Transmission of preferences and beliefs about female labor market participation: direct evidence on the role of mothers[*]

Jesús M. Carro[†], Matilde P. Machado[‡], Ricardo Mora[§]

Abstract

Recently, economists have established that culture—defined as a common set of preferences and beliefs—affects economic outcomes, including the levels of female labor force participation. Although this literature has argued that culture is transmitted from parents to children, it has also recognized the difficulty in empirically disentangling the parental transmission of preferences and/or beliefs from other confounding factors, such as technological change or investment in education. Using church registry data from the 18th and 19th centuries, our primary contribution is to interpret the effect of a mother's labor participation status on that of her daughter as the mother-to-daughter transmission of preferences and/or beliefs that are isolated from confounding effects. Because our data are characterized by abundant non-ignorable missing information, we estimate the participation model and the missing process jointly by maximum likelihood. Our results reveal that the mother's working status has a large and statistically significant positive effect on the daughter's probability of working. These findings suggest that intergenerational family transmission of preferences and/or beliefs played a decisive role in the substantial increases in female labor force participation that occurred later.

Keywords: Female labor market participation, intergenerational transmission of preferences and/orbeliefs, historical family data, church registry data, non-ignorable missingness, econometric methods for missing data.

JEL Classification: J22, J24, J16, J12

†   jcarro@eco.uc3m.es. Universidad Carlos III de Madrid
‡   matilde.machado@uc3m.es. Universidad Carlos III de Madrid and CEPR
§   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

1 Introduction

Increased female labor force participation, —particularly that segment consisting of married women—is one of the most important labor market transformations that has occurred over the last century. The contribution of technological change to such an increase is now generally acknowledged (e.g., Galor and Weil, 1996, Goldin and Katz, 2002, Greenwood, Seshadri, and Yorukoglu, 2005, and Goldin, 2006). Over the past decade, however, new theories have highlighted the important role of changes in preferences and/or beliefs in transforming the labor market. Within this literature, certain studies establish the effects of preferences and beliefs on economic outcomes (e.g., Guiso, Sapienza, and Zingales, 2006, Fernández, 2007, Tabellini, 2008, Fernández and Fogli, 2009, Alesina and Giuliano, 2010, Fernández, 2011, Alesina and Giuliano, Forthcoming), whereas others indicate that the intergenerational transmission of preferences and beliefs may have a potential multiplier effect on the evolution of attitudes and gender roles (e.g., Fernández, Fogli, and Olivetti, 2004, Morrill and Morrill, 2013, Alesina, Giuliano, and Nunn, 2013, Olivetti, Patacchini, and Zenou, 2013, and theoretical approaches by Bisin and Verdier, 2001 and Fernández, 2013). However, empirical studies reporting a relationship between parents' labor market behavior and attitudes and the behavior and attitudes of their children (e.g., Tsukahara, 2007, Farré and Vella, 2013, Bütikofer, 2013) cannot disentangle the transmission of preferences/beliefs from other family-specific effects, such as the transmission of or investment in human capital, which were also likely to occur in a world experiencing rapid technological changes.[1] Our paper belongs to this group of empirical studies but employs a novel approach to provide direct evidence of the intergenerational transmission of preferences and/or beliefs regarding labor market participation as transmitted through the mother-to-daughter channel.

Our approach investigates a time and place in which a mother's participation in the labor force has effects on her daughter's participation that can be separately identified

---

[1]It is possible that estimated intergenerational effects incorporate what is known in the literature as the "social multiplier" effect, i.e., one's neighbors peer effects. Maurin and Moschion (2009), for example, demonstrate that female neighbors' labor market participation has a substantial effect on women's participation in the labor market. More recently, Olivetti, Patacchini, and Zenou (2013) separately identify intergenerational (mother-to-daughter) and social network (friends' mothers) effects on hours worked.

from the effects of access to education or technological and cultural changes. In conducting this investigation, we employ historical parish registry data from four Portuguese locations in the 18th and 19th centuries. Importantly, technological and cultural changes occurred in Portugal only at the beginning of the 20th century (e.g., Lains, 2006 and the references therein). Although Portugal lagged behind the most developed countries in terms of technological innovation and cultural changes in the 19th century, it experienced changes in female labor force participation in the 20th century that closely followed those in the United States.[2]

Notably, our data include information on occupations and social status. We use this information to construct individual measures of labor force participation of the females in our data (the dependent variable) and that of their mothers (the independent variable of interest). As with most historical individual data, the information on occupations contains a substantial proportion of missing values, which affects both our dependent variable and our independent variable of interest. When missing observations are not random, statistical analysis using only observations with non-missing values suffers from sample selection bias. We consider two approaches to address this non-ignorability problem. First, informed by historical records and narratives, we conservatively impute missing values with the perceived predominant female labor market status at the time, i.e., non-participation, which sets the participation rate at an unrealistically low value relative to census data. Second, instead of imputing missing values, we apply a methodology based on that developed in Ramalho and Smith (2013), which allows models to be estimated in contexts in which missing data are abundant and non-random. This methodology involves Maximum Likelihood (ML) estimation using all observations—including those with missing values. Applying the Ramalho and Smith (2013) methodology offers considerable potential to use historical data that would otherwise remain unexplored.

---

[2]Census data from the beginning of the 20th century demonstrate that approximately 37% of women 15-65 years of age were employed (from the Portuguese "tem profissao", Portuguese census, 1911); by 1975, that figure increased to approximately 50%, and by the end of the century, the female participation rate had reached 67% and the male-female gap declined to 13.1% (INE, 1977, and Eurostat). In the United States, for example, female labor force participation was approximately 18.9% in 1890, 29.5% in 1950 and 60% in 1999, with a male-female participation gap of 12.9% by the end of the century (Goldin, 1990 and the Bureau of Labor Statistics).

We improve the identification of the parameters of the ML model by incorporating external information regarding aggregate female labor market participation from other historical sources into the likelihood function. Because these external data sources do not precisely match our locations and period, we present results under three scenarios: The *baseline*, which is based on the best external data proxy, and the *high* and *low* scenarios. Using the ML approach under the baseline scenario, we estimate a positive and statistically significant marginal effect of 32 percentage points of the mother's participation status on the daughter's probability of participation. This marginal effect is large in magnitude considering that the average value of female participation in our sample is only 14%. The marginal effect of the mother's participation status under the high and low scenarios—which we regard as extreme—is 47 and 22 percentage points, respectively. Estimating the model with the ML approach allows us to test the ignorability of the missing process. We reject ignorability. This implies that data with missing values should not be discarded. Our marginal effects estimates are considerably higher than those estimated with contemporaneous data for which other factors unrelated to the family—such as technological change—are major determinants of high female labor force participation; for example, Morrill and Morrill (2013) estimate a mother-to-daughter effect of approximately 7 percentage points.

We conduct a number of robustness checks. First, we re-estimate the model by restricting the data to the periods and locations for which the percentage of missing values is lower, and our results remain unchanged, which confirms that the ML procedure accords greater weight to those observations without missing values. Second, we address the potential issue of female migration and how it may bias our results by re-estimating the model using samples that differ with respect to the prevalence of migrants. Third, to ensure that our results capture only the transmission of preferences and/or beliefs, we discard the possibility that they reflect only those effects related to the transmission of property or social status. This would be the case, for example, if the economic activities of females were similar to those of their mothers as a result of property being transmitted mainly to daughters. In our baseline estimations, we include controls for influential fathers

and for mothers classified by the vicar as property owners or holders of titles of nobility. Moreover, we also re-estimate the model in samples in which property-owning mothers were discarded. Similar procedures were followed to dismiss the possibility that poverty traps transmitted from mother to daughter explain our results. Finally, we examine the exact occupations and professions that were reported by the vicars to assess the extent to which our estimated effects simply reflect the transmission of a particular skill or craft from mother to daughter. The low number of working women with the same profession as their mothers makes this possibility highly unlikely.

The remainder of this paper proceeds as follows. Sections 2 and 3 describe the historical background and the data set, respectively. Section 4 describes the econometric model and estimation methods. Section 5 presents our main estimation results, and Section 6 discusses the potential relevance of mechanisms beyond preferences and beliefs and evaluates the sensitivity of the main results to different samples. Section 7 concludes. Appendices A and B describe additional features of the data.

2   Historical Background

The four Portuguese locations from which our data are obtained include São Tiago de Ronfe (hereafter Ronfe), Ruivães, Horta, and São Mateus (hereafter S. Mateus). The villages of Ronfe and Ruivães are only 9 km apart and strategically located between the two historical administrative centers, Guimarães and Braga, in the Minho region in the northeast of Portugal. The coastal city of Horta at Fayal Island—a major stopping port on the journey to Brazil—and the rural village of S. Mateus at Pico Island are located in the Azores Archipelago.

The period studied pre-dates any technological change that occurred in Portugal (e.g., Lains, 2006). Additionally, the legal and social background of Portuguese society during the sample period does not favor the economic independence of women. The most relevant changes in the Portuguese legal system regarding women's rights occurred only after the proclamation of the First Republic in 1910.[3] Women were not only legally discriminated

---

[3]Regarding the evolution of women's legal rights in Portugal, see, for example, Solsten (1993).

against but also excluded from the main educational system.[4] They were, however, allowed to own and inherit property.

During our sample period, there were three succession systems. The first was a male primogeniture system referred to as *Morgadio* through which the oldest son inherited the land and the name (and title) of the property owner. The *Morgadio* was practiced only among the wealthiest families of landlords and aristocrats from the 13th century until it was abolished in 1863 (Moreira da Silva, 1983). The second norm, far more common than the *Morgadio*, applied to life-long rentals of aristocratic or ecclesiastic land. Life-long rentals had to be transmitted to a single heir and tended to favor spouses over children, male over female children, and older over younger children. In contrast to the *Morgadio*, daughters could inherit life-long rentals, as frequently occurred in the Minho region (Durães, 2009).

The third norm and general rule for property transmission was to divide two-thirds of the property (the *legitima*) equally among the legitimate heirs and to dispose of one-third (the *terço*) to benefit one of the children or the surviving spouse. Scholars describing the local customs report that the *terço*, which typically included the main house (or part of it) and the adjacent land, either became the property of the first marrying child (in which case daughters were more likely to receive it) or was bequested to a spouse or to unmarried children—who were frequently daughters (Brettell, 1991, Durães, 2009, Pina-Cabral, 1986, Matos, 2009).[5]

A feature of the four locations is the predominantly male emigration to Brazil beginning in the 16th century. As a consequence, the Minho region—to which Ronfe and

---

[4]According to the population census of 1864 , among boys between 6 and 15 years of age, the share attending primary educational institutions in Horta and Braga—the regions (*Distritos*) of our locations—is 13.4% and 18.0%, respectively. By contrast, the corresponding shares for girls are only 5.0% and 1.3%. Although some young women began to receive higher education late in the 19th century, a law was passed to allow for the creation of all-girls public schools for secondary education only in 1888.

[5]Daughters might be favored by the *terço* for several reasons. First, having land—or the promise of it—increased a woman's chances in the marriage market because landless women seldom married. The reason for the thin marriage market was the heavy male emigration to Brazil, which we describe in the next paragraph. Second, because married daughters 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 daughters-in-law. Third, it was also common for single daughters to inherit the *terço*, which, on the one hand, would guarantee them the means of survival and, on the other hand, would also guarantee that the parents would be cared for in their old age.

