The onset of female labor market participation and the role of the mothers[1]

Jesús Carro, Matilde P. Machado[2], Ricardo Mora[3]

Universidad Carlos III de Madrid

PRELIMINARY AND INCOMPLETE

February 15, 2013

[1]We acknowledge the support of the Ramón Areces Foundation (IV Concurso Nacional). We thank M. Norberta Amorim and staff at NEPS for assistance with the original data and important historical information. We are also thankful to Maristella Boticcini, Jaime Reis, Stefan Houpt and the audience of the CEPR conference on the Economics of Interactions and Culture for helpful comments.

[2]Universidad Carlos III de Madrid and CEPR

[3]Corresponding author: Ricardo Mora, Universidad Carlos III de Madrid, Department of Economics, calle Madrid 126, 28903 Getafe (Madrid), Spain. E-mail: ricmora@eco.uc3m.es

Abstract

Female labor participation has increased in most countries. Although the literature has advanced several explanations for this fact, the reasons for the onset of female labor market participation are not well understood. Among other factors, productivity differentials, and human capital accumulation have been proposed as fundamentals to understand the trends in female work status. Recently, economists have established that culture differences affect economic outcomes and in particular the levels of female labor force participation. Previous studies also showed that culture is transmitted from parents to children and, therefore, transitory shocks to preferences or beliefs may have long-run impacts on economic outcomes. To our knowledge, there are no studies which try to directly evaluate the role of intergenerational transmission of preferences and beliefs within the family in the onset of female labor market participation. We argue that intergenerational family data from parish records from the 18th and 19th century provide us with a nice opportunity to evaluate the importance of this mechanism for persistence in the work status of female workers. Using church registry data from several Portuguese settlements, we propose to estimate a reduced-form dynamic discrete choice model. Importantly, we address the problem of missing values, which so often affects historical records using a methodology recently proposed by Ramalho and Smith (2012). Preliminary results show a positive and statistically significant effect of the mother working status on the daughter's probability of working across different model specifications. We interpret such an effect as evidence of intergenerational family transmission of preferences and/or beliefs.

Keywords: Female Labor participation, dynamic discrete choice models, historical family data, non-ignorable missingness

JEL Classification: J22, J24, J16, J12

1 Introduction

The change in female work status is arguably one of the most important transformations in the labor market in the last century. As Goldin (1994) reports, in the United States, married female labor force participation is around 2% in 1880, 35% after World War II, and around 74% in 1999. The male-female participation gap by the end of the 20th century is only 12.9% with most women combining family and work. A rare event at the beginning of the 20th century becomes the mode by its end. Similar patterns are observed in other OECD countries. Portugal, the country on which this paper focuses, experiences a similar trend. According to the Portuguese Census, in 1911 around 37% of women between $15-65$ years old has a profession. By 1975, the figure increases to around 50% and by the end of the century Portugal's female participation rate reaches 67% and the male-female gap shrinks to 13.1% (INE, 1977, and Eurostat).

Several hypothesis have been considered to explain this increase in female labor force participation (Costa, 2000, Del Boca and Locatelli, 2006). These theories fall mainly into two broad categories: those which stress the role of technological changes (Goldin, 1994, Galor and Weil, 1996, Goldin and Katz, 2002, and Greenwood, Seshadri, and Yorukoglu, 2005) and those which emphasize the importance of changes in beliefs and/or preferences. The latter can be further split into studies that establish the importance of preferences and beliefs to economic outcomes (e.g. Fernández, Fogli, and Olivetti, 2004, Fernández, 2007, 2011b, Fernández and Fogli, 2009, Fernández, 2011a) and studies of intergenerational transmission of preferences and beliefs (e.g. Bütikofer, 2008, Farré and Vella, 2007, and Tsukahara, 2007). The latter studies find a relation between parents' labor market behavior and attitudes and those of their sons and daughters. Although these effects are likely to compound both changes in preferences/beliefs and the transmission of specific human capital, they are persuasive evidence of the importance of family effects. Our paper fits into this group but we use a novel approach to provide more direct evidence on the intergenerational transmission of preferences and beliefs.

We use historical parish registry data to identify and measure, in a reduced-form choice model, the effect of a mother's work-for-pay status on that of her daughter's. Our data comes from four Portuguese settlements from the 18th and 19th centuries gathered by the research institute NEPS (Núcleo de Estudos de População e Sociedade, Portugal). We argue that the historical and geographical settings—the onset of female labor market participation in Portugal—enable us to interpret this effect as evidence of mother-to-daughter transmission of preferences and/or beliefs. In our preferred specification, the probability that a woman works for pay increases by 8.6 percentage points if her mother also did. This effect is very large given that the estimated unconditional probability of working-for-pay in the sample is 11%.

The legal and social background of the Portuguese society during the sample period is one of hostility towards the economic independence of women. They were not only legally discriminated, but also excluded from the main educational system.[1] The real first changes in the Portuguese legal system regarding women's rights take place only after the proclamation of the First Republic in 1910. For example, the 3rd of November of 1910 Act legalized divorce and provided for the first time in Portugal some degree of gender equality in relation to marriage rights and the legal treatment of adultery.[2]

The period and country ensures that the cultural transmission effect is identified separately from the later technological and cultural changes that—in Portugal—only took place in the 20th century.

In sum, by looking for family effects in matrilineal transmission for a taste in female participation in Portugal in the 18th and 19th centuries.

, we effectively attempt to find socially unconventional behavior stemming from the family experience and where specific human capital accumulation cannot play a relevant role.

To assure our model of the transmission of labor market participation behavior from

[1]According to the population census of 1864, the share of boys between 6 and 15 years old attending primary education in Horta and Braga—the regions where the settlements are located—is 13.4% and 18.0%, respectively. In contrast, the same shares for girls are only 5.0% and 1.3%. Although some young women began to receive higher education late in the 19th century, only in 1888 a law was passed to allow for the creation of all-girls public schools for secondary education.

[2]The slow changes in Portuguese society translated into a gradual incorporation of women economic rights into Portuguese law. For example, women were given the right to became civil servants in 1911 and lawyers in 1918 but had to wait until 1966 for the right to have, for the same job, the same wage as a man, and until 1978 for the right to be a trader without the husband's consent. Full legal gender economic equality was recognized with the 1976 democratic constitution.

mothers to daughthers is only capturing the transmission of preferences or beliefs, we must discard the possibility that it is also capturing effects related to the transmission of wealth in the form of land or capital or social status. For example, if property was transmited mainly to daughthers, then the economic activity of females is bound to be similar to their mothers' economic activity. With respect to the transmission of property, in the north of Portugal and in particular in the Minho region where two of our villages are located, the norm was to divide two thirds of the property (the *legitima*) equally amongst the legitimate heirs and to dispose of one third (the *terço*) to benefit one of the children or the surviving spouse. Typically, the terço included the main house (or part of it) with adjacent land and was either given during life to the first child who got married as marriage bequests (daughters tended to be favored[3]) or left upon death to a spouse or unmarried children typically unmarried daughters (Brettell, 1991, Durães, 2009, Pina Cabral, 2986, Matos, 2009). Importantly, parents used bequests, such as the terço, and donations as an insurance contract to guarantee their maintenance and often also their power within the family at older ages (Brettell, 1991 and Durães, 2009, Matos, 2008). Moreover, in the Minho countryside, the rights within the household seemed to be "balanced between husband and wife" (Brettell, 1991) and the 'household couple' instead of the 'head of the household' was the reference regarding authority (Pina-Cabral, 1986). However, opposite to this culture of relative equality among heirs, at least regarding the distribution of the legitima, there existed two contemporanous property transmission norms which imposed great inequality amongst heirs and typically favored males. The first one, was a male Primogeniture or Majorat system (called *Morgadio* or *Morgado*) by which the oldest son inherited the land as well as the name (and title) of the property owner. The Morgadio was a common practice amongst the wealthiest families of landlords and aristocrats since the 13th century until it was abolished in 1863 (Moreira da Silva,

[3] Daugthers were often favored with the terço for several reasons. First, because having land or the promisse to it increased a woman's changes in the marriage market since landless women seldom married. Second, since married daugthers tended to live with their parents for a period of time (at least until the couple had their own house and land and/or until the next daughter married), they were more welcome in the house than daughthers-in-law. Third, it was also common for single daughthers to inherit the terço, which, on the one hand, would guarantee them the means of survival and, on the other hand, would guarantee the parents would be taken care in their old age.

1983). The Morgadio system avoided the fragmentation of the land from one generation to the other and kept the economic and social status of the family name unchanged. Under this system, the sons who did not inherited land would emigrate (mostly to Brazil), become members of the clergy, or pursue a military career (Moreira da Silva, 1983 and Scott, 1999). The Morgadio affected a very small percentage of families and it unlikely present in the four settlements. The second one, was applicable only to life-long rentals of aristocratic or ecclesiastic land during the 18th and part of the 19th century. The law made a distinction between own property which was to be transmitted according to the legitima and terço dispositions and the life-long rentals which had to be transmitted to a single heir. The rights to life-long rentals favored spouses over children, male over female children, and older over young children, in this order establishing an inequality amongst the heirs regarding distribution of land (Durães, 2009). Since rentals were a common way to obtain land, these inequalities may have been somewhat widespread until these rights were abolished.[4] Our results are robust to controls that reflect the mother and father social status as well as to re-estimation using subsamples where the mothers' property and social status are unlikely to bias our results.

The data was extracted from parishes in four settlements: São Tiago de Ronfe (hereafter Ronfe), Ruivães, Horta and São Mateus (S. Mateus). The villages of Ronfe and Ruivães are located in the Minho region in the northeast of Portugal, while Horta, a urban coastal settlement at the island of Faial, and S. Mateus, a rural village at the island of Pico, are located in the Azores archipelago. Ronfe and Ruivães are only 9 km away from each other and strategically located between two historical centers of activity and administrative power, Guimarães and Braga.

Although all settlements are part of Portugal and share the same Portuguese culture, they have distinctive economic characteristics in relation to women economic roles and are thus potentially interesting to study separately. Due to the predominantly male emigration to Brazil since the 16th century, the Minho region—to which Ronfe and Ruivães

[4] In a sample of roughly 1300 wills between 1720 and 1820 in the Braga region, Durães (2009) observes that a single heir was often privileged relative to the others also in the distribution of private land. When this was the case, the will would stipulate how the other heirs should be compensated with goods or money (Durães, 2009) to proxy for the equal share imposed in the legitima.

belong—was atypical in terms of the gender composition of the population with women outnumbering men since the beginning of the 18th century until the middle of the 19th century.[5] Scott (1999) argues that this unbalance between the number of men and women led to unorthodox situations which did not agree with the social and religious conventions. In particular, a relatively large number of women were head of households because their children were not conceived within a religious marriage. In Ronfe, for example, 20.7% of the head of households were single females in 1750 and 18% of the children baptized in 1700 were illegitimate (rate which declined to 3.9% by 1900, Scott, 1999).[6] The absence of males to perform many of the typically male jobs opened the way for women to intervene in the society. Moreover, given their status as head of households, they were more likely to take jobs to feed their children. In Scott's description of the period, one is led to think that there was transmission of preferences and attitudes from fathers to their unmarried daughters. In the second half of the 19th century, this region recovered and was one of the first in the country to develop due to the cotton manufacturing industry. In fact, while emigration to Brazil was a phenomenon affecting most of the country, this region saw the return of many emigrants. In contrast to Ronfe and Ruivães, Horta is a city as well as a stopping port in the journey to Brazil with an economy traditionally based on trade and human migrations between the two countries. Consequently, the professional choices and social structure of its inhabitants differ markedly from those of the dwellers of Ronfe and Ruivães. Finally, the rural village of S. Mateus was a traditional settlement and potentially provides information of a population which was not affected at the time by factors which could lead to higher female labor force participation although its male-female ratio was also low due to male emigration (Amorim and Santos, 2009).

