# *Online Appendix*

*La “Doña” è Mobile: The Role of Women in Social*

*Mobility in a Pre-modern Economy*

# A Appendix

# A.1 *Hidalgos* and *dons*

In Spain, the term *hidalgo* refers to the low aristocracy. It is typically reserved for people who are of noble origin but do not hold any nobility title. Dewald (1996) notes that during the 18*th* century there were two regions in Europe with different patterns regarding the number of aristocrats: Central Europe, with a percentage lower than 1%, and the periphery (Spain, Poland, Hungary and England), with more than 5%. Madramany y Calatayud (1788) mentioned that *hidalgos* were originally from “good places,” meaning places with little Arab influence in the north of Spain (Galicia, Asturias, Cantabria and the Basque Country). Therefore, lineage and place of origin, explain the high percentage of *hidalgos* in northern Spain. Moreover, the percentage in the remaining of Spain are closer to the 1% in central Europe. It is the high percentage in northern Spain makes the overall average closer to peripheral Europe.

According to the Floridablanca census (1787), 4*.*6% of the population in Spain were *hidalgos*. The northern provinces have over 10%, whereas Extremadura and Andalusia in the South have less than 1%. In the Kingdom of Aragon, there was traditionally a lower presence of *hidalgos* (Tomás Faci, 2015): 1*.*5% in Aragon, 0*.*14% in Valencia and 0*.*16% in Catalonia (INE, 2019). In the Region of Murcia, which belongs to the Kingdom of Castile, but shares many characteristics with Valencia and Catalonia, the number is 1*.*80%. In the city of Murcia, which is the interest of the article, the percentage of *hidalgos* was 0*.*45%. The fraction of *dons* that we observed in the data is an order of magnitude higher. This means, consistent with our interpretation, that the term *don*, by the 18*th* century, was no longer reserved for the nobility. On the contrary, the *nouveau riche*, the clergy and the higher ranks of the military also hold the honorary title. Dewald (1996) notices the sharp reduction in the percentage of aristocrats in Europe, beginning in the 18*th* century. In Spain, we observe a sharp decline between 1768 and 1787, from 7*.*9% to 4*.*6% (INE, 2019). In our data, however, we do not observe a sharp decline or rise in the fraction of brides or grooms carrying the honorific title of *don* or *doña*.

## A.2 Data description

Table A2 shows the fraction of the grooms and brides with the honorific, separately based on the status of the parents of the marrying couples. Approximately 6% of grooms and 8% of brides are recorded as high status. It is immediately clear that there is a high degree of persistence in this measure of social status. Of the 890 grooms with both a high-status father and mother, 89% of them were themselves a *don* at marriage. This fraction is similar for brides. Conversely, only about 1% of the much larger number of individuals with low status parents were themselves *dons* or *doñas*. In general, the transmission from parents to grooms versus that to brides is similar for various combinations of status of parents, with one major exception. If the father is low status but the mother is high status, then brides are much more likely to be high status than are grooms. In general, this suggests that the mobility *as measured at different ages* was downward. Of course, some of those individuals may later acquire *don* status.

In Table A3, we show the fraction of grooms and brides which marry a high-status partner, based on their own status and that of their own parents. Of grooms listed with *don* status, 90% are marrying brides who are *doñas*. The comparable number for *doñas* is 73%. The parents’ status matters in mating above their effect on the status of their children. In the first row, for example, we can see that while 92% of *don* grooms whose parents were themselves high status marry a high status *doña* woman, only 75% of *dons* whose parents were both low status have a high-status bride. For women, this difference is even stronger: While 96% of *doña* brides with high status parents marry a *don*, only 35% of “upwardly mobile” *doñas* with low status parents marry someone of high status. This is true for low status individuals as well: A downwardly mobile groom with high status parents has a 30% chance of marrying a high-status bride, versus only 1.4% of low status grooms with low status brides. This table thus reveals several interesting facts about assortative mating. Most obviously, much like the high intergenerational transmission of status, assortative mating on status is very strong. Second, and much less obviously, the degree of assortative mating depends on the status of the parents as well as that of the children. That is, a high-status groom and particularly a high-status bride who comes from higher status parents is more likely to marry a high-status individual than a high-status individual with low status parents. Goñi (forthcoming) shows how assortative mating among the high classes of England increased inequality, but we are not aware of any study of mating among the low classes.