Ruivães belong—and the Azores were atypical in Portugal in terms of the population's gender composition, with women substantially outnumbering men. According to the 1864 Census, the male to female ratio in the city of Horta, 75 men to 100 women, was the third lowest among the 32 largest Portuguese cities, and in Minho's Braga District, there were only 81 men for every 100 women.[6] A similar pattern has also been documented for the location of S. Mateus (Amorim and Santos, 2009). Researchers report, however, that there are differences in the sex ratio over time (see, e.g. Scott, 1999 for the evidence on Ronfe) and differences in emigration across regions and time (e.g. Reis, 2005).

3 Data

3.1 Parish data

The main data source is parish information that dates back to the end of the 16th century and was extracted from parish records in the villages of Ronfe, Ruivães, Horta, and S. Mateus by a research team led by M. Norberta Amorim at the Núcleo de Estudos de População e Sociedade (NEPS), which is a research institute associated with the Universidade do Minho.[7]NEPS staff collected the main data sets using all baptism, marriage, and death certificates found in the local churches and matched this family dataset with other church records known as *rol de confessados* (literally, "the list of the confessed").[8] Importantly, these lists contain information recorded by the local vicar regarding individual-specific occupations and/or social status.

The original baptism, marriage, and death certificate records allowed family linkages to be reconstructed within each location beginning in the 1550s through the 20th century (Amorim, 1991). Altogether, after some basic cleaning, the data set has entries for 92, 474 individuals. Individual records include information on birth, marriage, and death dates, gender, parents' identification codes, spousal identification codes, and children's identifi-

---

[6]Scott (1999) reports a male to female ratio of 0.64 in Ronfe in 1740.

[7]NEPS no longer exists. Currently, the Grupo História das Populações (Universidade do Minho) in the Centro de Investigação Transdisciplinar Cultura, Espaço e Memória, administers the genealogical database.

[8]These lists are reliable because they were used to monitor the administration of the sacrament of penance to the parishioners during Lent.

cation codes. Gender was originally inferred from the individual's first name in the Parish registry. Overall, 46,094 individuals were recorded as female (49.9%).

We have the year of birth for 60.9% of all records. By contrast, we have death information for only 33.2% 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 complete such records, we group all individuals for which the information is available into cohorts spanning 25 years. Observations for which the year of birth is missing are completed by sequentially examining the 25-year birth period of siblings, spouse, and children, in that order. When the cohort of the siblings or the spouse is identified, the record is completed with the 25-year birth period of the spouse or sibling. In the event that only cohorts of the children are identified, the 25-year period previous to the cohort of the eldest child is assigned to the missing record. This procedure is repeated until no changes are produced. As a result, 84.1% of the original data can be associated with a given 25-year period.

Our observations include individuals who do not survive childhood and individuals who migrate to other locations and for whom no further information is available. In theory, the former group of individuals would be identified using the death date information, whereas the latter group could be indirectly inferred by the absence of information on their deaths. However, the profusion of missing death dates hinders the precise identification of both early death and migration.

The data set originally includes 887 slaves, of whom 557 are women, and all but one are located in the Azores. Because our aim is to model the effect of a mother's labor force participation on that of her daughter, we drop all slaves and daughters of slaves from the sample. Additionally, we restrict our analysis to the 18th and 19th centuries because the small number of observations from the 16th and 17th centuries suggests that the parish registry is not complete for these periods. To minimize the number of observations for which the occupation was registered in the 20th century, we do not include individuals who were born in the last quarter of the 19th century. For purposes of consistency, we include those born in the last quarter of the 17th century. Thus, we denote as 18th-century

observations all observations with birth years in the 1675-1774 period and as 19th-century observations all those with birth years in the 1775-1874 period.

We are left with 24, 381 female observations. Panel A in Table 1 reports the number of observations by location and century of the resulting sample.

Table 1: NUMBER OF OBSERVATIONS BY LOCATION AND CENTURY

Panel A: Full Sample					
	Horta	Ronfe	Ruivães	S. Mateus	Total
18th century	5,310	1,357	799	2,731	10,197
19th century	6,327	1,741	1,151	4,965	14,184
Total	11,637	3,098	1,950	7,696	24,381
Panel B: Sample with Mother Identified					
	Horta	Ronfe	Ruivães	S. Mateus	Total
18th century	3,107	970	595	2,186	6,858
19th century	4,581	1,223	830	4,177	10,811
Total	7,688	2,193	1,425	6,363	17,669
Loss of obs. (%)	-33.9	-29.2	-26.9	-17.3	-27.5

Note: Observations are assigned to century by adding 25 years to the quarter-century of their birth year to better proxy for the period of their adult life. Hence 18th century observations include all women born between 1675-1774; 19th century observations include all women born between 1775-1874. Panel A above is constructed based on the sample with all women born between 1675-1874 except slaves and daughters of slaves. Panel B, restricts further the sample to women whose mothers are identified in the dataset. The last row labelled "Loss obs. (%)" shows the percentage loss in observations by location incurred due to the selection of women with identified mother.

Parish data containing birth, marriage, and death certificates were matched by NEPS with data from the church Lent census, the "*rol de confessados*" discussed above, which is a parochial census organized by the households of all residents older than seven years of age and produced by the parochial vicar during Lent to administer the sacrament of penance to the parishioners. The Lent census includes information on the economic activity and/or social status of adult individuals in the household.

3.2 Missing data and classification problems

The economic activity/social status information collected in the Lent census is not as complete as the baptism, marriage, and death registries. We proceed to describe and discuss the main problems related to both the Lent census data and the main parish data.

3.2.1 Non systematic classification of professions

Professions were not systematically classified across parochial vicars and across time. As a result, the original data include more than 500 categories, many of which are close substitutes for one another.[9] To make this information tractable, we construct a variable on economic activity using four major categories: *employee/farmer, professional/capital owner, domestic production*, and *unproductive*.[10] *Employee/farmer* includes all paid and unqualified jobs. *Professional/capital owner* includes landlords, liberal professions, traders, businesswomen, the self-employed and qualified and managerial jobs. *Domestic production* includes all women classified as housewives—from the Portuguese *domestica* recorded in 298 observations—and women to whom the vicar accorded the title of *Dona*, a term originally used to signal an upper class that was gradually adopted to also signal the bourgeoisie during the 18th and 19th centuries (933 observations). . The *unproductive* category includes the indigent (65 observations), individuals accorded a title of nobility by the vicar (3 observations), and others (14 observations). Based on these four major activity categories, we define labor market participation as being an employee/farmer or a professional/capital owner. Whereas this definition correctly includes as labor force participants all small capital owners who are self-employed, such as shop owners, it also includes large capital owners who live off of rents. Unfortunately, the vicar's description of the individual's occupation or social status does not allow us to accurately differentiate between these two types of capital owners. Thus, we will refer to this notion derived from the vicar's report on economic activity or social status as the individual's participation

---

[9]Table A1 in Appendix A reports the most common disaggregated female professions/social status reported by the vicars for females during the $1675 - 1874$ period.

[10]Five women from Horta have a second profession recorded in the registry. We conservatively adopted as valid only their first profession, which in all five records was "housewife".

status.

3.2.2 Lack of mother's identification

The second problem is that the mother is not identified for 6,712 women, which represents 27.5% of our sample (see Panel B in Table 1). Disregarding orphans and illegitimate children (annotations from the NEPS team allow us to identify only 1.74% of all cases with an unidentified mother as orphans/illegitimate), mothers may not be identified for two reasons. Identification fails if the mother's record precedes the first currently available registries, is illegible, or contains coding errors that prevent a match with the daughter's record. The older the records are, the more likely these occurrences are: 54.40% of the observations from the first quarter of the 18th century have an unidentified mother compared with only 27.5% in the last quarter of the 19th century. These registry issues are not systematically related to participation decisions or to the vicar's recording practices and are thus not a likely source of bias in our estimates.

Another reason for failing to identify the mother is that mothers who are not born, married, or deceased in the parish leave no personal records in the parish and, thus, cannot be found in our dataset. This situation arises, for example, when a woman migrates to one of our four locations. Migrations between villages were typically related to marriage. However, movements to cities, such as Horta, are also likely related to labor force participation.

In our benchmark estimation, we exclude observations for which the mother is not identified. Because we cannot rule out sample selection due to migratory flows, we perform two additional estimations as robustness checks, one including women whose mothers are not identified and another for Horta in which we include only women who were born in the city.

3.2.3 Non-ignorable missing occupational/social status

The third and, given its magnitude, more important problem is highlighted in Panel A in Table 2. Even in the location with the highest coverage, Horta, the proportion of ob-

11

servations with information on occupational/social status is small: 15.32%. Accounting for migrations and early deaths would increase the coverage for the population living in Horta, but it would not explain the nonresponse rate. The coverage is particularly poor in S. Mateus, for which only 0.09% of women have their participation status recorded. Participation status coverage also differs by century across locations. For example, whereas the coverage of participation status is 5.73% in Horta in the 18th century, it increases to 21.83% in the 19th century. The average coverage is 6.90%.

It would be tempting to interpret the low coverage of women's participation as connected with gender bias in the recording practices of the vicars. However, the coverage for males is only marginally higher (17.93% for Horta, 5.87% for Ronfe, 0.70% for Ruivães and 4.7% for S. Mateus). Thus, it appears that vicars are not less inclined to report women's economic activities.

Why would the vicar report some women's activity/social status and not that of others? One plausible explanation is that they reported this status to differentiate between common names. If there were an excessive number of "Marias", the women's economic activity or social status would help distinguish them. Frequency tables of participation status by given name, however, indicate that such is not the case, i.e., that the most common names are as likely among the reported as among the unreported. For example, among the reported, 30.2% were named "Maria", 9.0% "Ana", and 5% "Francisca" compared with 30.5%, 9.3% and 4.1%, respectively, among the unreported.[11]

Panel B of Table 2 shows that daughters of women whose participation status is not reported by the vicar have a 96.3% chance of not being reported. By contrast, daughters of women whose participation is reported have a 45.6% chance of having their participation also reported (and only 13.6% if the father's participation is reported). Underreporting among women whose mothers' participation is not reported suggests that missingness might be associated with dependence between mother's and daughter's status and potentially non-ignorable, as a consequence.

---

[11]The given names Maria and Ana were typically followed by a second given name.

Table 2: Information on Reported Participation Status

Panel A: Proportion (%) of Sample with Nonmissing Participation Status

	Horta	Ronfe	Ruivães	S. Mateus	Total
18th century	5.73	0.72	0.00	0.09	2.73
19th century	21.83	1.39	1.33	0.10	9.55
Total	15.32	1.09	0.77	0.09	6.90

Panel B: Proportion (%) of Daughters with Nonmissing Participation
By Mother's Missingness of Participation

	Mother's participation is:	
	Not Recorded	Recorded
Daughter's participation is Not recorded	96.3	54.4
Daughter's participation is recorded	3.7	45.6
Total	100	100

Note: Panels A and B above are constructed based on the sample of 17,669 women with identified mothers in the dataset, who were born between 1675-1874. Slaves and daughters of slaves are dropped from the sample. Observations are assigned to century by adding 25 years to the quarter-century of their birth year to better proxy for the period of their adult life. Hence 18th century observations include all women born between 1675-1774; 19th century observations include all women born between 1775-1874. Women have their participation recorded whenever the local vicar enters a description of her activity or social status in the Lent census (from the Portuguese "Rol de confessados") and that registry is matched by NEPS with the baptism, death or marriage certificates.