[5] In Ronfe, for example, the males/females ratio was only 0.67 in 1740 and by $1801 - 1802$ although the ratio improved there were still only 89 men for 100 women in the Minho region, the lower ratio in the country at the time (Scott, 1999).

[6] Due to the continuous flow of male emigration, single women had difficulty to find a husband, which explains the high definitive celibacy rates amongst women (between 10.3% and 35.5% in Ronfe during 1700-1900, Scott, 1999). Moreover, since having land (or the promisse to it) was a requisitive for marriage and this was often only attainable at the death of one or both parents, the average age at first marriage was extremely high for both women and men: around 27/28 years for women and men in Ronfe during 1700-1900 and 25 years for the Pico Island in the 19th century (Amorim and Santos, 2009). The tradition and laws applied to the transmission of land are also said to cause high rates of definite celibacy among men until 1850 (between 11.8% and 15.2% in Ronfe during 1700-1900 and around 15% in the 19th century, Scott, 1999). Note that celibacy rates are calculated for the 50-54 age group.

NEPS collected individual-specific information contained in baptism, marriage and death certificates from local parishes and used these records to reconstruct whole family linkages starting in the 16th century. Crucially for our objective, NEPS matched this family dataset with other church records called *rol de confessados* (literally "the list of the confessed") —which were used by the local vicar during Lent to monitor the administration of the sacrament of penance to the parishioners. The *rol de confessados* contains, among other, information recorded by the local vical on individual specific occupations or social status.

Like most historical data, ours suffer from a serious problem of missing values affecting primarily the information on occupation from the *rol de confessados* which is our main source to construct the labor force participation status of females (the dependent variable) and their mothers' (the independent variable of interest).[7] We propose two ways to deal with this problem. First, motivated by the historical records and narratives, we conservatively attribute predominantly labor market attitudes by gender to the missing values, i.e. agricultural activity for males and non-participation for females.[8] Second, we explore an estimation approach to dealing with the missing observations in a similar vein to Ramalho and Smith (2012). More specifically, we conduct Maximum Likelihood (ML) estimation of the labor force participation model accounting for nonignorability in the selection process triggering occupational nonresponse. The identification strategy for the parameters in the model is additionally based on the use in our ML estimation of predictions of aggregate female labor force participation obtained from external data sources, in particular the population census from 1864 to 1991.

Preliminary results based on simple probit estimation using conservative imputation methods for the missing occupational choices suggest a positive and statistically significant effect of the mother working status on the daughter's probability of working. These naive estimations, however, produce a wide range of values for the parameter of interest. In

[7] We also used information on titles and social status in general to inform us about female labor participation.

[8] This admittedly crude approach to imputation could be complemented by more sophisticated imputation methods. For example, we can randomly assign participation status in those observations with missing codes using as the random distribution for the probability of participation in those observations estimates using a small set of covariates such as time, age, and number of siblings.

contrast, results using the Ramalho and Smith methodology show a marginal effect for the working experience of the mother of 8.6 percentage points which represents an increase of 60 to 78 percent in the probability of working. The Ramalho Smith ML method also produced an estimate for the female labor force participation at the time which is consistent with our aggregate predictions using the census data.

The remainder of this paper proceeds as follows. Section 2 describes the data set. Section 3 provides the econometric model while Section 4 presents our estimation results. Section 5 concludes. Apendix A describes the estimation of female participation rates from census data used in the ML estimation. Appendix B contains Tables and Appendix C contains the derivatives of the ML function used in the estimation code to speed convergence.

2 Data

2.1 Parish data

The main source of the data used in this paper is parish-data information that goes back to the end of the 16th century extracted from parishes in the villages of Ronfe, Ruivães, Horta, and S. Mateus. The data has been gathered at the Núcleo de Estudos de População e Sociedade (NEPS), a research institute associated to the Universidade do Minho, by a research team led by M. Norberta Amorim. NEPS staff collected the main data sets using all baptism, marriage, and death certificates found in the local churches. These original church records have allowed the reconstruction of family linkages starting in the 1550s through the 20th century (Amorim, 1991). Altogether, the data set has entries for $92,474$ individuals.

Individual records prepared by NEPS include information on birth, marriage, and death dates, gender, parents identification code, identification code of the spouse, children identification code, and parish code. Gender was originally inferred from the individual's first name in the Parish registry. Therefore, most observations have this variable correctly recorded, the exception being babies who die before they are baptized. Overall, $46,094$

records are females (49.84%).

We aim to study the relation between mothers and daughters labor force participation decisions. It is thus essential to identify female lineages among free women.[9] The mother (grandmother) is identified in the sample for 28,360 (17,162) women. The figure is 28,355 for free women. We will refer to this sample as the *women sample* henceforth. Some observations do not have any female ancestors because they are orphans or observations at the beginning of the data period. For some observations female ancestors are missing in the sample because it was not possible to match the parish records. As it is reasonable to assume that the problems encountered by NEPS in matching parish records are not systematically related to the individuals' labor market behaviors, the sample for which we have female lineages is still a representative sample of the population in the four settlements.

Observations include individuals who do not survive childhood and also individuals who migrate to other locations and for whom no further information is available. In theory, the former group of individuals could be identified with the death date while the latter group could be indirectly inferred by the absence of information on their deaths. However, death dates are missing for the majority of the records, 66.78%, which hinders precise identification of both early death and migrant status. In contrast, we have the year of birth for 60.89% of all records. Consequently, we focus on exploiting birth-date information and attempt to complete records for which the date of birth is missing. To do so, we group all individuals for which the information is available into cohorts spanning 25 years. Observations for which the year of birth is missing can be completed by sequentially looking at the 25-year period of first the siblings, then the spouse, and finally the children. When the cohort of the siblings or the spouse is identified, the record is completed with the 25-year period of the spouse or sibling. In case only cohorts of the children are identified, the previous cohort of the oldest child is assigned to the missing record. This procedure is repeated until no changes are produced. As a result, 85.16% of the original

[9]The data set originally includes 887 slaves of which 557 are women, all but one located in the Azores. As we aim to model the effect of the mother's economic activity on individual choices, we do not include these observations in our study.

data can be associated to a given 25-year period. For the subsample of free women with identified mother, only 460 (1.62%) cannot be assigned to any particular 25-year period. It is important to note, however, that while the edited data is a reasonably good representation of the family links of the born free female population in each of the settlements, it should not be considered a sample of the living population who make professional choices in the four settlements because, as already mentioned, some women die before reaching the adult age and others migrate.

We restrict our analysis to the18th and 19th centuries, our final sample has 25,116 females. In Table 2.1 we show the number of observations by settlement and century of the resulting sample.

TABLE 2.1
Number of Observations by Settlement and Century

	Horta	Ronfe	Ruivães	S. Mateus	Total
18th century	5,882	1,450	820	3,268	11,420
19th century	5,667	1,911	1,330	4,788	13,696
Total	11,549	3,361	2,150	8,056	25,116

*Note: identified slaves, their ancestors and descendants were dropped from the sample.

2.2 Information on economic activity in the Lent Census

In a remarkable exercise of involved historical research, parish data containing birth, marriage, and death certificates were exhaustively matched by NEPS with data from a church Lent census, the so-called "*rol de confessados*" (literally, the list of the confessed). The latter information is a parochial census organized by households of all residents older than 7 years of age and produced by the parochial vicar during Lent to administer the sacrament of penance to the parishioners. Crucially to our objective, the Lent census includes information on the economic activity or social status of adult individuals in the household.

The economic information collected in the Lent census is, however, not as reliable as the baptism, marriage, and death registries as it is based on self-reported information and depends on the parochial vicar's recording practices on professional categories. In particular, there are at least two problems related with these data. First, professions were not systematically classified across parochial vicars and centuries. As a result, the original data includes more than 500 categories, many of them being close substitutes. To make this information tractable, we construct a variable on economic activity using four major categories: *employee/farmers*, *professional/capital owner*, *domestic production*, and *unproductive*.[10] *Employee/farmers* includes all paid and unqualified jobs while *professional /capital owner* includes landlords, liberal professions, traders, businessmen, and qualified and managerial jobs. *Domestic production* refers to a category, *domestica*, originally recorded in 310 observations while all women who are recorded by the priests as *Dona*, a term originally used to signal high social status, are classified as *unproductive* (805 observations).

The second problem is more challenging. Panel A in Table 2.2 shows the proportion of observations in the women sample containing economic activity information. Even in the settlement with the highest coverage— Horta —the proportion of nonmissing observations is very small: 12.46%. Accounting for migrations and early deaths would increase the coverage for the living population in Horta, but it is unlikely that they would explain the observed gap. The coverage is particularly poor in S. Mateus, for which only 0.24% of women have their occupations recorded. Occupational coverage also differs by century across settlements. For example, while the coverage of occupations is already relatively higher in Horta in the18th century, it increases to a surprisingly high level—18.44%—in the 19th century.

Given the limits of women economic rights, it would be tempting to postulate that the low coverage of women's economic activity was due to gender bias in the recording practices of the vicars. However, for males the reporting is only marginally higher (for example 17.15% for Horta and 0.70% for Ruivães). Thus, it seems that vicars have not been less inclined to report women economic activities. Differences in coverage are likely

[10] Usually, most individuals only have one professional category recorded in the registry. For Horta, however, 122 and 5 individuals (males and females) have two and three occupations recorded, respectively. For these individuals, we pick up either the profession for which we have the largest time span, or else the last observation.

more related to the parochial vicar's data collection behavior regarding social status of the professions reported and his own social network.

TABLE 2.2
INFORMATION ON FEMALE ECONOMIC ACTIVITY

	Horta	Ronfe	Ruivães	S. Mateus	Total
PANEL A: Proportion (%) of Sample with Recorded Economic Activity					
18th century	6.70	5.86	0.00	0.12	4.23
19th century	18.44	9.42	10.3	0.31	10.05
Total	**12.46**	**7.88**	**6.37**	**0.24**	**7.41**
PANEL B: Distribution (%) of Women Across Occupational Categories					
	Horta	Ronfe	Ruivães	S. Mateus	Total
Employee/Farmer	6.67	30.57	37.96	5.26	12.37
Professional/Capital owner	9.94	44.91	23.36	31.58	16.13
Domestic Production	83.18	1.13	38.69	0	67.37
Unproductive	0.21	23.40	0	63.16	4.14
Total	100	100	100	100	100

*Labor force participation is defined as either *employee/farmer* or *professional/capital owner*.