## A.3 Linking

In this section, we describe the procedure we used to link marriage records from one generation to the previous generation. The only link variables are names: The names of both parents of either the bride or groom (in the children’s generation), and the names of the marrying couple (in the parents’ generation). For each generation, we restrict the possible set of links in the parents’ generation to the set of marriages between 18 and 45 years before the marriage of the children; in other words, we assume that individuals marry between ages 18 and 45.

Historical records may contain errors. Errors could be due to the transcription from the physical to the digital sources or be made by the priests when they originally wrote the records. We now describe in detail the various cleaning methods we employed in the marriage records during the transcription. We detail those below:

* **Contractions**: There are many surnames that were originally created with the composition of an article and a name but got simplified over time. de la Fuente become la Fuente, and later becomes Lafuente. de la Jara becomes la Jara or Laxara, and later becomes Jara. One needs to be particularly careful with these names because some of them still have the archaic forms, i.e., some families are named Laxara and other families are named Jara during our period, even if their ancestors were all named de la Jara. During the transcription, we tried to stick to the actual spelling, except in the cases where two or more different spelling appear in the same document, i.e., father named la Fuente and his son named Lafuente.
* **Archaic Orthographic rules**: Many words appear in their archaic form, or in the local dialect form. For consistency, we wrote all names and words in their contemporaneous spelling, i.e., we write Alejandro instead of Alexandro. The most commons are: Ph 🡪 F; X 🡪 J; Ch 🡪 S; J 🡪 G; H 🡪 dissapears; Duplicates: Ss, Nn; S 🡪 Z; I, LL 🡪 Y; Y 🡪 I; N 🡪 Ñ; Z 🡪 C.
* **Prefixes**: Some names were originally written with prefixes, and they evolve over time and the prefix disappears, e.g., Escarabajal becomes Carvajal.
* **Vowel contractions**: In the local dialect, it is common to suppress a consonant and instead create a contraction between the adjacent vowels, e.g., Albadalejo becomes Albalaejo and Poveda becomes Povea.
* **Hernández and Fernández**: These are very common surnames in the area. They differ by only one letter. Hence, the use of fuzzy or probabilistic linking would link two individuals with these surnames, even if they are not the same individual. Nonetheless, in the original records, we found the same individual with different spelling in each record, e.g., Alejandro Fernández in the first record and Alejandro Hernández in the second record. We know he is the same individual due to the coincidence of the other names in both records, i.e., spouse’s name, parents’ names and place of origin. We fix these errors in the transcription but do not attempt to fix them with an algorithm or fuzzy/probabilistic linking.
* **Specific Names**: In addition to the general rules specified above, we use knowledge of the local dialect and names to fix spelling errors during the transcription. Table A1 shows the specific edits for names and surnames.

After the transcription, the linking is performed iteratively, similar to Ferrie (1996) and Abramitzky et al. (2014). In each step, the links are removed from the set of both parents and children linked later. In all links, we first perform a minimal set of cleaning after parsing the names into given names and surnames.[[1]](#footnote-1) This is limited to stripping out accents and the *ñ* and removing the words “Las”, “Los”, “La”, “De”, “Del”, and “De La”.

In the first step, we link on exact names, *i.e.*, we require the spelling to be identical, on four words: the given name and the surname of the parents, and disregard multiple links. Multiple links are relatively rare given the small area from which the marriages come and the use of four names to link. Thus, the iterative process is useful to create new links that were not created with the *raw* names. In the second step, we link replacing the names as transcribed and cleaned with a set of standardized spellings, and then again link on all four words. For example, for surnames both “Hespinosa” and “Hespinossa” are standardized to “Espinosa,” and for given names both “Josef” and “Joseph” are standardized to “Jose.” Finally, we use a set of standardized phonetic changes to clean the names, and repeat linking all four names. These substitutions were as follows: “Y” for “I” or “LL”, “B” for “V”, “C” for “Z”, “F” for “Ph”, “T” for “Th” or “TT”, “G” for “J” or “X”, “S” for “CH” or “SS”, “N” for “NN” or “Ñ”, “HU” for “GÜ”, “Q” for “Qu”, and removing any remaining “H”.[[2]](#footnote-2) It is worth noticing that the same individual could sometimes appear as *Fernández* and other times as *Hernández*. In those, more intricate cases, we have resorted to linking by hand before our iterative process.