In Table 3, we report the distribution of women across our four major activity categories by location. Distributions vary significantly across locations and may be the compound of location-specific economic factors and differentials in the incidence of missing information. In some of the locations, certain activities appear underrepresented (e.g., domestic production in Ronfe, Ruivães and S. Mateus), whereas others are surprisingly overrepresented (e.g., unproductives in S. Mateus). The share of employees/farmers is also unrealistically low for rural communities, particularly in S. Mateus, possibly because certain parochial vicars did not collect professional information on isolated farmers, as the Lent census was organized by household and gathered by the vicar door-to-door. More generally, it appears that vicars tended to record activity only for those whose labor or

social status was uncommon in the region and period, such as for civil servants or for the miller in a village of farmers. If such is the case, missingness is non-ignorable.

Table 3: DISTRIBUTION (%) OF WOMEN ACROSS FOUR MAJOR ECONOMIC CATEGORIES
Observations with Recorded Economic Activity or Social Status

	Horta	Ronfe	Ruivães	S. Mateus	Total
Employee/Farmer	6.5	16.7	45.5	16.7	7.0
Professional/Capital owner	10.9	79.2	54.6	33.3	12.7
Domestic Production	82.6	0.0	0.0	0.0	79.8
Unproductive	0.08	4.2	0.0	50.0	0.4
Total	100	100	100	100	100
Participating (as % reported)	17.3	95.8	100	50	19.8

Note: The Table is based on the sample of 1,219 women with identified mothers in the dataset who were born between 1675-1874 and had their activity or social status reported by the vicar. Slaves and daughters of slaves are dropped from the sample. Women's economic activity/social status was constructed using information from two variables from the original NEPS files: "Profissao" and "Titulo". While the former reports professions, the latter reports social status for example a nobility title such as "countess" or an ecclesiastic position such as "Abbess". Because the original data had so many different descriptions of professions and occupations, we aggregated them into the 5 categories described in this Table. The last row shows the percentage of those with reported activity/social status who we consider active in the labor market which corresponds to the sum of Employees/Farmers and Professional/Capital Owner.

Disregarding observations with missing participation status may bias upward estimates of the effect of mother's labor force participation on the daughter's probability of participation. A larger proportion of working daughters among those whose mothers worked would be incorrectly interpreted as intergenerational transmission when it might simply have arisen because those working women whose mothers were housewives would be underrepresented in the sample. In other words, the high marginal effect of the working status of the mother (approximately 42 percentage points) reported in the transition matrix in Table 4 may be an artifact of the missing process.

Table 4: Female Participation by the Mother's Participation Status
Sample with recorded participation status

| | Mother: | | |
	Did not participate	Participated	Total
Daughter:			
Does not participate	549 (94.66%)	18 (52.94%)	567 (92.35%)
Participates	31 (5.34%)	16 (47.06%)	47 (7.65%)
Total	580 (100%)	34 (100%)	614 (100%)

Note: The Table is based on the 614 mother-daughter pairs with participation status recorded in the sample and born between 1675-1874. In this sample the mothers, as well as their participation status, are identified in the dataset. A woman is considered participating if we classify her as an Employee/Farmers or a Professional/Capital Owner.

3.3 External data on female labor market participation rates

As we explain in Section 4, identifying the parameters of interest under non-ignorable missingness is improved when the model is estimated with the help of external information sources. We use census data on local and national female force participation rates as external sources of information. The census data have at least two limitations. First, the first census with information on labor market participation was taken at the end of our sample period in 1890. Second, census data relies on a definition of labor participation that changed throughout the second half of the 19th century and the first half of the 20th century, which led to an artificial U-shaped trend in the estimated participation rates (see Humphries and Sarasúa, 2012, and the references therein). The main difference is found in the concepts of "occupation" and "profession". Although the former concept, adopted in the early period, classifies most women as "active" in the labor market, the latter concept, adopted in the later period, does not. The difference applies most notably to the case of women working in the household or the family farm, who would only be included in the labor force in early censuses. Thus, we should regard the rates such as that obtained from the 1890 Census as likely over-counting the number of women participating in the labor force. Reis (2005) provides a national-level estimate of the female labor market

participation in 1864 that is substantially lower (19.12%) than the rate obtained for 1890 using census data (38.47%) (for a detailed description of the difficulties in homogenizing the data on female labor force participation across censuses, see Appendix B).

Given the lack of precision in the census data, we adopt a conservative approach and consider three alternative scenarios regarding the aggregate level of female labor force participation in the four locations during the sample period. In the scenario with the largest female participation rates, which we refer to as the "high scenario", we take the largest local participation rate of the closest city or region for which we have information in the 1890, 1900, or 1911 censuses. In the "baseline scenario", we construct interpolations for labor force participation rates using a log functional specification and all data available: the 1864 data from Reis (2005), and data from all censuses up to 1991.[12] In our "low scenario", we decrease the participation rate from the "baseline scenario" by 50%. Table 5 reports the average values of female labor force participation rates by location used in the three scenarios.

Table 5: EXTERNAL VALUES FOR FEMALE LABOR FORCE PARTICIPATION
Alternative scenarios

	baseline	high	Δbaseline	low
Horta	0.0908	0.2170	+139%	0.0454
Ronfe	0.3511	0.5523	+57%	0.1755
Ruivães	0.4644	0.6216	+34%	0.2322
S. Mateus	0.0516	0.1616	+213%	0.0258

Note: The baseline scenario is an average over our sample period constructed using data from Recenseamentos Gerais da População 1864-1991 and adjusted for different census definitions of participation and age range of labor force (See Appendix B for details). The variation during the period is minimal, hence our baseline values for the beginning and end of the sample period only vary to the 4th decimal point. The high scenario takes the largest participation rate of the closest place for which we have information in the 1890, 1900, and 1911 censuses. Δbaseline shows the percentage deviation of the high scenario with respect to the baseline scenario. The low scenario is 50% of the participation rates in the baseline scenario.

---

[12]See Appendix B for further details.

4 The econometric model

4.1 A reduced-form participation model

We propose a reduced-form model in which a woman's participation in the labor force depends on her mother's participation. Woman i chooses either to participate in the labor market, $y_i = 1$, or not, $y_i = 0$. We assume that the discrete choice is expressed in the following linear specification:

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

where dummy variable y_i^m indicates the predetermined labor force participation status of a woman's mother. Vector x_i includes location and time dummies, a dummy for a large number of siblings, a dummy for whether the father was influential (i.e., a rentier, a large land or business owner, a merchant, a high-ranking civil servant, or an officer), and a dummy that signals whether the mother was a property owner.

Parameter α is the parameter of interest and captures the effect of the mother's participation status on that of her daughter. Vector β captures systematic differences in female participation across quarter-centuries and regions, in addition to capturing other family and individual characteristics. The error term ϵ_i incorporates the effects of those variables for which we have no information. For example, this term may include the effect of human capital of woman i that is unrelated to her parents' economic activity decisions. Typically, other studies assume that children's educational choices are correlated with their parents' professional choices. This assumption is not realistic here because women were excluded from the main educational system. Moreover, throughout the sample period, the Portuguese educational system remained extraordinarily elitist, with illiteracy rates over 80%, and formal education was only accessible to a small and privileged group (see, for example, http://www.country-data.com/).[13] Thus, it is reasonable to assume that human capital was primarily acquired informally.

---

[13]Educational reforms initiated in 1822, 1835, and 1844 were primarily targeted at boys' education and were left incomplete and largely unimplemented (see footnote 4 for literacy figures for boys and girls in 1864).

Quasi-permanent factors, such as social status or genetic predisposition, are an alternative source of persistence and, if relevant to this choice, would also affect y_i^m. In our empirical specifications, we control for two proxies of social status (i.e., whether the mother is a property owner and whether the father is influential). If other effects from social and genetic quasi-permanent factors are present and not controlled for, they might bias the estimate of α. In Section 6, we also explore whether our results are robust to controlling for the existence of an intergenerational poverty trap and to excluding mothers who are capital owners.

4.2   Likelihood with missing information

Most observations have missing entries in $\{y_i, y_i^m\}$ (93.10% for y_i and 96.29% for y_i^m). Let us define a binary indicator I_i, which takes value 1 if the participation status for observation i is observed and 0 otherwise. Similarly, let I_i^m take value 1 if the participation status for the mother of i is observed and 0 otherwise.

Missing completely at random arises when the probability of a missing observation is independent of y_i, in which case the missing observations are ignorable in the sense that their omission from estimation does not bias the results. As argued in Section 3, priests tended to record participation only for those whose labor or social status was not common in the region and period. Thus, missing participation is likely related to individual characteristics and, in particular, to individual participation status. In this situation, the missing mechanism is non-ignorable because estimating the participation model using only the non-missing observations will generally lead to inconsistent estimates.

The aim is to estimate parameter vector $\theta \equiv \{\alpha, \beta\}$ where :

$$\Pr\{y_i = v \,|\, y_i^m = w, x_i\} = F(v, w, x_i; \theta) \tag{2}$$

for $v, w \in \{0, 1\}$. Assuming normality we have the conditional probit model:

$$F(v, w, x_i; \theta) \equiv \begin{cases} \Phi(\alpha w + x_i \beta) & \text{if } v = 1 \\ 1 - \Phi(\alpha w + x_i \beta) & \text{otherwise} \end{cases}.$$

The missingness mechanism and the participation process jointly define the probability of an event:

$$
\begin{aligned}
\Pr\left\{I_i = r, I_i^m = s, y_i = v, y_i^m = w, x_i\right\} = \ & \Pr\left\{I_i = r, I_i^m = s \,|\, y_i = v, y_i^m = w, x_i\right\} \times \\
& F\left(v, w, x; \theta\right) \times \Pr\left\{y_i^m = w, x_i\right\}.
\end{aligned}
\tag{3}
$$

for $r, s \in \{0, 1\}$. For an observation with non-missing information, the joint probability of non-missingness, i.e., $I_i = I_i^m = 1$, and the vector variables $\{y_i, y_i^m, x_i\}$ is $\Pr\{I_i = I_i^m = 1, y_i = v, y_i^m = w, x_i\}$, which is a particular case of equation (3).