In Panel B of Table 2.2, we report the distribution of women across economic activity categories by settlement. Distributions vary significantly across settlements and may confound location-specific economic factors with the result of differentials in the incidence of missing information. Some occupational categories are surprisingly underrepresented in the four settlements. For example, one would expect that domestic production would be more demographically important everywhere which leads us to conjecture that most vicars leave the labor force status for housewives unreported. The share of *employees/farmers* is also unrealistically low for rural communities. It is plausible that some parochial vicars do not collect professional information on isolated farmers as the Lent census was organized by households and gathered by the priest door to door. Some priests may also be more likely to underreport the incidence of professions and social status which carry a stigma, such as beggars and prostitutes. The most likely explanation, though, is that priests tended to record only the professions of those whose labor or social status was uncommon in the region and period such as civil servants or the miller in a village of farmers. Table 7 in Appendix B reports the most commom disaggregated female professions/social status reported by the vicars.

2.3 Dealing with Missing Information on Economic Activity

The seemingly underrepresentation of housewives and farmers implies that missingness is nonignorable. Disregarding missing observations about economic activity would likely bias upwards estimates of the effect of mother's labor force participation on the daughter's probability of participation. A higher proportion of working daughters among those whose mother worked would wrongly be interpreted as intergenerational transmission when it could simply arise because those working whose mothers are housewives would be underrepresented in the sample.

The traditional answer to nonignorable missingness is to carry out an imputation inference procedure in which the missing values are filled in (see, among others, Little and Rubin 2002). We start by implementing this procedure by assuming somewhat naively that women with missing occupations engage in domestic production (either housewives or non-paid farmers) and are, therefore, not participating in the labor force. Similarly, for fathers, with missing occupation we assume that they are *employees/farmers*. In Section 3, we show how we avoid the imputation procedure by implementing the method proposed by Ramalho and Smith (2012).

We define active labor force participation as being either *employee/farmer* or *professional/capital owner*. In Table 2.3, we report unconditional transition rates in labor force participation status from mothers to daughters before and after imputation of economic activity. In the original data without imputation the participation rate of those women whose mothers also work is unrealistically high, 51.16%. The figure goes down to 4.88% for the naive imputation procedure. Moreover, for the original dataset the proportion of working women increases from 6.52% to 51.16% —or 44.6 percentage points—when the mother also works. This change is substantially lower in the naive imputation—2.8 percentage points.

Importantly, compared to making no imputations, our imputation procedure lowers the differential in transition of participation rates from mother to daughters between mothers who worked and mothers who did not and therefore would bias downwards the effect of mothers' labor market participation.

TABLE 2.3
Distribution of Female Participation Conditional on the Mother's Participation Choice

	Non-working	Working	All
No Imputation			
Her mother did not work	545 (93.48%)	38 (**6.52%**)	583 (100%)
Her mother worked	21 (48.84%)	22 (**51.16%**)	43 (100%)
All	566 (90.42%)	60 (9.58%)	626 (100%)
Imputation 1 - all missings to inactive			
Her mother did not work	24,157 (97.94%)	508 (**2.06%**)	24,665 (100%)
Her mother worked	429 (95.12%)	22 (**4.88%**)	451 (100%)
All	24,586 (97.89%)	530 (2.11%)	25,116 (100%)

*Labor force participation is defined as either *employee/farmer* or *professional/capital owner*.

Imputation procedures are nonetheless subject to criticism as they are *ad hoc* procedures in the way they exploit available information on the missing data mechanism. As a consequence, the properties of the estimations obtained with the imputed data are not generally known. An alternative approach popular in econometrics is to propose a model for the missing data mechanism and estimate the model of interest subject to the missing data generation process. A seminal example is provided by Heckman (1976) within the literature of sample selection. Recently, Ramalho and Smith (2012) have proposed a likelihood-based approach to deal with missing data when the output variable is discrete, as it is in our case. In the next section, we propose a reduced-form participation model and illustrate how Ramalho and Smith (2012)'s procedure can be applied to account for missing information on economic activity.

2.4 External data on female labor market participation rates

The GMM procedure proposed by Ramalho and Smith allows for the incorporation of external data to help the identification of the parameters of interest. We are going to use census data on local and national female force participation rates as external sources of information. In Portugal, the first census was taken in 1864 and since then there are eleven more census from which we can gather relevant data. There are, however, difficulties in homogenizing the data on female labor force participation across censuses. We describe the sources and the computation of our measure of female participation in Appendix A.

3 The econometric model

3.1 A reduced-form participation model

We propose the estimation of a reduced-form participation model using church registry data at the individual level. Woman i chooses either to earn an income by participating in the labor market, $y_i = 1$, or to specialize in household production, $y_i = 0$, depending on which option gives her the greatest life-time expected utility. We assume that the discrete choice is expressed in the following linear specification:

$$y_i = \mathbf{1}\left\{\alpha y_i^m + x_i\beta + \nu_i > 0\right\} \tag{1}$$

where dummy variable y_i^m indicates the labor force participation status of her mother. Vector x_i includes discrete controls such as a dummy for large number of siblings, settlements dummies and quarter-century dummies.

Parameter α is the parameter of interest and captures the effect of the mother's participation choice on the daughter's participation decision. The vector β captures systematic differences in female participation across quarter-century and regions as well as other characteristics such as large families where we believe it to be more likely for daughters to have jobs. The error term ν_i incorporates effects of those variables for which we have no information. For example, this term may include the effect of educational choices by woman i unrelated to her parents economic activity decisions. Although in other studies children educational choices are usually assumed correlated with parents' professional choices, making interpretation of α difficult, this assumption is not realistic here: During the entire sample period the Portuguese educational system remained extraordinarily elitist, with illiteracy rates of over 80% and formal education being only accessible to a small privileged group (see, for example, http://www.country-data.com/).[11] Thus, it is reasonable to assume that in our data human capital is mainly acquired informally within

[11] Educational reforms initiated in 1822, 1835, and 1844 were mainly targeted towards boys' education and were left incomplete and largely unimplemented. Table 7 in Appendix B shows that ilitteracy rates among girls in the beggining of the 20th century were still very high in the North of Portugal although the pattern was different in the islands. Moreover, as stated in footnote 1 in the Introduction, secondary schools for girls were only approved after the 1888 educational reform.

the family and the parameter α can be interpreted as reflecting how the mother's participation in the labor market affects the daughter's preferences and/or beliefs towards participation in the labor market and, therefore, her net utility from working.

Other sources for the unobservable effect, ν_i, may be quasi-permanent factors, such as social status or genetic predisposition. These variables are an alternative source of persistence and, if relevant in this choice, would also affect y_i^m. When these quasi-permanent effects are present and not controlled for, they generally cause a bias in the estimate of the within-family cultural transmission parameter, α. The importance of these lineage-specific factors, however, may be in practice rather small, as both genetic and social traits are likely diluted quickly beyond the grandparents to grandchildren link. Unfortunately, a test on the relevance of lineage (quasi-permanent) versus parents (path dependent) effects seems difficult for theoretical and practical reasons. From a theoretical point of view, it is not clear how to identify the slowly decaying effect of the individual-specific full set of ancestors from the parents' direct effect. From a practical point of view, identification would require a large number of observations with the same lineage, i.e. a large number of siblings.[12]

Given the parsimonious specification for the determinants of y_i proposed in equations (1) and (**??**), one could argue that the estimate of α does not only capture the effect of the participation of the mother. Of course, a missing variable simultaneously correlated with the mother's and the daughter's participation decision would make the interpretation of the estimate as a causal effect invalid. In the Robustness Section, Section 4.3, we explore mechanisms beyond preferences and beliefs that could explain the effect of the mothers on the tranmission of female labor market participation.

3.2 Likelihood Approach to Deal with Nonignorable Missingness

Most observations in the women sample have missing entries in the set of discrete variables $\{y_i, y_i^m\}$ (see Table 2.2). Let us define a binary indicator I which takes value 1 if the participation status and occupational code for observation i is observed and 0 otherwise.

[12]Note that identification would still rely on the arguably strong assumption that siblings are affected by lineage in the same way, regardless of their birth order within the family, for example.

Similarly, let I^m take value 1 if the participation status and occupational code for the mother of i is observed and 0 otherwise. [13] Hence, we will consider the set of covariates $\{y_i, y_i^m, x_i\}$ where x_i is a vector of covariates which is always observable.

Random missingness arises when the probability of a missing observation is invariant to the set of available covariates in which case the missingness observations are ignorable in the sense that their omission from estimation does not bias the results. More realistically, missing occupational information is likely related to some of the individuals' characteristics as collection of this information by the vicar is influenced by them. For example, distant farms are less likely to be visited by an old vicar. In this situation, the missing mechanism is nonignorable in the sense that estimating the participation model only with the nonmissing observations will generally lead to inconsistent estimates.

Ramalho and Smith (2012) propose a likelihood-based approach to deal with nonignorable nonresponse in discrete choice models. More concretely, they propose a method to estimate a parametric conditional model of the discrete outcome when the marginal distribution of covariates is unknown but does not depend on the parameter vector of interest. As a special case they consider the situation whereby the missingness mechanism is conditionally dependent on the outcome variable and a discrete partition of the covariates. Hence, their approach is clearly valid in our setup.

Assume that variables y and y^m can only take up values $0, 1$ and that x is a discrete covariate vector. The aim is to estimate parameter vector θ in a conditional discrete choice model:

$$\Pr\{y = v \,| y^m = w, x\} = F\{v, w, x; \theta\}$$

where function F is known and $v, w \in \{0, 1\}$ and x denotes, for shortness, both the variable and a potential value. For an observation with nonmissing information, the joint probability of nonmissingness, i.e. $I_i = I_i^m = 1$, and the vector variables $\{y_i, y_i^m, x_i\}$ is without any loss of generality:

[13] In the rest of the paper, it will be assumed that missingness only affects the occupational variables. Missingness of other characteristics X (e.g. the father's occupation) is ignored, although the analysis can be extended to cover this more general case.

$$\begin{aligned}\Pr\{I_i = I_i^m = 1, y_i = v, y_i^m = w, x_i\} = \;& \Pr\{I_i = I_i^m = 1 \,|y_i = v, y_i^m = w, x_i\} \times \\ & F\{v, w, x; \theta\} \times \Pr\{y_i^m = w, x_i\}.\end{aligned} \tag{2}$$

The fundamental idea of Ramalho and Smith (2012) stems from the recognition that under certain circumstances missingness may be viewed as a modification of choice-based sampling. In our empirical application this situation is intuitively plausible. As argued before, some vicars may be more likely to underreport the incidence of professions that were common, such as farmers, and more likely to record the professions for those whose labor status was a differential characteristic (i.e. those with a profession other than the most common one in the area and period, like civil servants or the miller in a village of farmers), and leave the labor force status for housewives unreported. These considerations warrant the following:

Assumption 1 *(Daughter's Response Conditional Independence, DRCI) Nonmissingness in y is conditionally independent of I^m,y^m, and x; i.e.,*

$$\Pr\{I = 1 \,|I^m, y, y^m, x\} = \Pr\{I = 1 \,|y\}. \tag{3}$$

14

Given that the mother's information is likely collected early on and through a similar process, an assumption closely related to DRCI but referring to the availability of the mother's participation decision can also be made:

Assumption 2 *(Mother's Response Conditional Independence, MRCI) Nonmissingness in y^m is conditionally independent of I,y, and x; i.e.,*

$$\Pr\{I^m = 1 \,|I, y, y^m, x\} = \Pr\{I^m = 1 \,|y^m\}. \tag{4}$$

Let $H_1 \equiv \Pr\{I = 1, y = 1\}$, $H_0 \equiv \Pr\{I = 1, y = 0\}$and $H_1^m \equiv \Pr\{I^m = 1, y^m = 1\}$,

[14]As stressed by Ramalho and Smith (2012), a weaker assumption in which the conditional probability does depend on a subset of discrete covariates, such as period and location, is also implementable.