After the second step, we used a novel third step. We look for the small number of remaining multiple links. In this final population of multiple links, there is still some remaining information. In particular, we have many marriages where all the potential links all have the same status information on all relevant variables. For the regressions using grandparents, that means that the status as a *don* is the same for all six individuals listed in the candidate link. For the groom’s regressions, these six individuals are: the groom’s parents and grandparents. Notice, that we are not requiring those to be the same to the original marriage. As we explain above, status could change over time. Given the nature of the data, there are 26 = 64 possible combinations, *i.e.*, we have 6 individuals that could have a binary value.[[3]](#footnote-3) In practice, there is high correlation among the values for those 6 individuals and some combinations are extremely rare. Moreover, because *don* status is held by a relatively small minority of the population, the group where all 6 individuals are not *don* makes up a substantial fraction of the relatively small number of multiple links. This is particularly important in this, and other settings, where usually the individuals with low status are also those with common names, and they are out of the linked sample. Therefore, adding these final links helps in reducing the selection incurred when linking on names.

In the original dataset, we have 18,175 grooms and brides. In the first stage, when we link directly on exact names, we end up with 6,832. After the second stage, when we link the remaining observations using phonetically clean names, we end up with 6,832+41+48=6,921 matches, and 139 observations with multiple candidate matches. Of those 139 matches we are able to create 96 synthetic observations, because all the candidate matches have the same values for status on all individuals. Of those 96 synthetic observations, in 89 of the cases none of the 6 individuals in all the multiple candidates have the don status, and in 7 cases all 6 individuals have the don status.

In Table 1, in the article, we can compare the matched samples, for both grooms and brides who are matched, to the overall sample. As expected, among the matched samples the bride and grooms are less likely to be migrants. They are also less likely to be in a second marriage, although this number is generally small, at under 1%, even in the overall sample. Somewhat more notably, the percentage of individuals with high social status is lower in the linked sample. That said, the relative fraction of brides and grooms who are high status is similar across the overall sample, matched grooms, and matched brides. More precisely, the ratio of high status brides to grooms in the overall data is 1.22 (=7.79/6.38), vs. 1.19 in the linked grooms and 1.11 in the linked brides. This similarity, along with the similar coefficients in the regressions in columns (3) of Tables 2a and 3a and 2b and 3b, suggests that the differences between the unmatched and matches samples might not bias the coefficients of interest.

## A.4 Surnames

In this subsection, we study social mobility using the information contained in last names using an alternative method to studying social mobility prominently associated with Clark (2014). This technique looks at the relative representation of surnames within the elite, over time. That is, rather than rely on linked data, one regresses the share of individuals who possess some measure of high (or low) status among people with a particular surname on the share of individuals with that surname having high (or low) status in a previous generation. This avoids the need to link data. Clark (2014) argues that it also generates a better measure of an underlying measure of social status with high intergenerational persistence than the use of more standard measures such as income and occupation, which can be measured with considerable individual error. He generally finds estimates of between 0*.*7 and 0*.*8 across different societies, a pattern he names the “iron law of social mobility.”

To compare our estimates with his, we take the linked sample so that we can measure status for parents and children at consistent ages, to make the comparison as similar as possible. We take the “young” measure of status; that is, we use the measures of status taken at a groom or bride’s marriage and that of their parents, as in columns (4) through (6) of Table 3. For each observation, we calculate the *share* of individuals with a particular surname with status as a *don* or *doña* in each generation, and regress one on the other. The regressions are weighted by the number of individuals with each surname, to maximize comparison with the regression from the linked data.

The results from this regression are displayed in Table A4. The numbers for fathers are relatively similar to those from individual regressions, although they are still not in excess of 0*.*7*.* However, surname-based measures seem to give considerably lower estimates for the effect of mothers on social status on grooms and brides. This is not an artifact of naming conventions: For both grooms and brides, someone with the surname *Navarro* will have a father with this surname, and a mother with a different surname. Remember that in Spain, women do not change their name when they marry. The low values in the coefficients in Table A4, especially for women, contrast with the results in Solon (2018) that pseudo-links would tend to inflate the estimated intergenerational elasticity because they measure group, rather than individual, transmission of status.