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 observation $\{I_i = 0, I_i^m = 1, y_i^m = w, x_i\}$ decomposes into two event probabilities that are, again, particular cases of equation (3):

$$
\begin{aligned}
\Pr\left\{I_i = 0, I_i^m = 1, y_i^m, x_i\right\} = \ & \Pr\left\{I_i = 0, I_i^m = 1, y_i = 1, y_i^m, x_i\right\} \\
& + \Pr\left\{I_i = 0, I_i^m = 1, y_i = 0, y_i^m, x_i\right\}
\end{aligned}
\tag{4}
$$

The treatment of the second case, i.e., when the mother's information is missing but the daughter's is not, is similar to that of the first case.

$$
\begin{aligned}
\Pr\left\{I_i = 1, I_i^m = 0, y_i, x_i\right\} = \ & \Pr\left\{I_i = 1, I_i^m = 0, y_i, y_i^m = 0, x_i\right\} \\
& + \Pr\left\{I_i = 0, I_i^m = 1, y_i, y_i^m = 1, x_i\right\}
\end{aligned}
\tag{5}
$$

In the third case, i.e., when information for both the mother and the daughter is missing, the joint probability for observation $\{I_i = 0, I_i^m = 0, x_i\}$ decomposes into four event probabilities:

$$
\begin{aligned}
\Pr\left\{I_i = 0, I_i^m = 0, x_i\right\} = \ & \Pr\left\{I_i = 0, I_i^m = 0, y_i = 1, y_i^m = 1, x_i\right\} \\
+ \Pr\left\{I_i = 0, I_i^m = 0, y_i = 0, y_i^m = 1, x_i\right\} & + \Pr\left\{I_i = 0, I_i^m = 0, y_i = 1, y_i^m = 0, x_i\right\} \\
+ \Pr\left\{I_i = 0, I_i^m = 0, y_i = 0, y_i^m = 0, x_i\right\}&.
\end{aligned}
\tag{6}
$$

Define p_i as the probability of observation i.

$$
\begin{aligned}
p_i = \ & \left(\Pr\left\{I_i = I_i^m = 1, y_i, y_i^m, x_i\right\}\right)^{I_i I_i^m} \times \\
& \left(\Pr\left\{I_i = 0, I_i^m = 1, y_i^m, x_i\right\}\right)^{(1-I_i) I_i^m} \times \\
& \left(\Pr\left\{I_i = 1, I_i^m = 0, y_i, x_i\right\}\right)^{I_i (1-I_i^m)} \times \\
& \left(\Pr\left\{I_i = I_i^m = 0, x_i\right\}\right)^{(1-I_i)(1-I_i^m)}.
\end{aligned}
\tag{7}
$$

Ignorability of the Missing Process   Conditional on vector x_i, if the missing mechanisms of the mother and the daughter are independent of their participation decisions, then we can simplify (3):

$$
\Pr\left\{I_i = r, I_i^m = s, y_i = v, y_i^m = w, x_i\right\} = \Pr\left\{I_i = r, I_i^m = s\right\} F\left(v, w, x; \theta\right) \Pr\left\{y_i^m = w, x_i\right\},
$$

such that the probability of a non-missing observation is

$$
\Pr\left\{I_i = 1, I_i^m = 1, y_i = v, y_i^m = w, x_i\right\} = \Pr\left\{I_i = I_i^m = 1\right\} F\left(v, w, x; \theta\right) \Pr\left\{y_i^m = w, x_i\right\}
$$

and the probability conditional on the observation being non-missing is

$$
F\left(v, w, x_i; \theta\right) \times \Pr\left\{y_i^m = w, x_i\right\}.
\tag{8}
$$

Thus, θ can be consistently estimated by Maximum Likelihood using only the observations for which no information is missing, and the missing process is thus ignorable.

4.3   Imputation

The traditional solution to non-ignorable missingness is to perform a procedure in which the missing values are imputed (see, among others, Little and Rubin, 2002). This implies that certain events are known to have zero probability. In Section 5, we present estimations obtained after assuming that women with missing occupations engage in domestic production (either as housewives or as unpaid farmers) and are, therefore, not participating in the labor force. This imputation is consistent with the vicar being more likely to

report special activities and leaving as missing the most common activities, i.e., farming or household production. Further, housewives and unpaid farmers are more likely either to live geographically distant from the church or to not belong to the influential society that the vicar would be more likely to visit during Lent. Thus, we consider the case in which, independent of everything else, women with missing occupations do not participate in the labor market, i.e., $\Pr\{y_i = 1 | I_i = 0, I_i^m, y_i^m, x_i\} = 0$ and the same holds for the mother.[14]

As derived in Appendix C, the likelihood is equal to

$$
\begin{aligned}
\prod_i p_i = \ & \prod_i \Big\{ F\left(y_i, y_i^m, x_i; \theta\right)^{I_i I_i^m} F\left(0, y_i^m, x_i; \theta\right)^{(1-I_i)I_i^m} \times \\
& F\left(y_i, 0, x_i; \theta\right)^{I_i(1-I_i^m)} F\left(0, 0, x_i; \theta\right)^{(1-I_i)(1-I_i^m)} \Big\} \times \\
& \prod_i \Big\{ \left(\Pr\{I_i = 1, I_i^m = 1 \,|\, y_i, y_i^m, x_i\} \Pr\{y_i^m, x_i\}\right)^{I_i I_i^m} \times \\
& \left(\Pr\{I_i = 0, I_i^m = 1 \,|\, y_i = 0, y_i^m, x_i\} \Pr\{y_i^m, x_i\}\right)^{(1-I_i)I_i^m} \times \\
& \left(\Pr\{I_i = 1, I_i^m = 0 \,|\, y_i, y_i^m = 0, x_i\} \Pr\{y_i^m = 0, x_i\}\right)^{I_i(1-I_i^m)} \times \\
& \left(\Pr\{I_i = 0, I_i^m = 0 \,|\, y_i = 0, y_i^m = 0, x_i\} \Pr\{y_i^m = 0, x_i\}\right)^{(1-I_i)(1-I_i^m)} \Big\}.
\end{aligned}
\tag{9}
$$

Under the assumption that the imputation is correct, the likelihood to estimate the conditional model (2) uses only the first two lines of equation (9). If the remaining terms in equation (9) depend on the parameters of model (2), this conditional Maximum Likelihood estimator will be consistent but may not be efficient. Otherwise, it will be identical to the full Maximum Likelihood estimator.

Table 6 presents unconditional transition rates in labor force participation status from mothers to daughters after the imputation. In the original data, the participation rate of those women whose mothers also work is unrealistically high at 47.06% (see Table 4). The figure declines to 5.78% for the sample with imputed values. Moreover, in the original data, the proportion of participating women increases from 5.34% to 47.06%—or 42.6 percentage points—when the mother also participates. That increase is substan-

---

[14]These assumptions set as impossible those events in which either the mother or the daughter (or both) participate in the labor market and for which information is missing, i.e., $\{I_i = 0, I_i^m = 1, y_i = 1, y_i^m, x_i\}$, $\{I_i = 1, I_i^m = 0, y_i, y_i^m = 1, x_i\}$, $\{I_i = 0, I_i^m = 0, y_i = 1, y_i^m = 1, x_i\}$, $\{I_i = 0, I_i^m = 0, y_i = 1, y_i^m = 0, x_i\}$, and $\{I_i = 0, I_i^m = 0, \ \ y_i = 0, y_i^m = 1, x_i\}$.

tially lower after the imputation, i.e., 4.5 percentage points. By reducing the differential in the transition of participation rates from mother to daughters between mothers who participated and mothers who did not, the imputation might downward bias the effect of mothers' labor market participation on that of their daughters.

Table 6: FEMALE PARTICIPATION BY THE MOTHER'S PARTICIPATION STATUS
Sample with imputed participation status

| | Mother: | | |
	Did not participate	Participated	Total
Daughter:			
Does not participate	$17,167$ (98.71%)	261 (94.22%)	$17,428$ (98.64%)
Participates	225 (1.29%)	16 (5.78%)	241 (1.36%)
Total	$17,392$ (100%)	277 (100%)	$17,669$ (100%)

Note: The Table is constructed based on the sample of 17,669 women with identified mothers in the dataset, who were born between 1675-1874. A woman is considered participating if we classify her as a Employee/Farmers or a Professional/Capital Owner. Women with missing values in their activity (93.10% of the daughters and 96.29% of mothers) are assumed to be non-active in the labor market.

Imputation procedures are, nonetheless, subject to criticism because they are *ad hoc* procedures in the way that they exploit the available information on the missing data mechanism. For example, when we observe only the mother (the daughter) and she is working, and then we impute as non-working the missing value for the daughter (the mother), we might bias downward the estimates of the marginal effect of the mother's status on that of the daughter. An alternative approach is to propose a model for the missing data mechanism and estimate the model of interest subject to the missing data generation process. In the next section, we follow Ramalho and Smith (2013) and state weak assumptions regarding the missing data mechanism to identify reduced-form participation while controlling for potentially non-ignorable missing information.

4.4 The likelihood approach to address non-ignorable missingness

Ramalho and Smith (2013) propose a likelihood-based approach to address non-ignorability in discrete choice models. As a special case, they consider the situation in which the missingness mechanism is conditionally dependent on the outcome variable and a discrete partition of the covariates. In our empirical application, this situation is intuitively plausible. As argued above, vicars might have been more likely to under-report the incidence of professions that were common, such as farmers, and more likely to record the professions for those whose labor status was a differentiating characteristic. These considerations warrant the following:

Assumption 1 *(Daughter's Response Conditional Independence, DRCI) Non-missingness in y_i is conditionally independent of I_i^m, y_i^m, and x_i; i.e.,*

$$\Pr\left\{I_i = 1 \,|I_i^m, y_i, y_i^m, x_i\right\} = \Pr\left\{I_i = 1 \,|y_i\right\}. \tag{10}$$

Assumption 1 cannot be considered an imputation procedure because it does not replace the missing observations with any set of values. Because 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) Non-missingness in y_i^m is conditionally independent of I_i, y_i, and x_i; i.e.,*

$$\Pr\left\{I_i^m = 1 \,|I_i, y_i, y_i^m, x_i\right\} = \Pr\left\{I_i^m = 1 \,|y_i^m\right\}. \tag{11}$$

Let $H_1 \equiv \Pr\left\{I_i = 1, y_i = 1\right\}$, $H_0 \equiv \Pr\left\{I_i = 1, y_i = 0\right\}$ and $H_1^m \equiv \Pr\left\{I_i^m = 1, y_i^m = 1\right\}$, $H_0^m \equiv \Pr\left\{I_i^m = 1, y_i^m = 0\right\}$. Furthermore, the marginal distributions of the discrete variables y_i, y_i^m are denoted $\Pr\{y_i = 1\} = \Pi_1$ and $\Pr\{y_i^m = 1\} = \Pi_1^m$, respectively. Finally $\Pi_{w,x} \equiv \Pr\left\{y_i^m = w, x_i\right\}$, where the number of parameters in $\Pi_{w,x}$ is given by the number of different combinations of the variables y_i^m and x_i that are 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 (10) and (11) imply that

$$\Pr\left\{I_i = I_i^m = 1, y_i = y_i^m = 1, x_i = x\right\} = \left(\frac{H_1}{\Pi_1}\right)\left(\frac{H_1^m}{\Pi_1^m}\right) F\left\{1, 1, x; \theta\right\} \Pi_{1,x} \qquad (12)$$

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

Consider the case in which the daughter's information is missing but the mother's is not. The joint probability for $\{I_i = 0, I_i^m = 1, y_i^m = 1, x_i = x\}$, which corresponds to the first term in equation (7), is

$$\begin{aligned}
\Pr\left\{I_i = 0, I_i^m = 1, y_i^m = 1, x_i = x\right\} &= \left\{\Pr\left\{I_i = 0, I_i^m = 1, y_i = 1, y_i^m = 1, x_i\right\}\right. \\
&\quad \left. + \Pr\left\{I_i = 0, I_i^m = 1, y_i = 0, y_i^m = 1, x_i\right\}\right\} \\
&= \left(1 - \frac{H_1}{\Pi_1}\right)\left(\frac{H_1^m}{\Pi_1^m}\right) F\left\{1, 1, x_i; \theta\right\} \Pi_{1,x_i} + \left(1 - \frac{H_0}{1 - \Pi_1}\right)\left(\frac{H_1^m}{\Pi_1^m}\right) F\left\{0, 1, x_i; \theta\right\} \Pi_{1,x_i} \\
&= \left\{\left(1 - \frac{H_1}{\Pi_1}\right) F\left\{1, 1, x_i; \theta\right\} + \left(1 - \frac{H_0}{1 - \Pi_1}\right) F\left\{0, 1, x_i; \theta\right\}\right\}\left(\frac{H_1^m}{\Pi_1^m}\right) \Pi_{1,x_i}
\end{aligned} \qquad (13)$$

as

$$\Pr\left\{I_i = 0 \,|\, I_i^m = 1, y_i = 1, y_i^m = 1, x_i\right\} = \Pr\left\{I_i = 0 \,|\, y_i = 1\right\} = 1 - \frac{H_1}{\Pi_1}$$

and

$$\Pr\left\{I_i = 0 \,|\, I_i^m = 1, y_i = 0, y_i^m = 1, x_i\right\} = \Pr\left\{I_i = 0 \,|\, y_i = 0\right\} = 1 - \frac{H_0}{1 - \Pi_1}.$$