$H_0^m \equiv \Pr\{I^m = 1, y^m = 0\}$. Furthermore, the marginal distributions of the discrete variables y, y^m are parametrized by $\Pr\{y = 1\} = \Pi$ and $\Pr\{y^m = 1\} = \Pi^m$, respectively. Finally $\Pi_{w,x} \equiv \Pr\{y^m = w, x\}$ where the number of parameters in $\Pi_{w,x}$ is given by the number of different combinations between the variables y^m and x observed in the data with a maximum in our case of $2 \times comb(x)$ where $comb(x)$ is the theoretical number of combinations of the discrete vector x. Assumptions (3) and (4) imply that for example:

$$\Pr\{I_i = I_i^m = 1, y_i = y_i^m = 1, x_i\} = \left(\frac{H_1}{\Pi}\right)\left(\frac{H_1^m}{\Pi^m}\right) F\{1, 1, x_i; \theta\}\, \Pi_{1,x_i} \tag{5}$$

where Π_{1,x_i} is the parameter of the matrix $\Pi_{w,x}$ that corresponds to the specific combination of values of variables $(y^m, x_i) = (1, x)$.

There are three situations in which a given observation may have some missing information: when the daughter's information is missing but the mother's is not, when the mother's information is missing but the daughter's is not, and when information for both the daughter and the mother is missing. Consider the first case. The joint probability for $\{y_i^m = 1, x_i\}$ in the first case is:

$$\begin{aligned}
\Pr\{I_i = 0, I_i^m = 1, y_i^m = 1, x_i = x\} &= \{\Pr\{I_i = 0, I_i^m = 1, y_i = 1, y_i^m = 1, x_i\} \\
&\quad + \Pr\{I_i = 0, I_i^m = 1, y_i = 0, y_i^m = 1, x_i\}\} \\
&= \left(1 - \tfrac{H_1}{\Pi}\right)\left(\tfrac{H_1^m}{\Pi^m}\right) F\{1, 1, x_i; \theta\}\, \Pi_{1,x_i} + \left(1 - \tfrac{H_0}{1-\Pi}\right)\left(\tfrac{H_1^m}{\Pi^m}\right) F\{0, 1, x_i; \theta\}\, \Pi_{1,x_i} \\
&= \left\{\left(1 - \tfrac{H_1}{\Pi}\right) F\{1, 1, x_i; \theta\} + \left(1 - \tfrac{H_0}{1-\Pi}\right) F\{0, 1, x_i; \theta\}\right\}\left(\tfrac{H_1^m}{\Pi^m}\right) \Pi_{1,x_i}
\end{aligned} \tag{6}$$

Since:

$$\Pr\{I_i = 0 \,|I_i^m = 1, y_i = 1, y_i^m = 1, x_i\} = \Pr\{I_i = 0 \,|y_i = 1\} = 1 - \tfrac{H_1}{\Pi}$$

and

$$\Pr\{I_i = 0 \,|I_i^m = 1, y_i = 0, y_i^m = 1, x_i\} = \Pr\{I_i = 0 \,|y_i = 0\} = 1 - \tfrac{H_0}{1-\Pi}.$$

Following a similar argument, it is possible to show that the joint probability of an observation which does not have the mother's participation decision but has the daughther's and is given by $\{y_i = 1, x_i\}$ is of the form:

$$\Pr\{I_i = 1, I_i^m = 0, y_i = 1, x_i\} = \left(\frac{H_1}{\Pi}\right)\left\{\left(1 - \frac{H_1^m}{\Pi^m}\right) F\{1, 1, x_i; \theta\} \Pi_{1,x_i} + \left(1 - \frac{H_0^m}{1 - \Pi^m}\right) F\{1, 0, x_i; \theta\} \tag{7}$$

Finally, the joint probability of an observation which has neither the daughter's nor the mother's participation decision and only the $\{x_i\}$ is observable is:

$$\begin{aligned}\Pr\{I_i = 0, I_i^m = 0, x_i\} = &\left\{\left(1 - \tfrac{H_1}{\Pi}\right) F\{1, 1, x_i; \theta\} + \left(1 - \tfrac{H_0}{1-\Pi}\right) F\{0, 1, x_i; \theta\}\right\} \times \left(1 - \tfrac{H_1^m}{\Pi^m}\right) \Pi_{1,x_i} \\ &\left\{\left(1 - \tfrac{H_1}{\Pi}\right) F\{1, 0, x_i; \theta\} + \left(1 - \tfrac{H_0}{1-\Pi}\right) F\{0, 0, x_i; \theta\}\right\} \times \left(1 - \tfrac{H_0^m}{1-\Pi^m}\right) \Pi_{0,x_i}\end{aligned} \tag{8}$$

Expressions (5), (6), (7), and (8) can be more generally expressed as:

$$\begin{aligned}\Pr\{I_i, I_i^m, y_i, y_i^m, x_i\} = \;& \left(\left(\tfrac{H_{y_i}}{\Pi_{y_i}}\right)\left(\tfrac{H^m_{y_i^m}}{\Pi^m_{y_i^m}}\right) F\{y_i, y_i^m, x_i; \theta\} \Pi_{y_i^m, x_i}\right)^{I_i I_i^m} \times \\ & \left(\left\{\textstyle\sum_{v\in\{0,1\}} \left(1 - \tfrac{H_v}{\Pi_v}\right)\left(\tfrac{H^m_{y_i^m}}{\Pi^m_{y_i^m}}\right) F\{v, y_i^m, x_i; \theta\} \Pi_{y_i^m, x_i}\right\}\right)^{(1-I_i) I_i^m} \times \\ & \left(\textstyle\sum_{w\in\{0,1\}} \left\{\left(\tfrac{H_{y_i}}{\Pi_{y_i}}\right)\left(1 - \tfrac{H_w^m}{\Pi_w^m}\right) F\{y_i, w, x_i; \theta\} \Pi_{w, x_i}\right\}\right)^{I_i (1 - I_i^m)} \times \\ & \textstyle\sum_{v,w\in\in\{0,1\}} \left\{\left(1 - \tfrac{H_v}{\Pi_v}\right)\left(1 - \tfrac{H_w^m}{\Pi_w^m}\right) F\{v, w, x_i; \theta\} \Pi_{w, x_i}\right\}^{(1-I_i)(1-I_i^m)}.\end{aligned} \tag{9}$$

with $\Pi_0 = \Pi$ and $\Pi_1 = 1 - \Pi$ and similarly for Π^m and where the meaning of subscript i in a given variable is that the function is to be evaluated at the value of the variable at observation i and whenever the subscript i is used in parameters it means that the relevant parameter is the one that corresponds to the value at that observation. Hence, for example, $F\{y_i, w, x_i; \theta\}$ in the third row of (9) should be evaluated at the value of y and x of observation i and a running value for w (that is not necessarily the one of observation i). Extending the vector of parameters to include in addition to θ, the probabilities$\{H_v\}$, for $v \in \{0, 1\}$,$\{H_w^m\}$, for $w \in \{0, 1\}$, and $\{\Pi_{w,x}\}$ which has as many parameters as the possible combinations of the values of the vector y^m and x found in the data, then (9) represents the likelihood for any given observation i. Maximum Likelihood estimation is subject to the following constraints:

$$\begin{aligned}
&\textstyle\sum_{w,x} \Pi_{w,x} = 1 \text{ for } w = 0, 1 \\
&\Pi_v = \textstyle\sum_{w,x} F\{v, w, x; \theta\} \Pi_{w,x} \text{ for } v = 0, 1 \\
&\Pi_w^m = \textstyle\sum_x \Pi_{w,x} \text{ for } w = 0, 1
\end{aligned} \tag{10}$$

will give consistent and asymptotically efficient estimates for vector parameter of interest θ. One of the difficulties produced by this model is that the vector of parameters $\Pi_{w,x}$ grows with the number of exogenous variables in x and depends on the number of possible combinations of all values in x. It is reasonable to reduce the number of parameters to estimate by introducing some restrictions in $\Pi_{w,x}$. In our specific framework, we are going to decompose the vector x into two types of variables x_1 and x_2 where x_1 is location and quarter-century and x_2 is everything else.

$$\Pi_{w,x} = \Pi_{x_2|w,x_1} \Pi_{w,x_1}$$

Our simplifying assumption is that the distribution of x_2 only depends on the location and quarter-century and not on the mother working status which implies:

$$\Pi_{w,x} = \Pi_{x_2|w,x_1} \Pi_{w,x_1} = \Pi_{x_2|x_1} \Pi_{w,x_1}$$

Finally, the accuracy of our estimated parameters may be improved upon by adding estimators of Π_{w,x_1} which were based on external data sources.

3.3 GMM Estimation using external information

Following Ramalho and Smith (2012) in their application of Imbens and Lancaster (1994), we can improve our estimates of θ by incorporating additional information on the marginal probabilities Π and Π^m In Appendix A we take census data for the end of the 19th century and 20th century on national and regional female participation rates, we project these values using a reduced form model to get estimates for the female participation rates for our sample period in our four sites. These estimates of $\hat{\Pi}$ and $\hat{\Pi}^m$ and their standard errors play the role of $\hat{Q}_v$ and $\hat{\Sigma}$ in Ramalho and Smith (2012), section five.

4 Results

4.1 ML estimates using the imputation methods

In this Section we present the results of probit estimations of the model (1). Table 4.1 presents the coefficient estimates of parameter α for different samples and different missing imputation methods including no imputation.[15] The first row shows results for the full sample, the seond row restricts the sample to Horta, where the problem of missings is less acute and finally, the last row further restricts the results to Horta in the 19th century where the percentage of missings is the lowest.[16]

Column I of Table 4.1 reports estimates of α using the naive imputation or "Imputation 1" where we impute all the missings as unpaid farmers/domestic production, i.e. non-participating in the labor market. Column II shows the coefficient estimates when imputation 1 is restricted to the sample where mothers are identified, denoted by "Imputation 1b)". Imputation 1b) implies loosing approximately 5000 observations relative to imputation 1. Column III shows results obtained when imputation 1 is applied only to rural settlements which implies leaving the missing observations from Horta out of the sample. We denote this imputation by "Imputation 2". The idea here is to impute only non-paid farmers. Since it is less likely to find farmers in Horta, we do not impute any of the missings in Horta. Column IV, follows the approach often used by historians and impute all wives of farmers as non-paid farmers. We denote this imputation by "Imputation 3". Finally, on the last column of Table 4.1 we do not impute any of the missing values, which reduced the sample size drastically. All regressions in Table 4.1 include quarter-century dummies, location dummies —more precisely a dummy for the Minho

[15] Table XX in Appendix B shows the complete set of coefficient estimates [Table to be added].

[16] A second reason to restrict the sample to the 19th century lies in the lower standard errors of the predicted Π_{w,x_1} for this sample. As we explain in Appendix A, $\hat{\Pi}_{w,x_1}$ are obtained from regressions that use census data from the end of the 19th century till the 1991. The standard errors of the out-of-sample prediction of Π_{w,x_1} increase as we go further backwards in time.

region (included Ronfe and Ruivães) and a dummy for the islands (includes Horta and S. Mateus)— and a dummy for large family size which takes value 1 if the female has more than four siblings.

