **Rosen, J.**, “Gender quotas for women in national politics: A comparative analysis across development thresholds,” *Social Science Research*, 2017, 66, 82–101
- [55] **Rotkirch, A., and K. Janhunen**, “Maternal guilt,” *Evolutionary Psychology*, 2010, 8 (1), 90–106.
- [56] **Ruhm, C.J.**, “The economic consequences of parental leave mandates: lessons from Europe,” *The Economic Journal*, 1998, 113 (1), 285–317.
- [57] **Schwartz, C.R., and R.D. Mare**, “Trends in educational assortative marriage from 1940 to 2003,” *Demography*, 2005, 42 (4), 621–646.
- [58] **Steinhauer A.**, “Identity, working moms, and childlessness: evidence from Switzerland”, mimeo 2013.
- [59] **Young, H.P.**, “The evolution of conventions”, *Econometrica*, 1993, 61, 57–84.
- [60] **Young, H.P.**, *Individual Strategy and Social Structure*, Princeton NJ: Princeton University Press, 1998.
- [61] **Young, H.P.**, “Self-knowledge and self-deception,” JohnHopkins University mimeo, 2008.