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

$$\Pr\left\{I_i = 1, I_i^m = 0, y_i = 1, x_i\right\} = \left(\frac{H_1}{\Pi_1}\right)\left\{\left(1 - \frac{H_1^m}{\Pi_1^m}\right) F\left\{1, 1, x_i; \theta\right\} \Pi_{1,x_i}\right.$$
$$\left. + \left(1 - \frac{H_0^m}{1 - \Pi_1^m}\right) F\left\{1, 0, x_i; \theta\right\} \Pi_{0,x_i}\right\} \tag{14}$$

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

$$\Pr\left\{I_i = 0, I_i^m = 0, x_i\right\} = \left\{\left(1 - \frac{H_1}{\Pi_1}\right) F\left\{1, 1, x_i; \theta\right\} + \left(1 - \frac{H_0}{1 - \Pi_1}\right) F\left\{0, 1, x_i; \theta\right\}\right\} \times \left(1 - \frac{H_1^m}{\Pi_1^m}\right)\Pi_{1,x_i}$$
$$\left\{\left(1 - \frac{H_1}{\Pi_1}\right) F\left\{1, 0, x_i; \theta\right\} + \left(1 - \frac{H_0}{1 - \Pi_1}\right) F\left\{0, 0, x_i; \theta\right\}\right\} \times \left(1 - \frac{H_0^m}{1 - \Pi_1^m}\right)\Pi_{0,x_i}$$
$$\tag{15}$$

Thus, under assumptions (10) and (11), equation (7) becomes

$$p_i = \left(\left(\frac{H_{y_i}}{\Pi_{y_i}}\right)\left(\frac{H_{y_i^m}^m}{\Pi_{y_i^m}^m}\right) F\left\{y_i, y_i^m, x_i; \theta\right\} \Pi_{y_i^m, x_i}\right)^{I_i I_i^m} \times$$
$$\left(\left\{\sum_{v \in \{0,1\}} \left(1 - \frac{H_v}{\Pi_v}\right)\left(\frac{H_{y_i^m}^m}{\Pi_{y_i^m}^m}\right) F\left\{v, y_i^m, x_i; \theta\right\} \Pi_{y_i^m, x_i}\right\}\right)^{(1 - I_i)I_i^m} \times \tag{16}$$
$$\left(\sum_{w \in \{0,1\}}\left\{\left(\frac{H_{y_i}}{\Pi_{y_i}}\right)\left(1 - \frac{H_w^m}{\Pi_w^m}\right) F\left\{y_i, w, x_i; \theta\right\} \Pi_{w, x_i}\right\}\right)^{I_i(1 - I_i^m)} \times$$
$$\sum_{v, w \in \in \{0,1\}}\left\{\left(1 - \frac{H_v}{\Pi_v}\right)\left(1 - \frac{H_w^m}{\Pi_w^m}\right) F\left\{v, w, x_i; \theta\right\} \Pi_{w, x_i}\right\}^{(1 - I_i)(1 - I_i^m)}.$$

where $\Pi_0 = 1 - \Pi_1$ and $\Pi_0^m = 1 - \Pi_1^m$. A clarification concerning our notation is perhaps in order. 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 indicates that the relevant parameter is that which corresponds to the value at that observation. Thus, for example, $F\left\{y_i, w, x_i; \theta\right\}$ in the third row of (16) should be evaluated at the value that variables y_i and x_i have at observation i and a running value w instead of y_i^m (i.e., a value w that is not necessarily that observed for y_i^m at observation i). The vector of parameters includes, 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 combinations of the values of y_i^m and x_i in the data. Equation (16) represents the

likelihood $\mathcal{L}_i$ for any given observation i. The log-likelihood function results from the sum of the log of $\mathcal{L}_i$, $\log(\mathcal{L}) = \sum_{i=1}^{N} \log(\mathcal{L}_i)$. Subject to the following constraints

$$\sum_{w,x} \Pi_{w,x} = 1 \tag{17}$$

$$\Pi_v = \sum_{w,x} F\{v,w,x;\theta\} \Pi_{w,x} \text{ for } v = 0,1 \tag{18}$$

$$\Pi_w^m = \sum_x \Pi_{w,x} \text{ for } w = 0,1 \ , \tag{19}$$

Maximum Likelihood estimation will yield consistent and asymptotically efficient estimates of θ.

One of the difficulties associated with this model is that the vector of parameters $\Pi_{w,x}$ grows in the number of exogenous variables in x and depends on the number of possible combinations of all values in x. It is thus reasonable to reduce the number of parameters by introducing additional restrictions on $\Pi_{w,x}$. Decompose vector x into two types of variables x_1 and x_2, where x_1 includes location and time dummies and x_2 includes other variables, such as the dummy for a large number of siblings, a dummy for whether the father was influential, and a dummy that signals whether the mother was a property owner:

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

Our simplifying assumption is that the distribution of x_2 depends only on the location and time dummies and not on the mother's working status:

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

Because our set of controls is discrete, the ML estimator coincides with the GMM estimator proposed by Ramalho and Smith (2013) for more general cases. Although the model is identified, a very large proportion of missing observations likely affects the concavity of the likelihood function. In the next section, we describe how we improve the sample identification of the parameter vector by using external information from census

data in constructing $\Pi_{w|x_1}$.

Use of external information   Due to the large incidence of missingness in our data, identifying the parameters of the participation process is challenging. To improve identification, census data provide direct values for $\Pi_{w|x_1}$ that, following Imbens and Lancaster (1994), we can plug into the likelihood function, which results in a reduction of the set of parameters. Census data provide also information on Π_1 (i.e., the unconditional probability of participation) that implies an additional restriction on the likelihood function. Given an estimate of all other parameters, we use this restriction in conjunction with equation (18) to identify the constant β_0 in equation (2). As a result of the problems associated with census data that we discussed in Section 3.3, we consider the three alternative scenarios for female participation rates that we labeled the baseline, low, and high scenarios in Section 3.3 above.

In the remaining sections of the paper, we refer to this approach in addressing non-ignorable missing observations as the *likelihood approach*.

5   Main results

In this section, we report the marginal effect estimates of the participation status of the mother in equation (1). Table 7 presents the estimates of the parameters in equation (1) for the non-missing sample, for the sample with imputed values, and for the three alternative scenarios using the likelihood approach described in the previous section.

We find a positive and statistically significant effect of the mother's participation status on the daughter's probability of participating—parameter α in equation (1)—in all the estimations. Comparing Table 7 with Tables 4 and 6, the introduction of controls reduces the size of the marginal effect in both the non-missing sample and the sample with imputed values. Nevertheless, the marginal effects are positive and significant in all specifications. As anticipated when discussing Table 4 above, the imputation of all missing observations as inactive reduces substantially the effect of mothers' labor market participation on that of their daughters. Whereas the marginal effect estimated using

Table 7: Parameter Estimates, Average Marginal Effect, and Ignorability Test

	Ignorability	Imputation	Likelihood Approach		
			baseline	high	low
mother participated	1.4184***	0.6007***	1.2286***	1.6049***	1.0946***
	(0.379)	(0.196)	(0.211)	(0.209)	(0.191)
Horta	-1.8378***	0.7670***	0.8209***	0.7774***	0.7986***
	(0.641)	(0.072)	(0.122)	(0.138)	(0.100)
19th Century	1.1730***	0.5718***	0.6775***	0.7565***	0.6269***
	(0.377)	(0.081)	(0.108)	(0.126)	(0.095)
last quarter	0.4820**	0.5790***	0.7240***	0.9164***	0.6276***
	(0.193)	(0.064)	(0.095)	(0.120)	(0.080)
more than 4 siblings	0.3836**	-0.1049	-0.2489***	-0.2357**	-0.1965**
	(0.184)	(0.066)	(0.090)	(0.112)	(0.079)
father influential	0.0459	0.1652	0.8404***	0.9894***	0.6094***
	(0.198)	(0.104)	(0.109)	(0.159)	(0.099)
mother property owner	-0.3554	-0.0517	-1.4602***	-1.9058***	-1.1086***
	(0.439)	(0.270)	(0.196)	(0.199)	(0.181)
Avg. marginal effect	0.2891**	0.0302**	0.3227***	0.4709***	0.2189***
	(0.113)	(0.015)	(0.056)	(0.047)	(0.048)
Ignorability test statistic			74.0195	75.8044	180.9174
			[0.000]	[0.000]	[0.000]
Conditional probabilities:					
$\hat{Pr}\{I_i = 1 \,|\, y_i = 0\}$			0.0573	0.0591	0.0565
$\hat{Pr}\{I_i = 1 \,|\, y_i = 1\}$			0.4014	0.2134	0.6520
$\hat{Pr}\{I_i^m = 1 \,|\, y_i^m = 0\}$			0.0711	0.0843	0.0654
$\hat{Pr}\{I_i^m = 1 \,|\, y_i^m = 1\}$			0.1051	0.0555	0.2099
Avg. participation (%)	7.65	1.36	14.14	27.45	7.07
Number of observations	614	17669	17669	17669	17669

Note: "Ignorability" refers to the estimation results using only the sample without missing observations. "Imputation" displays the results using the sample after imputing non-activity in the labor market to all missing values. "Likelihood Approach" reports the estimates using the estimator for both the participation and missing processes using external information that is described in Section 4.4. "Avg. Marginal Effect" refers to the average change in the probability of participation when the mother participation status changes from no-participation to participation. The baseline scenario uses as external information the aggregate level of female labor force participation that we obtain from interpolations for the participation rates using census data. The high scenario takes the largest participation rate of the closest place for which we have information in the 1890, 1900, or 1911 censuses. The low scenario is 50% lower participation rates relative to the baseline scenario. Standard errors are in parenthesis and $p-$values are in brackets. "Ignorability Test Statistic" is the Wald statistic for the two equalities $\Pr\{I_i = 1 \,|\, y_i = 0\} = \Pr\{I_i = 1 \,|\, y_i = 1\}$ and $\Pr\{I_i^m = 1 \,|\, y_i^m = 0\} = \Pr\{I_i^m = 1 \,|\, y_i^m = 1\}$. "Conditional Probabilities" reports estimates using the Likelihood Approach estimator for both the participation and missing processes.

only non-missing observations is 28.91 percentage points, the estimated effect when using imputed observations decreases to only 3.02 percentage points. Nevertheless, the effect of the mother is positive and statistically significant even in this case.

The size of the marginal effect using the likelihood approach depends on the scenario considered. With low participation rates, as in the low scenario, the point estimate of the effect is smaller than when considering high participation rates, as in the high scenario. Similarly, the baseline scenario results in an estimate that lies between the estimates of the other two scenarios. The high scenario should be regarded as an extreme case scenario with arguably unrealistically high participation rates. By contrast, both the lower and baseline scenarios describe more plausible situations and do not result in substantially different estimates of the marginal effect than the estimation with the non-missing observations. The estimated marginal effects from both the baseline and low scenarios are very high, considering that the average value of female labor force participation in the baseline scenario, $\hat{\Pi}$, is only 0.141.

Despite what these results might suggest regarding the similarity of the effect of mothers' participation, ignoring the missing observations is not equivalent to the likelihood approach. For example, the effect of the number of siblings has a negative sign under the likelihood approach and under imputation but a positive sign in the sample with non-missing observations. Similarly, living in Horta increases participation under the likelihood approach and decreases it in the sample with non-missing observations.