TABLE 4.1 Probit Estimations of α

	Imp. 1	Imp 1b	Imp. 2	Imp. 3	No Imp.
Full Sample	0.121	0.369^{***}	0.43^{***}	1.313^{***}	0.926^{***}
	(0.106)	(0.109)	(0.116)	(0.206)	(0.239)
Horta	0.433^{***}	0.406^{***}	1.132^{***}	1.134^{***}	1.132^{***}
	(0.153)	(0.155)	(0.251)	(0.251)	(0.252)
Horta 19th century	0.443^{***}	0.427^{***}	1.132^{***}	1.134^{***}	1.132^{***}
	(0.154)	(0.156)	(0.251)	(0.251)	(0.251)

Imp 1b is Imp 1 for observations with identified mother. Regressions include: quarter-century dummies, dummy for Minho, dummy for number of siblings>4

Importantly, we find a positive and statistically significant effect of the mother working status on the daughter's probability of working (α) in all estimations but in Imputation 1-full sample. Moreover, as predicted when discussing Table 2.3 above, the imputation missings to inactivity goes against finding an effect of the mothers which is generally speaking what can be observed from Table 4.1 where the estimated α increases as we move from left to right. Notice that for Horta, imputation 2, 3 and no imputation are quite similar and that is why the results are virtually identical.

4.2 Dealing with missing occupations via ML estimation

In this Section, we show the results obtained with the Ramalho and Smith (2012) methodology developed for our particular case in Section 3.2. In the first five columns of Table 4.2 we report the estimated α obtained from simple probits under the different imputation methods (same as those shown in Table 4.1). In the last column, we report the estimated α obtained through the Ramalho & Smith method. The upper pannel of the Table shows the estimations done with the full sample and the bottom part shows the estimations done when restricting to the 19th century sample. In the second row of each pannel, we report the estimated marginal effects. In both panels, the estimated value of α and its marginal effect is within the range obtained with the probit estimations and the different imputations. Comparing the value of $\hat{\alpha}$ obtained with the Ramalho and Smith methodology with the values obtained from the probit estimations, we may conclude that

Imputation 2 is rather unbiased. It is important, however, to notice that the the marginal effects of the two estimation procedures are rather different. Hence, estimating a probit model using Imputation 2 would have led us to conclude that the marginal effect of a working mother on the probability of participation in the labor market was about a third of the true value; a substancial underestimation of the true effect.

TABLE 4.2

	Simple Probits					R&S
	Imp. 1	Imp 1b	Imp. 2	Imp. 3	No Imp.	No Imp.
	All observations					
$\hat{\alpha}$	0.121	0.369***	0.430***	1.313***	0.926***	**0.487*** **
	(0.106)	(0.109)	(0.116)	(0.206)	(0.239)	(0.218)
Mg. Eff.	0.006	0.018	0.028	0.189	0.193	0.086
Pr. LFP	–	–	–	–	–	11.0%
Nobs	25116	18974	14173	2431	523	25116
	19th century					
$\hat{\alpha}$	0.165	0.383***	0.526***	1.327***	0.927***	**0.640*** **
	(0.109)	(0.111)	(0.123)	(0.208)	(0.239)	(0.222)
Mg. Eff.	0.013	0.030	0.049	0.267	0.231	0.082
Pr. LFP	–	–	–	–	–	27.6%
Nobs	13696	10681	8419	1549	410	13696

Regressions include: quarter-century dummies, dummy for Minho, dummy for number of siblings>4. Predicted LFP 1700-1900 using census data is 14.3 percent. A marginal effect of 0.086 represents an increase of 60-78.2 percent.

Notice also that the marginal effect of this variable is smaller than the unconditional effect reported in Table 2.3, and is also smaller than the effect found if we use the original data with no imputation of missing occupations.

4.3 Robustness Section - Looking for other Mechanisms beyond Preferences

5 Conclusions and future work+

In this paper, we use historical parish registry data from four Portuguese settlements from the 18th and 19th centuries to estimate a female labor force participation model to identify the effect of mothers' labor market experience on that of their daughters. The period and locations chosen allow us to interpret the effect of mothers on daughters labor force participation as intergenerational transmission of preferences and beliefs. First, they were characterized by a very low male-female population ratio which open a window of opportunities to women. Second, the period studied is previous to any technological change. Third, given the high levels of illiteracy, particularly among women, it is unlikely that direct human capital transmission or indirect transmission through educational choices may explain the mother's effect on daughters choices.

After controlling for settlement and century dummies and some family covariates, we find a positive effect of the mother's working status on the daughter's decision to participate in the labor market. This effect is smaller than the unconditional effect, but it is strongly statistically significantly different from zero and is robust to the imputation method we use to address the problem of missing information on occupational status.

Future Work: Although we have shown that the naive imputation method does not directly drive the main result of the study, i.e. the positive and significant effect of the mother working condition on the daughter's participation decision, there are at least two ways by which we could further study to what extent are the results sensitive to alternative imputation methods. First, we could simulate more sophisticated imputation methods based as described in footnote 8. Second, the adaptation and implementation of Ramalho and Smith (2012)'s approach to our dynamic setup, as presented in Section 3 and complemented with external historical data.

Given the small set of determinants of y_{il} that we observe, $\alpha_1$would not include only the causal effect between the participation of the mother and the participation of the daughter, but would also capture the transmission of other determinants that would affect y_{il} on their own. These would include both unobservable matrilineage time invariant heterogeneity and the conditioning effect of endogenous sample selection due to migration of some descendants or ancestors. This latter effect invalidates the causal effect interpretation for the estimate for α_1 if the migration decision is affected by unobserved factors that also affect the labor force participation decision, such as ability.

A more detailed control for unobserved heterogeneity is the next step towards being able to interpret our results as causal effects and be sure that the estimated AME of ym_{il} is the magnitude of the effect of intergenerational human capital transmission. The most obvious extension, in line with the usual practice in panel data literature, is to allow for a lineage-specific permanent unobserved factor. These factors will compose a special unobservable part that is constant for all i in the same l. ν_{il} would be decomposed in two terms:$\nu_{il} = \mu_l + \varepsilon_{il}$. μ_l and can be separately identified because we have several observations for each matrilineage. On the other hand, it is plausible that the lineage-specific variables mentioned do not have a permanent effect throughout all descendants, but are diluted rather quickly as we go further down in the lineage, making them not permanent and only relevantly transmitted from parents to children or grandparents to grandchildren. Furthermore, the fact that adding the participation of the grandmother does not change our conclusions limits the presence of these factors to one generation (parents to children). In this case, the only way of identifying μ separately from ym_{il} is to use siblings. The problem is then that the number of observations with the same μ is very small.

We want to interact $ym_i$with period to compute to what extent $\alpha_1$changes with time. We want to check the effects of grandfathers and mothers in law.

Another important way in which specification (1) may be too restrictive is in the homogeneity assumption of all the coefficients. A relaxation of this strong assumption does not only control for additional unobservable effects, but also provides an econometric specification that does not rule out economic models. See Browning and Carro (2007) and Browning and Carro (2009) for more information on this.

material deleted from introduction

Theories of female labor participation which invoke the importance of transmission of beliefs and/or preferences through family experience are not easy to test. Fernández and Flogi (2009) and Fernández (2007) use country-level female labor force participation rates from the parents' country of origin to explain differences in paid working hours of daughters of US immigrants. These papers represent an effort to separate culture from the environment by studying women in a common institutional and economic environment, what Fernández denotes as the "epidemiological approach" (Fernández, 2011). The premise is that "culture," i.e. a set of preferences and beliefs common to a group, is transmitted from parents to children. Costa et al (2001), however, show that gender differences in personality vary across cultures but are small relative to individual variation within genders. Part of this individual variation may be explained through family effects, which can be disentangled from general social norms and attitudes. Fernandez, Fogli and Olivetti's (2004) model of the change in social norms, for example, can be interpreted as a process that takes place at the family level where the presence of a working mother during childhood changes preferences of sons and/or daughters towards female labor force participation. Some studies find a relation between parents' labor market behavior and attitudes and those of their sons and daughters (see, for example, Bütikofer, 2011, Farré and Vella, 2007, and Tsukahara, 2007). These effects are likely to compound both changes in preferences/beliefs and the transmission of specific human capital. The indirect evidence provided by these studies is rather persuasive of the importance of family effects. However, direct evidence remains, to our knowledge, nonexistent. In this paper, we use a novel approach based on historical data at the onset of female labor market participation to attempt to provide more direct evidence on the intergenerational transmission of preferences and beliefs.

6 Appendix A

In this Appendix we describe the data used to obtain estimations of female labor participation which are included in the GMM estimation of our model in Section 3.3. Our main data source in this Section is the population censuses (*Recenseamentos Gerais da População*) for the period 1890-1991. The data for Portugal 1864 was taken from Reis (2005). The first Portuguese census was taken in 1864 and since then there are eleven censuses more or less periodical that contain information on female labor activities. The smallest geographical level for which demographic data is collected is the Borough (*Concelho*), followed by the District (*Distrito*), and the Province. Most censuses also publish information on economic activity and even professions of men and women above a certain age at various levels of geographical aggregation, which unfortunately varied over the censuses. The censuses collected data for all regions of Portugal including the archipelagos of Madeira and Azores.

For the purpose of estimating the female labor force participation, we only assemble data from regions that were either more representative of the villages in our sample or of Portugal as a whole. The data assembled are a mixture of Borough and District level data and data for Portugal as a whole. Specifically, the regions are: 1) Portugal (including Azores and Madeira and excluding colonial territories); 2) the district of Braga to which Ronfe and Ruivães belong; 3) the district of Horta to which the city of Horta and São Mateus belong; 4) the Boroughs of Lisbon and Oporto which are the largest cities in the country; 5) the Borough of Guimarães to which Ronfe belongs; 6) the Borough of Vila Nova de Famalicão to which Ruivães belong; 7) the Borough of Horta; 8) and finally the Borough of Madalena in Pico island to which São Mateus belongs.

Table 6 shows our best rough calculation of the female labor market participation for nine regions in Portugal over the period 1864-1991 based on the censuses data. Although we use a single data source (with the exception of 1864), the participation rates reported in Table 6 show large jumps across census —which are most likely caused by the changes in the definition of "active" female population during the period— and large differences across regions for a given census. Carrilho (1996) provides a good guide for the different definitions of active female population across the censuses. One of the main differences regards the population of reference. For example, while until 1930 the reference population was considered the "present" population (*população de facto*), after 1940 the censuses considered the "resident" population. Moreover, while until 1930, all individuals were included in the reference population, after 1940 only those above a certain age (either 10 or 12 depending on the census) were considered. The other main difference refers to who was considered "active". The earlier censuses (roughly until 1950) defined as active population everyone with an occupation regardless of whether or not that occupation was a profesion—for example, domestic or agricultural unpaid work was considered an occupation although not a profession—. This definition implied very high and unrealistic participation rates for women, specially in the less urban districts and boroughs (e.g. compare Lisbon with Guimarães). The islands are somewhat of an exception where the rural borough of Madalena has a very low participation even lower than the capital of the district Horta. Since 1960, the unemployed who are looking for a job are also included as part of the active population. Note that for all figures in the Table the denominator allways includes the elderly in the reference population. That is the reason why the values in the table are so low compared to the values of female participation rates of 67% in the nineties given in the Introduction.