Surname-based methods appear in this case to give little information as to the role of women in social mobility, even when they are feasible to use. Comparing columns (1) and (2), and similarly for columns (4) and (5), we can see the drop in the size of the coefficient, and an R-squared that is an order of magnitude smaller. In other words, columns (2) and (5) would seem to imply that mothers’ status have very little effect on their children’s status.

This result underscores the importance of having not only information on women’s status, but that this information is about the individual women. If we were to have the surname for all the women in our sample, but not their individual status, based on the results of Table A4, we would have wrongly concluded that they played little role in determining the status of their sons and daughters.

**A.5 Dons and Income**

In this section, we analyze the characteristics of individuals with status as *don* with a different source: The 1756 Ensenada *Catastro*.[[4]](#footnote-4) This source is described in greater detail in Espín-Sánchez et al. (2019). Briefly, this was a large-scale census of Castile, taken for the purposes of gathering information on income to institute a tax reform. For our purposes, the most important result of the Ensenada *Catastro* is that it provides an additional source with people identified by the honorific *don*, but also with the status identifiers more often used to study social mobility, in particular occupation and, unusually for this period, income.

In the following analysis, we restrict attention to males, although some women were listed in the *Catastro*. We also exclude individuals listed as “single” to better align the sample with the marriage records. This naturally excludes those in religious orders or the priesthood, who have the honorific *don* by custom. Finally, we exclude individuals with zero income, and use income from individuals’ primary job to measure income. We first plot the distribution of incomes for the remaining individuals for both *dons* and non-*dons* in Figure A1. Note that while the distribution of income is higher for *dons*, there is substantial overlap in the numbers. For reference, an unskilled worker has an annual income of 480 *reales.* The median income for *dons*, 1912.5 *reales*, is just over double that of non-*dons*, who have a median income of 900 *reales*. The difference in mean income is slightly less, with respective figures of 2643 and 1612.The 25th and 75th percentile for *dons* is 1095 and 3300, while the respective figures for non-*dons* are 600 and 1440. By any measure, *dons* have a higher income than non-*dons*, but there are clearly high-income non-*dons*, and conversely low income *dons*.

An alternative measure of social status is occupation. We classify the occupations into the HISCLASS occupational system of van Leeuwen and Maas (2012), as described in further detail in Espín-Sánchez et al. (2019). To ease visual interpretation, we collapse the categories into four broad occupational classifications: “High professional,” “Lower professional,” “Skilled workers,” and “Low skilled or unskilled workers.”[[5]](#footnote-5) We exclude the small number of farmers from the analysis. In Figure A2, we plot the proportion of individuals who are *dons* by occupational group, with the same sample restrictions as before. It is immediately apparent that *don* status is highly correlated with occupational status, with almost half of higher professionals with the honorific, and very few non-professional workers.