The estimates for the remaining controls present the same signs regardless of the approach. Not surprisingly, time has a positive effect on participation. Finally, both *father influential* and *mother property owner* capture wealth and social status effects on participation. Controlling for these effects and further robustness checks that we present below allows us to discard the transmission of wealth and status as the explanation for the results concerning the marginal effect of the participation of the mother.

Ignorability tests

A comparison of point estimates and standard errors between the likelihood approach and the ignorability approach is not a test of the ignorability assumption because estimates differ in several dimensions. First, the samples are different because the likelihood approach exploits information from all observations, including those for which the participation of the mother or the participation of the daughter is missing. Second, the likelihood approach also employs external information. Finally, the likelihood function when ignoring missing observations, which is obtained from equation (8), assumes that the missing mechanism is independent of the participation decisions of mothers and daughters, i.e., it assumes $\Pr\{I_i = 1 \,|\, y_i = 0\} = \Pr\{I_i = 1 \,|\, y_i = 1\}$ and $\Pr\{I_i^m = 1 \,|\, y_i^m = 0\} = \Pr\{I_i^m = 1 \,|\, y_i^m = 1\}$. Here, we test in each of the three scenarios these two equalities, i.e., we test whether missingness is ignorable provided that DRCI and MRCI hold.

From Section 4.4, we know that $\Pr\{I_i = 1 \,|\, y_i = 0\} = \frac{H_0}{1 - \Pi_1}$, $\Pr\{I_i = 1 \,|\, y_i = 1\} = \frac{H_1}{\Pi_1}$, $\Pr\{I_i^m = 1 \,|\, y_i^m = 0\} = \frac{H_0^m}{1 - \Pi_1^m}$, and $\Pr\{I_i^m = 1 \,|\, y_i^m = 1\} = \frac{H_1^m}{\Pi_1^m}$, where H_1, H_0, H_1^m, and H_0^m, are a subset of the parameters estimated in the likelihood approach. Therefore, we use those estimates and their standard errors to test ignorability of the missing process using a Wald Test. The estimates of H_1, H_0, H_1^m, and H_0^m can be found in Table A2 in Appendix A. Table 7 contains the probabilities $\Pi_1 = \Pr\{y_i = 1\}$, $\Pi_1^m = \Pr\{y_i^m = 1\}$, $\Pr\{I_i = 1 \,|\, y_i = 0\}$, $\Pr\{I_i = 1 \,|\, y_i = 1\}$, $\Pr\{I_i^m = 1 \,|\, y_i^m = 0\}$, and $\Pr\{I_i^m = 1 \,|\, y_i^m = 1\}$ that result from the likelihood approach estimates, as well as the Wald test statistic of the ignorability of the missing process. The null hypothesis of ignorability is clearly rejected in all three scenarios.

6 Extensions and robustness checks

In this section, we first conduct robustness checks to evaluate whether our main results are sensitive to sampling issues. Next, we discuss the relevance of mechanisms beyond preference and beliefs that might account for the observed effect of the status of the mother.

6.1 Sampling issues

Horta and the 19th Century

In contrast to the villages of Ronfe and Ruivães, Horta is a city, and the professional choices and social structure of its inhabitants may differ from those of the inhabitants of Ronfe and Ruivães. Moreover, it is sensible to estimate the model using data for Horta alone, as it is where most of the records of economic activity/social status are registered. When we restrict the sample to the location of Horta, both the marginal effect estimate of the mother's participation status and its standard error remain similar to the results presented in Table 7. For example, in the baseline scenario, the average marginal effect becomes 0.3527 (vs. 0.3227 for the estimate using the full sample), and the standard error becomes 0.058 (vs. 0.056 using the full sample). Similarly, when we restrict the sample to data from the 19th century, the period in which the percentage of missing observations is lowest, the results are also nearly unchanged (0.2933 for the average marginal effect and 0.060 for its standard error).

Not surprisingly, because most of the identification comes from observations from Horta and observations from the 19th century, restricting the sample to these cases has little effect on the results. The estimation technique, thus, correctly assigns greater weight to observations with non-missing values. Moreover, for both the Horta and 19th century samples, we reject the null hypothesis of ignorability.

Migration

As we report in Panel B of Table 1, 27.5% of the original sample did not have their mothers identified in the dataset. It is plausible that a significant share of this percentage is explained by the migration of mothers or their daughters across villages. Certain women might have moved to neighboring villages or cities to work. If a share of the migration is related to labor force participation, then we may have a sample selection problem when excluding those women with non-identified mothers. Raw statistics from Horta, where we have information on place of birth, further confirm that those with non-identified mothers are more likely to be immigrants; as we show in the last two columns in Panel A of Table 8,

the percentage of observations born outside the city of Horta decreases substantially when we restrict the sample to observations with identified mothers. Additionally, if we compare women with and without identified mothers, we observe that, despite being almost as likely to have their participation status reported (6.90% and 7.60%, respectively), those with non-identified mothers are much more likely to participate (37.8%), conditional on being reported, than those with identified mothers (19.8%).

Table 8: DISTRIBUTION OF WOMEN BY PLACE OF BIRTH
Sample of Horta

Panel A: Observations by Place of Birth

	All observations		With identified mothers	
	No. of obs.	%	No. of obs.	%
Born in Horta city (matriz)	5,717	49.1	5,084	66.1
Born in Borough Horta	2,012	17.3	942	12.3
Born in other parts of Azores	3,028	26.0	1,373	17.9
Other birth place	99	0.85	39	0.51
Missing birth place	781	6.71	250	3.25
Total	11,637	100	7,688	100

Panel B: Distribution (%) of Reported and Participating by Place of Birth

	Reported		Participating	
	No	Yes	No	Yes
Born in Horta city (matriz)	88.22	11.78	87.48	12.52
Born in Borough Horta	80.47	19.53	73.91	26.09
Born in other parts of Azores	77.35	22.65	76.21	23.79
Other birth place	48.72	51.28	95.00	5.00
Missing birth place	84.68	15.32	90.63	9.38

Note: The Table is constructed based on the observations for the location of Horta only. The first two columns in Panel A include all female observations from the dataset (except slaves and their daughters) who were born between 1675-1874, while the last two columns include only those whose mothers are identified in the dataset. In Panel B, "Reported" refers to whether the activity and/or social status is reported by the vicar while "Participating" refers to the participation status if reported.

To assess whether the results presented thus far suffer from sample selection bias arising from migration, we re-estimate the model and include the women with non-identified mothers in the estimation sample. We include a dummy variable for a non-identified mother as an additional control because the missing process of the mother's participation

status is likely to be different from the missing process of the identity of the mother. The estimated coefficient for this dummy variable in the participation equation is positive and significant. That estimate, in conjunction with the lower estimate for the marginal effect of the mother's participation, is compatible with the notion that women with non-identified mothers are more likely to migrate to find work and that their mothers also worked, which would imply that migration is a potential source of bias for our estimated effect. Naturally, because we lack complete information on migrations, we cannot know the sources of the mothers' identity missingness process, and therefore, other interpretations of the results may be valid. In any event—and more importantly for our goal—the transmission effect from mother to daughter obtained when women with non-identified mothers are included in the estimation sample remains high (albeit lower) and highly statistically significant: in the baseline scenario, the average marginal effect is 0.1828 with a standard error of 0.051.

Returning to the sample with women with identified mothers, more direct evidence regarding migrants and their reporting and working status is presented in Panel B of Table 8 for location Horta. Considering all those born outside the city of Horta as immigrants, we conclude that immigrants have a lower proportion of missing information on participation status and are more likely to participate, conditional on that information being reported. To determine the extent to which our estimated transmission effect from mother to daughter is due to the presence of immigrants, we re-estimated our model using data exclusively from those born in the city of Horta. The estimates in the baseline scenario of the marginal effect of the mother's participation status and of its standard deviation are 0.3992 and 0.077, respectively. The transmission effect here is not significantly different from that estimated using the entire Horta sample or the sample based on all locations.

6.2 Beyond preferences and beliefs

6.2.1 On property transmission and poverty traps

To ensure that our model for the transmission of labor market participation behavior from mothers to daughters only captures the transmission of preferences or beliefs, we

must discard the possibility that it also captures effects related to the transmission of wealth in the form of land or capital. For example, when property is transmitted to the daughter, and mother and daughter have the same participation behavior, this behavioral similarity may be due to the transmission of property and not caused by the transmission of preferences. In all models presented thus far we already included a dummy for mother property owner. To isolate completely our estimates from property transmission effects we drop those observations for which the mother is a property owner (131 observations) from the estimation sample.[15] The estimate of the effect of mother's participation remains significant and positive in all cases. The point estimates are highly similar to those obtained with the full sample, both when ignoring missing observations and when using the imputed values. However, we obtain even larger point estimates when using the likelihood approach (for example, under the baseline scenario, the point estimate of the marginal effect of the mother increases to 0.51 percentage points with a standard error of 0.064), a result that is consistent with the negative sign of the slope coefficient of *mother property owner* in Table 7.

In contrast to property transmission, very poor mothers may transmit poverty to their daughters. In a society in which marriage and household production may be the social norm, day-worker landless women (from the Portuguese *jornaleiras*) may have resulted from a poverty trap that persists for generations. With unfavorable male-to-female ratios and no land, these landless women frequently do not marry and become single mothers.[16] They and their daughters must work for a living and will appear as participating in our data. The relevant question is then whether our results presented thus far are driven by this mechanism. We believe that such is not the case; in our main sample, after excluding observations for which the mother is not identified, we find only 8 daughters whose mothers are *jornaleiras* . Not surprisingly, dropping these observations does not affect the results.

---

[15]In the interest of brevity, these results are not presented here but are available upon request.

[16]Scott (1999) argues that the imbalance between the number of men and women led to unorthodox situations that did not conform to the social and religious conventions of the nation. In particular, a relatively large number of women were heads of households because their children were not conceived within a religious marriage. In Ronfe, for example, 20.7% of the heads of households were single females in 1750 and 18% of the children baptized in 1700 were illegitimate.

After restricting the sample to observations with identified mothers, the very small number of *jornaleiras* might be the result of *jornaleiras* being particularly mobile because of the unstable nature of their work. If that were the case, they would be particularly overrepresented in the population of mothers who are not identified. However, even when we estimate by including women with non-identified mothers, the estimated marginal effect of the mother's working status remains positive and statistically significant. Additionally, there are fewer *jornaleiras* in Horta, where we find a larger effect. All the foregoing suggests that poverty traps in the population of *jornaleiras* cannot be driving our main results.

6.2.2 Skills

In this section, we assess the extent to which our intergenerational effect is due to the transmission of skills. For example, the daughter of a seamstress would learn from her mother and become a seamstress herself. Lacking formal training, it is plausible that specific human capital was transmitted through family linkages. Do our results simply reveal the mother-to-daughter transmission of a particular craft or occupation? Alternatively, are daughters taking on the occupations of their fathers (Hellerstein and Morrill, 2011)?

Identifying the effect of the mother's participation status could be directly linked to those observations for which the profession of the daughter coincides with that of the mother. Conditional on participating in the labor market and on the mother's (father's) participation status being reported, in 19.1% (26.2%) of cases, the daughter has the same profession as the mother (father). Nevertheless, the actual number of observations with the same occupation is very small: 14 observations for which the occupations of the mother and daughter coincide and 16 observations in which the occupations of the daughter and the father coincide. Dropping these observations from the estimation sample does not affect our results. Thus, although we cannot exclude the possibility that there is intergenerational skill transmission in observations for which occupational choice is missing, we find that our results on the effect of the mother's participation status are not driven by the inclusion of daughters whose profession is identical to that of their mothers.