Altgough in Table 6's data is as much homogeneous as possible there remain inconsistencies in the definitions of female participation (see footnote to Table 6 for a detailed description of our definition in each census year). Given these difficulties, we decided to use controls in the regression model to correct for those cases where full homogenization is not possible. We intend to use the regression model to predict backwards values of female participation in the four settlements (Ronfe, Ruivães, Horta and São Mateus) in the 8 quarter-centuries of the 18th and 19th century.

In Figure 1, we ploted the data from Table 6 for Portugal (solid line), the fitted values obtained from the regression model and their 99% confidence interval. In order to stress the differences in definitions across censuses we connected only the data points which have similar definitions (from 1890-1911 and from 1930 onwards). The data point corresponding to 1864 is singled out because it was taken from a different source (Reis,2005) however it looks perhaps surprisingly similar to the values from 1930 onwards. Notice that the data values for 1890, 1900, 1911 look somewhat inflated compared to the rest of the data points. The reason why these values are so elevated is due to the inclusion of unpaid farmers and landlords, whose only source of income was rents, in the definition of active females. After 1940, the censuses clearly makes a distinction between occupation and profession: whilst most women have an occupation only a few have a profession. In 1930, we are able to obtain a reasonable estimate by excluding the housewives and the family members. The decreasing trend in years 1890-1911 is probably due to the lower relative importance of agriculture. The 1890-1911 data, however, gives useful information regarding the differences across regions in Portugal. Therefore, in our econometric model we estimate a separate trend for this period (as seen in discontinuous line in Figure 1) which is shut in our predictions as can be seen by the smooth line of predicted values.

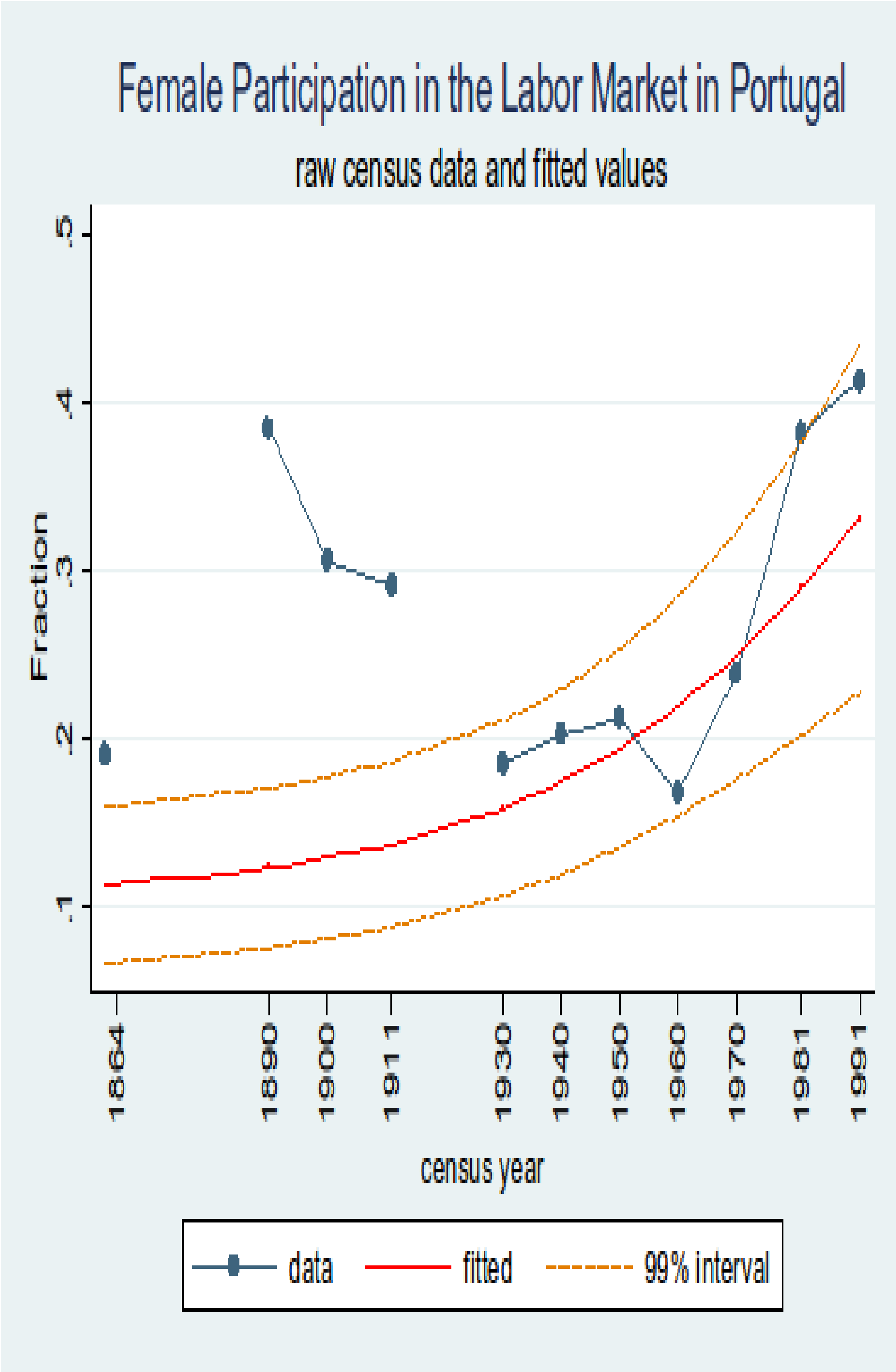

Figure 1: Female Participation in Portugal 1864-1991

TABLE 6 - Data on Female Participation Rates

	Portugal	Braga (Dist.)	Lisbon	Oporto	Horta (Dist.)	Guimarães	V. N. Famalicão	Horta (city)	Madalena
1864	0.191	*n.a*	*n.a*	*n.a*	*n.a*	*n.a*	*n.a*	*n.a*	*n.a*
1890	0.385	0.513	0.230	0.338	0.103	0.552	0.622	0.111	0.038
1900	0.307	0.534	0.276	0.380	0.146	0.532	0.511	0.187	0.100
1911	0.292	0.413	0.281	0.403	0.188	0.539	0.364	0.217	0.162
1925(*)	*n.a*	*n.a*	0.284	0.259	*n.a*	*n.a*	*n.a*	*n.a*	*n.a*
1930	0.186	0.335	0.261	0.317	0.079	0.375	0.322	0.092	0.062
1940	0.203	0.329	0.258	0.332	0.065	*n.a*	*n.a*	*n.a*	*n.a*
1950	0.213	0.301	0.287	0.343	0.061	0.423	0.366	0.078	0.054
1960	0.169	0.237	0.315	0.365	0.055	0.336	0.288	0.084	0.029
1970	0.240	0.334	0.328	0.391	0.073	*n.a*	*n.a*	*n.a*	*n.a*
1981	0.382	0.483	0.439	0.466	*n.a*	0.590	0.532	0.217	0.146
1991	0.437	*n.a*	0.433	0.462	*n.a*	0.575	0.550	0.325	0.286

Source: Authors computations from Recenseamentos Gerais da Populaçao, INE, Lisbon.

Notes: (*) The 1925 census was a special census restricted to the cities of Lisbon and Oporto

(1) In 1864, active female population was taken from the total number (all ages) of working women (*mão-de-obra feminina*) in Reis (2005). In the denominator: female population older than 10.

(2) In 1890, 1900, 1911 active female population = No. of females - housewives (unpaid domestic work) - improductive + servants. Information taken from Table V (1890, 1900), Table 3 (1911). In the denominator: female population older than 10.

(3) In 1925, active female population = No. of females - No. female children between 0-9 years old - housewives (unpaid domestic work) - females without profession - female beggars. In the denominator, female population older than 10.

(4) In 1930, active female population = sum the three first columns under "populaçao activa" i.e. females who work for the administration (local or state) + females who work in the private sector + self-employed females- housewives (unpaid domestic work) who are considered in the column of active and self-employed. Information from Table 1. In the denominator, female population older than 10.

(5) In 1940, active female population = No. of active females ("activas") - No. females in non-professional activities. Information from Table 24. Female population older than 10.

(6) In 1950, active female population = No. of active females ("activas") with profession. Information from Table 1. Female population older than 12

(7) In 1960, active female population = No. of active females ("activas") with profession - active females with profession less than 15 years of age. Information from Table 1 (Tomo 5, vol III). Female population older than 15

(8) In 1970, active female population = No. of active females ("activas") with profession. Information from Table 8. Female population older than 15 years old.

(9) In 1981, active female population = No. of active females ("activas") older than 12 years old. Information from Table 6.13. In the denominator, female population older than 14

(10) In 1991, active female population = No. of active females with economic activity (minimum age 12 years old). Information from Table 6.16.1. (pp138). Female population older than 12.

The regression model estimated is the following:

$$y_{it} = c + \delta_1 f(t) + \delta_2 g(t) + \sum_{r=1}^{R} \gamma_r D_r + \sum_{k} \gamma_k D_k + \varepsilon_{it} \tag{11}$$

where $y_{it} = ln\left(\frac{p_{it}}{1-p_{it}}\right)$where p_{it}corresponds to the values of the female participation rate given in Table 6 for region i and census year t. $f(t)$corresponds to a logit transformation of the census year t, specifically $f(t) = \frac{\exp\left(\frac{(t-2000)}{40}\right)}{1+\exp\left(\frac{(t-2000)}{40}\right)}$ where the inflexión year of the logistic was chosen such as to better fit the data. The function $g(t) = 1(t \leq 1940)(t-1940)$ is a specific trend for the values before 1940 in order to prevent the inflated values of our data before 1940 and its unlikely trend to contaminate our predictions, whist keeping the information regarding the differences across regions for those years. The $D_r's$ represent regional dummies, specifically, a dummy for the Braga district (encompasses Braga district, Guimaraes and V.N. Famalicão), a dummy for the Horta district (encompasses, Horta district, the city of Horta and Madalena), a dummy for a large city (encompasses Lisbon and Oporto), a dummy for Guimarães, a dummy for V.N. Famalicão, a dummy for Horta (city), a dummy for Madalena. The $D_k's$ are other dummies used to take into account specificities of the particular censuses such as a dummy for years where the participation rates are applied to women older than 14 years old (1960, 1970, 1981), and a specific dummy for 1981 where there is a particular high jump in the participation rate due the course way in which the data is reported in the census. Finally, $\varepsilon_{it}$is an error term.

We estimate the model (11) by OLS and use the estimated parameters to predict participation rates for Ronfe, Ruivães, Horta, and São Mateus for our sample period $1700-1900$. Since the dependent variable is a nonlinear function of p_{it}, we first transform the model so that our predicted values for p_{it} are unbiased estimates. For each of the four regions, we activate the dummies of their borough and district e.g. for Ronfe, we set the dummy for the Braga district and the dummy for Guimarães Borough to 1. We further, set the $g(t)$ function to zero as well as the dummy for older than 14, the 1981 dummy, and the dummy for large city.