**Table A1. Names and Surnames equivalents.**

|  |
| --- |
| **Names** |
| ALEXANDRA | ALEJANDRA | JOAQUIM | JOAQUIN |
| ALEXANDRO | ALEJANDRO | JOSEF | JOSE |
| ALPHONSA | ALFONSA | JOSEPH | JOSE |
| ALPHONSO | ALFONSO | JOSEPHA | JOSEFA |
| ANNA | ANA | LAURIANO | LAUREANO |
| ATHANASIA | ATANASIA | LUISSA | LUISA |
| ATHANASIO | ATANASIO | MATHEA | MATEA |
| BARTHOLA | BARTOLA | MATHEO | MATEO |
| BARTHOLO | BARTOLO | MATHIAS | MATIAS |
| BARTHOLOME | BARTOLOME | MICHAELA | MICAELA |
| BARTHOLOMEA | BARTOLOMEA | NICACIO | NICASIO |
| BAPTISTA | BAUTISTA | PASQUAL | PASCUAL |
| BENTURA | VENTURA | PASQUALA | PASCUALA |
| BERNABEL | BERNABE | PHELIPA | FELIPA |
| BICENTE | VICENTE | PHELIPE | FELIPE |
| CALISTO | CALIXTO | RITTA | RITA |
| CATHALINA | CATALINA | ROSSA | ROSA |
| CHRISONTOMO | CRISTONTOMO | SICILIA | CECILIA |
| CHRISTOBAL | CRISTOBAL | SINFOROSSA | SINFOROSA |
| DEOGRACIAS | DESGRACIAS | SIZILIA | CECILIA |
| DOROTHEA | DOROTEA | TERESSA | TERESA |
| DOROTHEO | DOROTEO | THEODOSIO | TEODOSIO |
| ERMENEGILDO | HERMENEGILDO | THERESA | TERESA |
| FELIZ | FELIX | THOMAS | TOMAS |
| FELIZIANA | FELICIANA | THOMASA | TOMASA |
| GAVIERA  | JAVIERA | URSOLA | URSULA |
| GERTRUDES | GERTRUDIS | XAVIER | JAVIER |
| GINESSA | GINESA | XAVIERA | JAVIERA |
| JERTRUDIS | GERTRUDIS | ZACHARIAS | ZACARIAS |
| JIL | GIL | ZINFOROSA | SINFOROSA |
| JINES | GINESA | ZINFOROSSA | SINFOROSA |
| JINESA | GINESA | ESTASIA | ESTACIA |
| JINESSA | GINESA | LAURIANA | LAUREANA |
| JOACHIN | JOAQUIN | ROCHA | ROSA |
| JOACHINA | JOAQUINA |  |  |
| **Surnames** |
| AIALA | AYALA | PIÑARANDA | PEÑARANDA |
| ALBALAEJO | ALBADALEJO | PIZALO | PISALO |
| ALBAREZ | ALVAREZ | POBEA | POVEDA/POBEDA |
| ALBAREZ | ALVAREZ | PONZE | PONCE |
| ALVARRACIN | ALBARRACIN | POVEA | POVEDA/POBEDA |
| ALVORNOZ | ALBORNOZ | POVEDA | POBEDA |
| AMORES | AMOROS | QUADRADO | CUADRADO |
| ANDUXAR | ANDUJAR | QUIXADA | QUIJADA |
| AVELLAN | ABELLAN | RAMIRES | RAMIREZ |
| BACAS | VACAS | RAYGAL | RAIGAL |
| BALERO | VALERO | REBERTE | REVERTE |
| BAPTISTA  | BAUTISTA | RESAL | RESALTE |
| BEGARA | VERGARA | RESALT | RESALTE |
| BELASCO | VELASCO | REYNEL | REINEL |
| BELMUDEZ | BERMUDEZ | RIQUERO | RIQUEJO |
| BERDU | VERDU | RIQUEXO | RIQUEJO |
| BILA | VILA | ROXO | ROJO |
| BILLOTE | VILLOTE | RUVIRA | ROVIRA |
| BONZALBEZ | GONZALEZ | SABATER | ZAPATER |
| BURRUESO | BURRUEZO | SAMBUDIO | ZAMBUDIO |
| BUTIERREZ | GUTIERREZ | SAURIN | SAORIN |
| BUXALON | BUJALON | SELLES | CELIS |
| CALLEXAS | CALLEJAS | SELLIS | CELIS |