7 Conclusions

In this paper, we use historical parish registry data from four Portuguese locations from the 18th and 19th centuries to estimate a female labor force participation model that identifies the effect of mothers' labor market participation on that of their daughters. Importantly, we address the problem of missing values, which frequently affects historical records, by following the methodology recently proposed by Ramalho and Smith (2013). By allowing the estimation of models in contexts in which missing data are abundant and non-random, this methodology confers considerable potential to the examination of historical data that have yet to be explored.

Our results show a large and positive statistically significant effect of the mother's working status on the daughter's decision to participate in the labor market. In our preferred specification, the probability that a woman participates in the labor market increases by 32 percentage points if her mother also works, a very large effect given that the probability of female participation in the estimated model is 14.1%. A way to assess to what extent is the estimated marginal effect important for the evolution of the female labor market participation is to simulate the long-run impact on this rate of a technological change with and without mother-to-daughter transmission of preference/beliefs. Our simulation involves the following simple exercise: Consider a woman of the nineteenth century whose characteristics are those of the reference woman in our model. The probability that such a woman works given that her mother did not work, which was the most likely at the time, is 9%. Now imagine, there is a permanent technological change which doubles that probability, for example by increasing the estimate of the constant term in our model from -2.015 to -1.6. In the absence of mother-to-daughter transmission of preference/beliefs, the expected long-run female labor market participation would be 18%. On the contrary, when mother-to-daughter transmission occurs, the unconditional probability of working for the generation that follows the technological change is 24%, and as an increasing proportion of descendants influence the participation decisions of their own daughters the long-term participation rate increases to 32%. We argue that the existence of such a transmission mechanism served as an important catalyst for the

increase in female labor force participation when the technological change took place in the 20th century.

References

ALESINA, A., AND P. GIULIANO (2010): "The power of the family," *Journal of Economic Growth*, 15(2), 93–125.

——— (Forthcoming): *Family Ties*North Holland.

ALESINA, A., P. GIULIANO, AND N. NUNN (2013): "On the origins of gender roles: Women and the plough," *The Quarterly Journal of Economics*, 128(2), 469–530.

AMORIM, M. N. (1991): "Uma metodologia de Reconstituiçao de Paróquias," Discussion paper, Universidade do Minho.

AMORIM, N., AND C. SANTOS (2009): "Marriage Strategies in Azorean Communities of Pico Island (19th Century)," in *The Transmission of Well-Being: Gendered Marriage Strategies and Inheritance Systems in Europe (17th-20th Centuries)*, ed. by A. Margarida Duraes, L. Fauve-chamoux, and J. K. Ferrer. Peter Lang Pub Incorporated, Bern.

BISIN, A., AND T. VERDIER (2001): "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory*, 97(2), 298–319.

BRETTELL, C. B. (1991): "Kinship and contract: Property transmission and family relations in northwestern Portugal," *Comparative Studies in Society and History*, 33(03), 443–465.

BÜTIKOFER, A. (2013): "Revisiting 'mothers and sons' preference formation and the female labor force in Switzerland," *Labour Economics*, 20(0), 82 – 91.

CARRILHO, M. J. (1996): "População Activa: Conceito e Extensão através dos Censos," Discussion paper, Lisbon: INE.

DURÃES, M. (2009): "Providing Well-Being to Women through Inheritance and Succession," in *The Transmission of Well-Being: Gendered Marriage Strategies and Inheritance Systems in Europe (17th-20th Centuries)*, ed. by A. Margarida Duraes, L. Fauve-chamoux, and J. K. Ferrer. Peter Lang Pub Incorporated, Bern.

FARRÉ, L., AND F. VELLA (2013): "The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labour Force Participation," *Economica*, 80(318), 219–247.

FERNÁNDEZ, R. (2007): "Women, work, and culture," *Journal of the European Economic Association*, 5(2-3), 305–332.

——— (2011): "Does Culture Matter?," *Handbook of Social Economics*, 1A.

FERNÁNDEZ, R. (2013): "Cultural change as learning: The evolution of female labor force participation over a century," *The American Economic Review*, 103(1), 472–500.

FERNÁNDEZ, R., AND A. FOGLI (2009): "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American economic journal. Macroeconomics*, 1(1), 146–177.

FERNÁNDEZ, R., A. FOGLI, AND C. OLIVETTI (2004): "Mothers and sons: Preference formation and female labor force dynamics," *The Quarterly Journal of Economics*, 119(4), 1249–1299.

GALOR, O., AND D. N. WEIL (1996): "The Gender Gap, Fertility, and Growth," *American Economic Review*, 86(3), 374–87.

GOLDIN, C. (1990): *Understanding the Gender Gap: an Economic History of American Women*. Oxford University Press.

GOLDIN, C. (2006): "The Quiet Revolution That Transformed Women's Employment, Education, And Family," *American Economic Review*, 96(2), 1–21.

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).

GREENWOOD, J., A. SESHADRI, AND M. YORUKOGLU (2005): "Engines of liberation," *The Review of Economic Studies*, 72(1), 109–133.

GUISO, L., P. SAPIENZA, AND L. ZINGALES (2006): "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 20(2), 23–48.

HELLERSTEIN, J. K., AND M. S. MORRILL (2011): "Dads and Daughters: The Changing Impact of Fathers on Women's Occupational Choices," *Journal of Human Resources*, 46(2), 333–372.

HUMPHRIES, J., AND C. SARASÚA (2012): "Off the Record: Reconstructing Women's Labor Force Participation in the European Past," *Feminist Economics*, 18(4), 39–67.

IMBENS, G. W., AND T. LANCASTER (1994): "Combining micro and macro data in microeconometric models," *The Review of Economic Studies*, 61(4), 655–680.

INE (1977): *Estatísticas do Emprego*. Instituto Nacional de Estatística, Lisboa.

LAINS, P. (2006): "Growth in a protected environment: Portugal, 1850–1950," *Research in Economic History*, 24, 119–160.

LITTLE, R. J., AND D. B. RUBIN (2002): *Statistical analysis with missing data*, vol. 333. Wiley-Interscience.

MATOS, P. L. (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)*, ed. by A. Margarida Duraes, L. Fauve-chamoux, and J. K. Ferrer. Peter Lang Pub Incorporated, Bern.

MAURIN, E., AND J. MOSCHION (2009): "The Social Multiplier and Labor Market Participation of Mothers," *American Economic Journal: Applied Economics*, pp. 251–272.

MOREIRA DA SILVA, R. F. (1983): "Contraste e mutações na paisagem agrária das planícies e colinas minhotas," *Studium Generale. Estudos contemporâneos*, 5, 9–115.

MORRILL, M. S., AND T. MORRILL (2013): "Intergenerational links in female labor force participation," *Labour Economics*, 20(1), 38–47.

OLIVETTI, C., E. PATACCHINI, AND Y. ZENOU (2013): "Mothers, Friends and Gender Identity," Discussion paper, CEPR Discussion Paper No. 9712.

PINA-CABRAL, J. D. (1986): "Sons of Adam, daughters of Eve: The peasant worldview of the Alto Minho," *Sons of Adam, daughters of Eve: the peasant world view of the Alto Minho.*

RAMALHO, E. A., AND R. J. SMITH (2013): "Discrete Choice Non-Response," *The Review of Economic Studies*, 80(1), 343–364.

REIS, J. (2005): *O Trabalho*Imprensa de CiÃšncias Sociais, Lisboa.

SCOTT, A. S. V. (1999): *Famílias, formas de união e reprodução social no noroeste português (séculos XVIII e XIX)*. NEPS-Universidade do Minho.

SOLSTEN, E. (1993): "Portugal: A country study," Federal Research Division Library of Congress.

TABELLINI, G. (2008): "PRESIDENTIAL ADDRESS INSTITUTIONS AND CULTURE," *Journal of the European Economic Association*, 6(2-3), 255–294.

TSUKAHARA, I. (2007): "The effect of family background on occupational choice," *Labour*, 21(4-5), 871–890.

Appendix A

Table A1: MOST COMMON PROFESSIONS/SOCIAL STATUS FOR WOMEN
Lent Files

	No. of obs.	%
Madam	708	58.08
Housewife	269	22.07
Poperty owner/landlord	84	6.89
Independent worker	60	4.92
Seamstress	44	3.61
Craftsman textile	19	1.56
Domestic servant	7	0.57
Assistant	5	0.41
Farmer	4	0.33
Landless worker, day laborer	3	0.25
Laundress	2	0.16
Teacher	2	0.16
Other	9	0.74
Total	1219	100

Note: Sample of women born between 1675-1874 with identified mothers. For this table we aggregated similar professions/social status within the same category. Some professions/social status, however, such as "Madam" (from the portuguese "Dona") or "Housewife" (from the portuguese "Doméstica") had enough observations such that no aggregation was needed. "Other" includes professions/social status with only 1 observation (actress, midwife, carder, winder, woman who irons, shopkeeper, religious, baroness, women who look after the house/property of others).

Table A2: ESTIMATES OF THE JOINT PROBABILITIES

	Likelihood Approach		
	baseline	high	low
$H_1 \equiv \Pr\{I = 1, y = 1\}$	0.0568	0.0586	0.0461
	(0.0070)	(0.0079)	(0.0057)
$H_0 \equiv \Pr\{I = 1, y = 0\}$	0.0492	0.0429	0.0525
	(0.0015)	(0.0014)	(0.0016)
$H_1^m \equiv \Pr\{I^m = 1, y^m = 1\}$	0.0149	0.0152	0.0148
	(0.0009)	(0.0009)	(0.0009)
$H_0^m \equiv \Pr\{I^m = 1, y^m = 0\}$	0.0610	0.0612	0.0607
	(0.0018)	(0.0018)	(0.0018)

Note: "Likelihood Approach" reports the estimates using the estimator for both the participation and missing processes using external information described in Section 4.4. The baseline, high, and low scenarios are defined in Section 3.3. Standard errors are in parenthesis.

Appendix B

In this appendix, we describe the external data used to obtain aggregate female participation rates that are included in the likelihood approach as the "baseline" scenario. The first Portuguese census was administered in 1864, and since then, 11 censuses have been conducted more or less periodically. In most censuses, the smallest geographical area for which demographic data are collected is the Borough (*Concelho*), followed by the District (*Distrito*) and the Province. Most censuses also publish information regarding economic activity and even the professions of men and women above a certain age at various levels of geographical aggregation, which unfortunately varies across censuses. The census collects data for all regions of Portugal.

For the purpose of obtaining female labor force participation rates, we only assemble data from Portugal as a whole, the two largest cities (Lisbon and Oporto) and from regions that are more representative of the villages in our sample. The data assembled are a combination of Borough- and District-level data and the national totals for Portugal as a whole. Specifically, the regions are: 1) Portugal (including the 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 S. Mateus belong; 4) the Boroughs (cities) of Lisbon and Oporto; 5) the Borough of Guimarães to which Ronfe belongs; 6) the Borough of Vila Nova de Famalicão to which Ruivães belongs; 7) the Borough (city) of Horta; 8) and, finally, the Borough of Madalena (on Pico Island) to which S. Mateus belongs.