Table 6 shows the estimated female participation rates for Ronfe, Ruivães, Horta, and

S. Mateus for our sample period 1700-1900. The values are low, as expected, and quite constant.

TABLE 6 - Predicted Participation Rates for our sample period

	1700-1725	1800-1825	1875-1900
Ronfe	0.210	0.214	0.242
	(.0672)	(.0682)	(.0740)
Ruivães	0.265	0.270	0.301
	(.087)	(.0873)	(.0924)
Horta	0.047	0.049	0.056
	(.0079)	(.0080)	(.0083)
São Mateus	0.026	0.027	0.030
	(.0335)	(.0342)	(.0342)

*Note: Given that estimated participation rates are rather constant, we decided to show only 3 representative moments of our sample period. Predictions are obtained from the predicted coefficients of regression (11) of female participation rates from 12 censuses since 1864.

7 Appendix B

TABLE 7 - % of boys and girls 10-15 years old who can read

Borough	Population Census 1911			
	age 5-9		age 10-14	
	Girls	Boys	Girls	Boys
Guimarães (Ronfe)	6.6%	10.5%	22.6%	37.2%
V. N. Famalicão (Ruivães)	6.6%	10.8%	19.4%	37.8%
Horta	19.9%	13.6%	70%	41.2%
Madalena (São Mateus)	19.9%	19.3%	58.2%	45.9%

*Source: recenseamentos Gerais da População, 1911, INE, Lisbon, Portugal. Note: the figure for girls 10-14 in Horta seems exagerated when compared to boys of the same age. Higher literacy rates are expected in the cities such as Horta, however, this does not explain the relatively lower rate for boys in the Azores. Possible explanations may be related to the low presence of boys in the schools or home due to either early male migration or work at the sea.

Most common Professions/Social Status for Women in Lent Files			
18th and 19th centuries			
	Freq.	Percent	Cum.
lady not noble	921	45.82	45.82
housewife	335	16.67	62.49
slave	147	7.31	69.8
owner/landlord	132	6.57	76.37
craftsman textile	127	6.32	82.69
non-qualified worker	76	3.78	86.47
domestic servant	57	2.84	89.3
indigent/beggar	57	2.84	92.14
landless worker, day laborer	38	1.74	93.88
non-qualified worker manufacture/industry	23	1.14	95.02
farmer	20	1	96.02
high society	18	0.9	96.92
assistant	6	0.3	97.21
miller - owner	6	0.3	97.51
high rank clergy member	6	0.3	97.81
non-qualified worker construction	5	0.25	98.06
teacher	5	0.25	98.31
owner/manufacture/industry	5	0.25	98.56
craftsman wood	4	0.2	98.76
craftsman of tools	4	0.2	98.96
business owner	3	0.15	99.25
landless worker, day laborer - domestic	2	0.1	99.35
small business owner	2	0.1	99.45
clergy member	2	0.1	99.55
other	9	0.1	100

8 Appendix C - Gradient of log likelihood Function

In this appendix we write down the gradient of the likelihood function with respect to the parameters of the model. For some of the parameters of the model, the gradient programmed in the estimation code is a transformation of these gradients to accommodate restrictions such as that probabilities need to be between 0 and 1.

$$\frac{\partial ln\mathcal{L}}{\partial H_y} = I_i I_i^m \left(\mathbb{1}\left(y_i = y\right) \frac{1}{H_y} \right) + \left(1 - I_i\right) I_i^m \frac{-\frac{1}{\Pi_y} F\left(y, y_i^m, x_i; \theta\right)}{\sum_{v\in\{0,1\}} \left(1 - \frac{H_v}{\Pi_y}\right) F\left(v, y_i^m, x_i; \theta\right)} +$$
$$I_i \left(1 - I_i^m\right) \left(\mathbb{1}\left(y_i = y\right) \frac{1}{H_y} \right) + \left(1 - I_i\right)\left(1 - I_i^m\right) \frac{\sum_{w\in\{0,1\}} -\frac{1}{\Pi_y} \left(1 - \frac{H_w^m}{\Pi_w^m}\right) F\left(y, w, x_i; \theta\right) \Pi_{wx_i}}{\sum\sum_{v,w\in\{0,1\}} \left(1 - \frac{H_v}{\Pi_v}\right) \left(1 - \frac{H_w^m}{\Pi_w^m}\right) F\left(v, w, x_i; \theta\right) \Pi_{wx_i}} \tag{12}$$

$$\frac{\partial ln\mathcal{L}}{\partial H_{y^m}^m} = I_i I_i^m \left(\mathbb{1}\left(y_i^m = y^m\right) \frac{1}{H_{y^m}^m} \right) + \left(1 - I_i\right) I_i^m \left(\mathbb{1}\left(y_i^m = y^m\right) \frac{1}{H_{y^m}^m} \right) +$$
$$I_i \left(1 - I_i^m\right) \frac{-\frac{1}{\Pi_{y^m}^m} F\left(y_i, y^m, x_i; \theta\right) \Pi_{y^m x_i}}{\sum_{w\in\{0,1\}} \left(1 - \frac{H_w^m}{\Pi_w^m}\right) F\left(y_i, w, x_i; \theta\right) \Pi_{wx_i}} +$$
$$\left(1 - I_i\right)\left(1 - I_i^m\right) \frac{\sum_{v\in\{0,1\}} -\frac{1}{\Pi_{y^m}^m} \left(1 - \frac{H_v}{\Pi_v}\right) F\left(v, y^m, x_i; \theta\right) \Pi_{y^m x_i}}{\sum\sum_{v,w\in\{0,1\}} \left(1 - \frac{H_v}{\Pi_v}\right) \left(1 - \frac{H_w^m}{\Pi_w^m}\right) F\left(v, w, x_i; \theta\right) \Pi_{wx_i}} \tag{13}$$

$$
\frac{\partial ln\mathcal{L}}{\partial\theta} = I_i I_i^m \left\{ -\frac{\sum\sum_{w\in\{0,1\},x} F_\theta\left(y_i, w, x; \theta\right)\Pi_{wx}}{\Pi_{y_i}} + \frac{F_\theta\left(y_i, y_i^m, x_i; \theta\right)}{F\left(y_i, y_i^m, x_i; \theta\right)} \right\} + \left(1 - I_i\right) I_i^m \times
$$

$$
\left\{ \frac{\sum_{v\in\{0,1\}} \left[\left(1-\frac{H_v}{\Pi_v}\right) F_\theta\left(v, y_i^m, x_i; \theta\right) + \frac{\partial\Pi_v}{\partial\theta}\frac{H_v}{\Pi_v}\frac{1}{\Pi_v} F\left(v, y_i^m, x_i; \theta\right) \right]}{\sum_{v\in\{0,1\}} \left(1-\frac{H_v}{\Pi_v}\right) F\left(v, y_i^m, x_i; \theta\right)} \right\}
$$

$$
I_i\left(1 - I_i^m\right) \left\{ -\frac{\sum\sum_{w\in\{0,1\},x} F_\theta\left(y_i, w, x; \theta\right)\Pi_{wx}}{\Pi_{y_i}} + \frac{\sum_{w\in\{0,1\}} \left(1-\frac{H_w^m}{\Pi_w^m}\right) F_\theta\left(y_i, w, x_i; \theta\right)\Pi_{wx_i}}{\sum_{w\in\{0,1\}} \left(1-\frac{H_w^m}{\Pi_w^m}\right) F\left(y_i, w, x_i; \theta\right)\Pi_{wx_i}} \right\} +
$$

$$
\left(1 - I_i\right)\left(1 - I_i^m\right) \frac{1}{\sum\sum_{v,w\in\{0,1\}} \left(1-\frac{H_v}{\Pi_v}\right)\left(1-\frac{H_w^m}{\Pi_w^m}\right) F\left(v, w, x_i; \theta\right)\Pi_{wx_i}} \times
$$

$$
\left\{ \sum_{v\in\{0,1\}} \sum_{w\in\{0,1\}} \left[\left(1-\frac{H_v}{\Pi_v}\right)\left(1-\frac{H_w^m}{\Pi_w^m}\right) F_\theta\left(v, w, x_i; \theta\right)\Pi_{wx_i} + \frac{\partial\Pi_v}{\partial\theta}\frac{H_v}{\Pi_v}\frac{1}{\Pi_v}\left(1-\frac{H_w^m}{\Pi_w^m}\right) F\left(v, w, x_i; \theta\right)\Pi_{wx_i} \right] \right.
$$

$$
\text{where } \frac{\partial\Pi_v}{\partial\theta} = \left(\sum \sum_{w\in\{0,1\},x} F_\theta\left(v, w, x; \theta\right)\Pi_{wx} \right) \quad (14)
$$

$$\frac{\partial ln\mathcal{L}}{\partial \Pi_{zx^0}} = I_i^m \mathbb{1}\left(y_i^m = z\right)\mathbb{1}\left(x_i = x^0\right)\frac{1}{\Pi_{zx^0}} - I_i I_i^m \left(\frac{\partial \Pi_{y_i}}{\partial \Pi_{zx^0}}\right)\frac{1}{\Pi_{y_i}} - I_i I_i^m \mathbb{1}\left(y_i^m = z\right)\left(\frac{\partial \Pi_z^m}{\partial \Pi_{zx^0}}\right)\frac{1}{\Pi_z^m} +$$

$$\left(1 - I_i\right) I_i^m \left\{ -\mathbb{1}\left(y_i^m = z\right)\left(\frac{\partial \Pi_z^m}{\partial \Pi_{zx^0}}\right)\frac{1}{\Pi_z^m} + \frac{\frac{\partial\left(\sum_{v\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)F\left(v,y_i^m,x_i;\theta\right)\right)}{\partial \Pi_{zx^0}}}{\sum_{v\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)F\left(v,y_i^m,x_i;\theta\right)} \right\} +$$

$$I_i\left(1 - I_i^m\right)\left\{ -\left(\frac{\partial \Pi_{y_i}}{\partial \Pi_{zx^0}}\right)\frac{1}{\Pi_{y_i}} + \frac{\frac{\partial\left(\sum_{w\in\{0,1\}}\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(y_i,w,x_i;\theta\right)\Pi_{wx_i}\right)}{\partial \Pi_{zx^0}}}{\sum_{w\in\{0,1\}}\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(y_i,w,x_i;\theta\right)\Pi_{wx_i}} \right\} +$$

$$\left(1 - I_i\right)\left(1 - I_i^m\right)\frac{\frac{\partial\left(\sum\sum_{v,w\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(v,w,x_i;\theta\right)\Pi_{wx_i}\right)}{\partial \Pi_{zx^0}}}{\sum\sum_{v,w\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(v,w,x_i;\theta\right)\Pi_{wx_i}}$$

where:

$$\left(\frac{\partial \Pi_{y_i}}{\partial \Pi_{zx^0}}\right) = F\left(y_i, z, x^0; \theta\right)$$

$$\left(\frac{\partial \Pi_{y^m}^m}{\partial \Pi_{zx^0}}\right) = \begin{cases} 1 \text{ if } z = y^m \\ 0 \text{ if } z \neq y^m \end{cases}$$

$$\frac{\partial\left(\sum_{v\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)F\left(v,y_i^m,x_i;\theta\right)\right)}{\partial \Pi_{zx^0}} = \sum_{v\in\{0,1\}}\frac{H_v}{\left(\Pi_v\right)^2}\left(\frac{\partial \Pi_v}{\partial \Pi_{zx^0}}\right)F\left(v,y_i^m,x_i;\theta\right) =$$