| CARABAJAL | CARVAJAL/CARBAJAL | SEREZO | CEREZO |
| CARAVAJAL | CARVAJAL/CARBAJAL | SOREJANO | SOBEJANO |
| CARBAXAL | CARVAJAL/CARBAJAL | SUABE | SUAVE |
| CARVAXAL | CARVAJAL/CARBAJAL | TEVAR | TEBAR |
| CAVALLER | CABALLER | THEBAR | TEBAR |
| CAVALLERO | CABALLERO | TRUXILLO | TRUJILLO |
| CERBERA | CERVERA | UJENA | UGENA |
| CERNA | SERNA | VAIADOLID | VALLADOLID |
| CLABEL | CLAVEL | VALEDIANA | VALERIANA |
| COSSIO | COSIO | VALERIANA | VALERIOLA |
| ECHAVARRIA | ECHEVARRIA | VEGARA | VERGARA |
| ERRERA | HERRERA | VELMUDEZ | BERMUDEZ |
| ESCARABAJAL | CARVAJAL/CARBAJAL | VELTRAN | BELTRAN |
| ESCARABAXAL | CARVAJAL/CARBAJAL | BERASTEGUI | VERASTEGUI |
| ESCARBAJAL | CARVAJAL/CARBAJAL | VERMUDEZ | BERMUDEZ |
| ESCARBAXAL | CARVAJAL/CARBAJAL | VOLARIN | BOLARIN |
| ESPINOSSA | ESPINOSA | XAREÑO | JAREÑO |
| EXEVARRIA | ECHEVARRIA | XEA | GEA |
| FAXARDO | FAJARDO | XIMENEZ | JIMENEZ |
| FUSTEL | FUSTER | YORCA | LLORCA/LORCA |
| GARBI | GARVI | YTA | HITA |
| GAVARRON | GABARRON | ZAPATEL | ZAPATER |
| GILABERTE | GILABERT | ZELDRAN | CELDRAN |
| GONZALBEZ | GONZALEZ | ZELIS | CELIS |
| GUARTE | HUARTE | ZERBERA | CERVERA |
| GUERTA | HUERTA | ZEREZUELA | CEREZUELA |
| GUETE | HUETE | ZERRANO | SERRANO |
| GUIRAO | GUIRADO | ZERVERA | CERVERA |
| GUSMAN | GUZMAN | ZEZAR | CESAR |
| HABRIL | ABRIL | ZORI | AZORI |
| HESPINOSA | ESPINOSA | TOLMOS | TORMOS |
| HESPINOSSA | ESPINOSA | MANCHON | MENCHON |
| HINOJOSSA | HINOJOSA | ARVANO | ALBANO |
| HOJOS | OJOS | MONTOIA | MONTOYA |
| HYTA | HITA | SERRAJO | CERROJO |
| IDALGO | HIDALGO | SERROJO | CERROJO |
| INOJOSA | HINOJOSA | GALLA | GAYA |
| INOJOSSA | HINOJOSA | PAMUS | PAMUZ |
| ITA | HITA | MATAZ | MATAS |
| JARREÑO | JAREÑO | BUZAN | PUZAN |
| JEA | GEA | PRAST | PRATS |
| JILABERTE | GILABERT | TISON | TIZON |
| JUAREZ | SUAREZ | CRISPIN | CRESPIN |
| JUSTAMANTE | BUSTAMANTE | ELCID | DEL CID |
| LAXARA, LA XARA, DE LA XARA | LAJARA, LA JARA, DE LA JARA | GALLAR | GAYAR |
| LIMAS | LIMA | PARMAÑES | PARMAÑEZ |
| LISSON | LISON | ELVIRA | ALVIRA |
| MACANAS | MACANAZ | MONUERA | MUNUERA |
| MANDONADO | MALDONADO | AGULLON | AGULLO |
| MANRRESSA | MANRRESA | HABELLAN | ABELLAN |
| MARQUES | MARQUEZ | BRIZUELA | BRISUELA |
| MATAIX | MATAIS | TENZA | ATIENZA |
| MESIA | MEJIA | VIÑABAT | BIÑABAT |
| MUNOZ | MUÑOZ | VIÑABATE | BIÑABAT |
| NABARRO | NAVARRO | BIÑABATE | BIÑABAT |
| NAXAR | NAJAR | HARNAU | ARNAU |
| OXOS | OJOS | ERGUETA | ELGUETA |
| PENALBA | PEÑALVA | MARZILLA | MARCILLA/MARSILLA |
| PENALVA | PEÑALVA | YER | AYER |