Table B1 reports our best approximate calculation of the female labor market participation rates for nine regions in Portugal over the 1864-1991 period based on the census data. Although we use a single data source (with the exception of 1864, for which we also use data published in Reis, 2005), the participation rates reported in Table B1 exhibit large jumps across censuses—which are likely caused by changes in the definition of "active" female population during the period—and substantial differences across regions for a given census. Carrilho (1996) provides a detailed guide for the different definitions of active female population across the censuses. One of the main differences regards the

	Country	City			District		Borough		
	Portugal	Lisbon	Oporto	Horta	Braga	Horta	Guimarães	V. N. Famalicão	Madalena
1864	0.191	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a
1890	0.385	0.23	0.338	0.111	0.513	0.103	0.552	0.622	0.038
1900	0.307	0.276	0.38	0.187	0.534	0.146	0.532	0.511	0.1
1911	0.292	0.281	0.403	0.217	0.413	0.188	0.539	0.364	0.162
1925	n.a	0.284	0.259	n.a	n.a	n.a	n.a	n.a	n.a
1930	0.186	0.261	0.317	0.092	0.335	0.079	0.375	0.322	0.062
1940	0.203	0.258	0.332	n.a	0.329	0.065	n.a	n.a	n.a
1950	0.213	0.287	0.343	0.078	0.301	0.061	0.423	0.366	0.054
1960	0.169	0.315	0.365	0.084	0.237	0.055	0.336	0.288	0.029
1970	0.24	0.328	0.391	n.a	0.334	0.073	n.a	n.a	n.a
1981	0.382	0.439	0.466	0.217	0.483	n.a	0.59	0.532	0.146
1991	0.437	0.433	0.462	0.325	n.a	n.a	0.575	0.55	0.286

Note: Authors computations from Recenseamentos Gerais da População (INE, Lisbon) and Reis (2005). The 1925 census was a special census restricted to the cities of Lisbon and Oporto. In 1864, the active female population is taken from the total number (all ages) of working women (mão-de-obra feminina) in Reis (2005). Until 1950, participation rates are computed using as denominator the female population older than 10. For 1950, the denominator is the female population older than 12. For 1960 and 1970, the denominator is the female population older than 15, while for 1981 and 1991 it is the female population older than 14 and 12, respectively. In 1890, 1900, 1911 the active female population is defined as females (including the servants) minus the housewives (unpaid domestic work) and the improductive (data taken from Table V in 1890 and 1900 and from Table 3 in 1911. In 1925, the actives are the females minus the female children between 0-9 years old, the housewives (unpaid domestic work), those without profession, and the female beggars. In 1930, the data comes from Table 1 and the actives are the sum of the three first columns under "população activa" (i.e. females who work for the administration, in the private sector, and the self-employed) minus the housewives (unpaid domestic work) who are considered in the column of active and self-employed. In 1940, the actives are the number of active females ("activas") minus the number of females in non-professional activities (information from Table 24). In 1950, the actives are the active females ("activas") with profession (information from Table 1). In 1960, the actives are the active females ("activas") with profession minus active females with profession less than 15 years of age (Table 1, Tomo 5, vol III). In 1970, the actives are the active females ("activas") with profession (Table 8). In 1981 and 1991, the actives are the active females ("activas") (Tables 6.13 and 6.16.1, respectively).

reference population. For example, whereas until 1930, the reference population was considered the "present" population (população de facto), after 1940, the censuses considered the "resident" population. Moreover, although until 1930, all individuals were included in the reference population, after 1940, only those above a certain age (either age 10 or 12, depending on the census) were considered. The other main difference refers to the definition of "active" population. The earlier censuses (until circa 1950) defined as active population everyone with an occupation regardless of whether that occupation was a profession—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, particularly in the districts and boroughs that were less urban (e.g., compare Lisbon with Guimarães). The islands are somewhat of an exception, where the rural borough of Madalena has a very low participation rate, even lower than the capital of the district, the city of Horta. Since 1960, unemployed individuals seeking a job are included in the active population. The values for female participation presented in Table

B1 may appear relatively low, when compared, for example, to the figures presented in the Introduction because the denominator of all rates includes the elderly as part of the reference population.

Although the data reported in Table B1 are as homogeneous as possible, inconsistencies remain in the definitions of female participation. To correct for these inconsistencies, we interpolate the female participation values for the baseline specification of our model using the following function:

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

where $y_{it} = ln\left(\frac{p_{it}}{1-p_{it}}\right)$ and p_{it} correspond to the values of the female participation rate presented in Table B1 for region i and census year t. The function $f(t)$ corresponds to a logit transformation of the census year t, specifically, $f(t) = \frac{\exp\left(\frac{(t-2000)}{\gamma}\right)}{1+\exp\left(\frac{(t2000)}{\gamma}\right)}$, where $\gamma = 15$ (selected such that the sigmoid shape of $f(t)$ better fits the evolution of the national female participation rate). The function $g(t) = 1(t \leq 1940)(1940 - t)$ is a specific trend for the values before 1940 to prevent inflated values of our data before 1940 and its unlikely trend to contaminate our predictions, while retaining the information regarding the differences across regions for those years. The $D_r's$ represent regional dummies, specifically, a dummy for the Braga district (encompassing Braga District, Guimaraes and V.N. Famalicão), a dummy for the Horta District (encompassing, Horta District, the city of Horta and Madalena), a dummy for a large city (encompassing Lisbon and Oporto), a dummy for Guimarães, a dummy for V.N. Famalicão, a dummy for Horta (city), and a dummy for Madalena. The $D_k's$ are other dummies used to account for the specificities of the particular censuses, such as a dummy for years in which the participation rates are applied to women older than 14 years of age (1960, 1970, 1981), and a specific dummy for 1981, when there is a particularly large jump in the participation rate due to the manner in which the data are reported in the census.

The unknown parameters of function B.1 are obtained by fitting this function to the observed data in Table B1 using a quadratic loss criterion. The interpolation then consists

of activating the relevant dummies. For example, for Ronfe, we set the dummy for the Braga district and the dummy for Guimarães Borough to 1, we set the $g(t)$ function to 0, and we set the dummies for older than 14, 1981, and large city to zero as well. The resulting interpolated female participation rates for Ronfe, Ruivães, Horta, and S. Mateus during our sample period are presented in Table 5 in Section 3.3 under the "baseline" column.

Appendix C

In this appendix, we derive the likelihood function for the traditional imputation procedure. From equations (4), (5), and (6), we can rewrite p_i as

$$
\begin{aligned}
p_i = \ & \left(\Pr\left\{ I_i = I_i^m = 1, y_i, y_i^m, x_i \right\} \right)^{I_i I_i^m} \times \\
& \left(\sum_{v \in \{0,1\}} \Pr\left\{ I_i = 0, I_i^m = 1, y_i = v, y_i^m, x_i \right\} \right)^{(1-I_i)I_i^m} \times \\
& \left(\sum_{w \in \{0,1\}} \Pr\left\{ I_i = 1, I_i^m = 0, y_i, y_i^m = w, x_i \right\} \right)^{I_i(1-I_i^m)} \times \\
& \left(\sum_{v \in \{0,1\}} \sum_{w \in \{0,1\}} \Pr\left\{ I_i = I_i^m = 0, y_i = v, y_i^m = w, x_i \right\} \right)^{(1-I_i)(1-I_i^m)} .
\end{aligned}
\tag{C.1}
$$

The traditional imputation procedure in which missing values are filled in implies that certain events are known to have zero probability. After imputing the missing observations, equation (C.1) simplifies to

$$
\begin{aligned}
p_i = \ & \left(\Pr\left\{ I_i = I_i^m = 1, y_i, y_i^m, x_i \right\} \right)^{I_i I_i^m} \times \\
& \left(\Pr\left\{ I_i = 0, I_i^m = 1, y_i = \overline{v}, y_i^m, x_i \right\} \right)^{(1-I_i)I_i^m} \times \\
& \left(\Pr\left\{ I_i = 1, I_i^m = 0, y_i, y_i^m = \overline{w}, x_i \right\} \right)^{I_i(1-I_i^m)} \times \\
& \left(\Pr\left\{ I_i = I_i^m = 0, y_i = \overline{v}, y_i^m = \overline{w}, x_i \right\} \right)^{(1-I_i)(1-I_i^m)}
\end{aligned}
\tag{C.2}
$$

where $y_i = \overline{v}$ ($y_i^m = \overline{w}$) denotes that the only admissible value $\overline{v}$ ($\overline{w}$) is imputed in the observation, and y_i and y_i^m denote observed values. By equation (3), we can write p_i in equation (C.2) as:

$$
\begin{aligned}
p_i = \ & \left(\Pr\left\{ I_i = 1, I_i^m = 1 \,|\, y_i, y_i^m, x_i \right\} \times F\left(y_i, y_i^m, x_i; \theta \right) \times \Pr\left\{ y_i^m, x_i \right\} \right)^{I_i I_i^m} \times \\
& \left(\Pr\left\{ I_i = 0, I_i^m = 1 \,|\, y_i = \overline{v}, y_i^m, x_i \right\} \times F\left(\overline{v}, y_i^m, x_i; \theta \right) \times \Pr\left\{ y_i^m, x_i \right\} \right)^{(1-I_i)I_i^m} \times \\
& \left(\Pr\left\{ I_i = 1, I_i^m = 0 \,|\, y_i, y_i^m = \overline{w}, x_i \right\} \times F\left(y_i, \overline{w}, x_i; \theta \right) \times \Pr\left\{ y_i^m = \overline{w}, x_i \right\} \right)^{I_i(1-I_i^m)} \times \\
& \left(\Pr\left\{ I_i = 0, I_i^m = 0 \,|\, y_i = \overline{v}, y_i^m = \overline{w}, x_i \right\} \times F\left(\overline{v}, \overline{w}, x_i; \theta \right) \times \Pr\left\{ y_i^m = \overline{w}, x_i \right\} \right)^{(1-I_i)(1-I_i^m)}
\end{aligned}
\tag{C.3}
$$

such that, rearranging those factors, the likelihood is equal to

$$
\begin{aligned}
\prod_i p_i = \quad & \prod_i \Big\{ F\left(y_i, y_i^m, x_i; \theta\right)^{I_i I_i^m} F\left(\overline{v}, y_i^m, x_i; \theta\right)^{(1-I_i) I_i^m} \times \\
& \quad F\left(y_i, \overline{w}, x_i; \theta\right)^{I_i (1-I_i^m)} F\left(\overline{v}, \overline{w}, x_i; \theta\right)^{(1-I_i)(1-I_i^m)} \Big\} \times \\
& \prod_i \Big\{ \left(\Pr\left\{I_i = 1, I_i^m = 1 \,|\, y_i, y_i^m, x_i\right\} \Pr\left\{y_i^m, x_i\right\}\right)^{I_i I_i^m} \times \\
& \quad \left(\Pr\left\{I_i = 0, I_i^m = 1 \,|\, y_i = \overline{v}, y_i^m, x_i\right\} \Pr\left\{y_i^m, x_i\right\}\right)^{(1-I_i) I_i^m} \times \\
& \quad \left(\Pr\left\{I_i = 1, I_i^m = 0 \,|\, y_i, y_i^m = \overline{w}, x_i\right\} \Pr\left\{y_i^m = \overline{w}, x_i\right\}\right)^{I_i (1-I_i^m)} \times \\
& \quad \left(\Pr\left\{I_i = 0, I_i^m = 0 \,|\, y_i = \overline{v}, y_i^m = \overline{w}, x_i\right\} \Pr\left\{y_i^m = \overline{w}, x_i\right\}\right)^{(1-I_i)(1-I_i^m)} \Big\}.
\end{aligned}
$$
$$\text{(C.4)}$$

If independent of everything else, women with missing occupations do not participate in the labor market—as we consider in Section 4.3— then $\overline{v} = \overline{w} = 0$ and equation (9) follows.