$$\sum_{v\in\{0,1\}}\frac{H_v}{\left(\Pi_v\right)^2}F\left(v,z,x^0;\theta\right)F\left(v,y_i^m,x_i;\theta\right)$$

$$\frac{\partial\left(\sum_{w\in\{0,1\}}\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(y_i,w,x_i;\theta\right)\Pi_{wx_i}\right)}{\partial \Pi_{zx^0}} = \frac{H_z^m}{\left(\Pi_z^m\right)^2}\underbrace{\left(\frac{\partial \Pi_z^m}{\partial \Pi_{zx^0}}\right)}_{=1}F\left(y_i,z,x_i;\theta\right)\Pi_{zx_i} +$$

$$\mathbb{1}\left(x_i = x^0\right)\left(1 - \frac{H_z^m}{\Pi_z^m}\right)F\left(y_i,z,x^0;\theta\right)$$

$$\frac{\partial\left(\sum\sum_{v,w\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(v,w,x_i;\theta\right)\Pi_{wx_i}\right)}{\partial \Pi_{zx^0}} =$$

$$\sum\sum_{v,w\in\{0,1\}}\frac{H_v}{\left(\Pi_v\right)^2}\left(\frac{\partial \Pi_v}{\partial \Pi_{zx^0}}\right)\left(1-\frac{H_w^m}{\Pi_w^m}\right)F\left(v,w,x_i;\theta\right)\Pi_{wx_i} +$$

$$\sum_{v,\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)\frac{H_z^m}{\left(\Pi_z^m\right)^2}\underbrace{\left(\frac{\partial \Pi_z^m}{\partial \Pi_{zx^0}}\right)}_{=1}F\left(v,z,x_i;\theta\right)\Pi_{zx_i} +$$

$$\sum_{v,\in\{0,1\}}\left(1-\frac{H_v}{\Pi_v}\right)\left(1-\frac{H_z^m}{\Pi_z^m}\right)\mathbb{1}\left(x_i = x^0\right)F\left(v,z,x^0;\theta\right)$$

References

[1] Amorim, Maria Norberta (1991), Uma metodologia de Reconstituiçao de Paróquias, Universidade do Minho.

[2] Amorim, Maria Norberta (2009): "Marriage Strategies in Azorean Communities of Pico Island (19th Century), Differentiated Female Behavior in Choosing a Lifelong Partner," in The Transmission of Well-Being: Gendered Marriage Strategies and Inheritance Systems in Europe (17th-20th Centuries), Margarida Duraes, Antoinette Fauve-chamoux, Llorenc Ferrer, Jan Kok (Eds). Peter Lang Pub. Inc.

[3] Andersen, E.B. (1973), Conditional Inference and Models for Measuring, Copenhagen, Mentalhygiejnisk Forlag.

[4] Anker, R. (1998): Gender and Jobs. Sex Segregation of Occupations in the World, International Labor Office, Geneva.

[5] Anker, R. and C. Hein (1985): "Why Third World Urban Employers Usually Prefer Men," International Labour Review, 104, 85:105.

[6] Arellano, M. and R. Carrasco (2003): "Binary choice panel data models with predetermined variables," Journal of Econometrics 115 (1), pages 1-197.

[7] Arellano, M. and J. Hahn (2007): "Understanding Bias in Nonlinear Panel Models: Some Recent Developments." Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress, Volume 3, edited by R. Blundell, W. Newey, and T. Persson. Cambridge University Press.

[8] Arellano, M. and B. Honoré (2001): "Panel Data Models: Some Recent Developments," in Heckman, J.J. and E. Leamer (eds) Handbook of Econometrics, V, Ch. 53, Elsevier Science, Amsterdam.

[9] Baptista, Virgínia do Rosário (1999): "As mulheres no Mercado de Trabalho em Portugal: Representaçoes e Quotidianos (1890-1940)," Orfanizaçoes nao Governa-

mentais do Conselho Consultivo da Comissão para a Igualdade e para os Direitos das Mulheres. Lisboa.

[10] Becker, G. (1971): The Economies of Discrimination. Chicago: Chicago University Press, 2nd ed.

[11] Bergmann, B.R. (1974): "Occupational Segregation, Wages and Profits When Employers Discriminate By Race Or Sex," Eastern Economic Journal, 1: 103-110.

[12] Beller, A. (1985): "Changes in the Sex Composition of U.S. Occupations, 1960-1981," Journal of Human Resources 20, 235-250.

[13] Blau, F. (1977): "Equal Pay in the Office," Lexington, Heath.

[14] Blau, F. and Hendricks, W. (1979): "Occupational Segregation by Sex: Trends and Prospects," Journal of Human Resources 12 (1979), 197-210.

[15] Blau, F. (1984): "Occupational Segregation and Labor Market Discrimination," in Sex segregation in the Work Place, ed. by B. Reskin, National Academy Press.

[16] Blau, F., Simson, P. and Anderson, D. (1998): "Continuing Progress? Trends in Occupational Segregation Over the 1970s and 1980s," Feminist Economics 4, 29-71.

[17] Black S.E., and C. Juhn (2000). "The Rise of Female Professionals: Are Women Responding to Skill Demand?", American Economic Review, 90 (2), pp. 450-455.

[18] Borghans, L., and Groot, L. (1999): "Educational Presorting and Occupational Segregation", Labour Economics 6, 375-395.

[19] Brettell, Caroline (1991b): "Kinship and Contract: Property Transmission and Family Relations in the Northwestern Portugal," Comparative Studies in Society and History, 33 (3), pp 443-465.

[20] Browning, M. and J. M: Carro (2007), "Heterogeneity and microeconometrics modeling", joint with Martin Browning

[21] Browning, M. and J. M: Carro (2007), "Heterogeneity and microeconometrics modeling"; in Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress, Volume 3, edited by R. Blundell, W. Newey, and T. Persson. Cambridge University Press.

[22] Browning, M. and J. M: Carro (2009), "Dynamic binary outcome models with maximal heterogeneity". *mimeo.* Departamento de Economia, Universidad Carlos III de Madrid.

[23] Bütikofer, Aline (2010): "Revisiting 'Mothers and Sons': Preference Formation and Female Labor Force in Switzerland." Mimeo.

[24] Carrilho, Maria José (1996): "População Activa: Conceito e Extensão Através dos Censos," Instituto Nacional de Estatística, Lisbon, Portugal.

[25] Chamberlain, G. (1980): "Analysis of Covariance with Qualitative Data," Review of Economic Studies, 47(1), 225-38.

[26] Chamberlain,G.(1985): "Heterogeneity,Omitted Variable Bias,and Duration Dependence," in Longitudinal Analysis of Labor Market Data ,ed. J.J.Heckman and B.Singer.Cambridge:Cambridge University Press, 3:38.

[27] Costa, D.L. (2000), From Mill Town to Board Room: The Rise of Women's Paid Labor, Journal of Economic Perspectives, 14, 101-122.

[28] Costa, P., A. Terracciano, and R.R. McCrae, (2001): "Gender Differences in Personality Traits Across Cultures: Robust and Surprising Findings," Journal of Personality and Social Psychology. Vol 81(2), 322-331.

[29] Del Boca, D. and M. Locatelli (2006): "The Determinants of Motherhood and Work Status: A Survey," Discussion Paper No. 2414, IZA.

[30] Doeringer, P.B. and M.J. Piore (1971): Internal Labor Markets and Manpower Analysis. Massachusetts. D.C. Heath and Company.

[31] Durães, Margarida (2009): "Providing Well-Being to Women through Inheritance and Succession, Portugal in the 18th and 19th Centuries," in The Transmission of Well-Being: Gendered Marriage Strategies and Inheritance Systems in Europe (17th-20th Centuries), Margarida Duraes, Antoinette Fauve-chamoux, Llorenc Ferrer, Jan Kok (Eds). Peter Lang Pub. Inc.

[32] Fernández, R. (2011), "Does Culture Matter?," in Handbook of Social Economics, vol 1A, Jess Benhabib, Alberto Bisin, Matt Jackson, eds., Elsevier, 2010r.

[33] Fernández, R. (2007), "Women, Work and Culture," Journal of the European Economic Association, 5 (2-3), 305-332 .

[34] Fernández, R., and A. Fogli (2009), "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," American Economic Journal: Macroeconomics, 1(1), 146-177.

[35] Fernández, R., A. Fogli, and C. Olivetti (2004), "Mothers and Sons: Preference Formation and Female Labor Force Dynamics," Quarterly Journal of Economics 119(4), 1249-1299.

[36] Galor, O. and D.N. Weil (1996): "The Gender Gap, Fertility, and Growth," American Economic Review, 86 (3), 374:387.

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

[38] Goldin, C. and L. Katz (2002): "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," Journal of Political Economy 110(4), 730-770.

[39] Greenwood, J., Seshadri, A. and M. Yorukoglu (2005): "Engines of Liberation," Review of Economic Studies, 72, 109:133.

[40] Gross, E. (1968): "Plus ça change...? The Sexual Structure of Occupations Over Time," Social Problems 16, 198-208.

[41] Hakim, C., (1996): Key Issues in Women's Work , University Press, Cambridge.

[42] Honore, B.E., and E.Kyriazidou (2000): "Panel Data Discrete Choice Models with Lagged Dependent Variables," Econometrica 68, 839:874.

[43] Jacobsen, J., The Economics of Gender (1998), second edition, Blackwell, Oxford.

[44] Matos, Paulo Lopes (2009): "Female Life Courses and Property Transmission in the Azorean Periphery (Portugal): The Case of the Island of São Jorge in the 19th Century," in The Transmission of Well-Being: Gendered Marriage Strategies and Inheritance Systems in Europe (17th-20th Centuries), Margarida Duraes, Antoinette Fauve-chamoux, Llorenc Ferrer, Jan Kok (Eds). Peter Lang Pub. Inc.

[45] Moreira da Silva, Rosa Fernanda, (1983): "Contraste e Mutaçoes na Paisagem Agrária das Planícies e Colinas Minhotas," Estudios Contemporâneos, Ministério da Cultura, Delegaçao do Norte, Centro de Estudos Humanísticos. Comunidades Rurais, Estudos Interdisciplinares. N^o 5, Porto.

[46] Miller, P. (1987). "The Wage Effect of Occupational Segregation of Women in Britain," Economic Journal, 97, 885-896.

[47] Pina-Cabral, João de (1986): "Sons of Adam, Daughters of Eve: The Peasant World View of the Alto Minho." Clarendon Press, Oxford.

[48] Ramalho, Esmeralda, and Smith, Richard (2012): "Discrete Choice Nonresponse", Review of Economic Studies, April, 6, 2012, doi:10.1093/restud/rds018.

[49] Reis, Jaime (2005): "O Trabalho" in Historia Económica de Portugal 1700-2000, vol II, Pedro Lains e A. F. Silva (Eds).

[50] Roos, P. A., Gender and Work: A Comparative Analysis of Industrial Societies (1985), State University of New York Press, New York.

[51] Scott, A. S. Volpi (1999), "Famílias, formas de união e reprodução social no noroeste português (séculos XVIII e XIX)". Guimarães : NEPS.

[52] Tsukahara, I. (2007), "The Effect of Family Background on Occupational Choice," LABOUR, vol. 21(4-5), pages 871-890, December.