**Table A2: Percent *Don/ Doña* by Parents Status**

|  |
| --- |
| Father and Mother's Status |
|  | All | High/High | Low/Low | High/Low | Low/High |
| Groom | 6.38 | 89.7 | 0.955 | 71.2 | 30 |
| Bride | 7.79 | 92.7 | 1.26 | 67.3 | 68.2 |
| **Number** Grooms | 18175 | 859 | 16745 | 139 | 432 |
| Number Brides | 18175 | 852 | 16712 | 141 | 470 |

*Notes:* Value in the table is the percent of grooms and brides with *don*/*doña* status based on the status of their (own) parents, given in the column header as the status of the father and status of the mother. Number grooms/brides refers to the total number of grooms and brides respectively with the parental status combination shown above.

**Table A3: Percent Spouse *Don*/*Doña* Based on Parent’s and Own Statuses**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Parent Status: | All | High/High | Low/Low | High/Low | Low/High |
|  |  |  |  |  |  |
| PERCENT SPOUSE HIGH STATUS |
| Groom Don | 89.9 | 92.3 | 75.6 | 90.9 | 92.3 |
| Groom Not Don | 2.19 | 29.5 | 1.36 | 7.5 | 39 |
| Bride Doña | 73.6 | 96.4 | 34.7 | 96.7 | 50.4 |
| Bride Not Doña | 0.698 | 48.7 | 0.237 | 19.5 | 5.15 |
|  |  |  |  |  |  |
| NUMBER OF SPOUSES |
| Number Groom Don | 1160 | 771 | 160 | 99 | 130 |
| Number Groom Not Don | 17015 | 88 | 16585 | 40 | 302 |
| Number Bride Doña | 1417 | 738 | 348 | 93 | 238 |
| Number Bride Not Doña | 16758 | 121 | 16397 | 46 | 194 |

*Notes:* Value in the table is the percent of spouses with *don*/*doña* status based on the combination of the status of their (own) parents, given in the column header as the status of the father and status of the mother, and their own status. Number grooms/brides refers to the total number of grooms and brides respectively with the parental status combination shown above.

**Table A4: Surnames Regression**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Rel. share groom father | 0.655\*\*\* |  | 0.651\*\*\* |  |  |  |
|  | (0.044) |  | (0.043) |  |  |  |
|  |  |  |  |  |  |  |
| Rel. share groom mother |  | 0.242\*\*\* | 0.023 |  |  |  |
|  |  | (0.032) | (0.022) |  |  |  |
|  |  |  |  |  |  |  |
| Rel. share bride father |  |  |  | 0.367\*\*\* |  | 0.349\*\*\* |
|  |  |  |  | (0.012) |  | (0.016) |
|  |  |  |  |  |  |  |
| Rel. share bride mother |  |  |  |  | 0.222\*\*\* | 0.076\*\* |
|  |  |  |  |  | (0.041) | (0.028) |
|  |  |  |  |  |  |  |
| Parish FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Decade FE | Yes | Yes | Yes | Yes | Yes | Yes |
| R-squared | 0.617 | 0.071 | 0.617 | 0.326 | 0.089 | 0.333 |
| N | 3070 | 3070 | 3070 | 3931 | 3931 | 3931 |

*Notes:* OLS regression for status as a *don* or *doña* based on surname status. Dependent variable is the fraction of individuals with a surname with high status, and the independent variables are these fractions for the surnames of an individuals father and mother. Status is measured at the time of marriage for both generations. Regressions are weighted by the number of individuals in the younger generation with a particular surname. All standard errors are clustered at the parish level. \*, \*\*, and \*\*\* represent p < 0.1, p < 0.05, and p < 0.01 respectively.

Figure A1: Income Distribution of Individuals in 1756 Ensenada



*Notes*: Plots of kernel densities of the wealth of individuals described as a *don* or not as a *don* in the 1756 Ensenada *Catastro* described in Appendix A.5. The sample is restricted to married males who report a positive income. The density uses an Epanechinikov kernel and Stata-default bandwidth selection.

Figure A2: Percent Don by Social Class, 1756 Ensenada



*Notes*: Percent *don* in the 1756 Ensenada *Catastro,* separately by HISCLASS as described in Section A.5. Sample is restricted to married males. For simplicity, the HISCLASSes are compressed into four categories. “High professional” includes HISCLASSes 1 and 2 (originally “higher managers” and “higher professionals”), “Lower professional” HISCLASSes 3, 4, and 5 (“lower managers”, “lower professionals”, and “lower clerical and sales”), “Skilled” HISCLASS 7 (“skilled workers”), and “Low/unskilled” HISCLASSes 9 and 11 (“lower skilled workers” and “unskilled workers”). Job classifications are described in greater detail in Espín-Sánchez et al. (2019). HISCLASS classification system based on that of van Leeuwen and Maas (2012).

1. This parsing is non-trivial, with the number of names for individuals varying and some name, such as Garcia, potentially being either a given name or a surname. We also exclude from the set of names special constructions such as “De La Cruz” or “Del Carmen,” which are not consistently recorded for individuals. [↑](#footnote-ref-1)
2. In our experience, using *Soundex* with Spanish names did not perform as well. This is not surprising, because *Soundex*is built for Germanic languages, like English and German, which are phonetically based on consonants. Latin languages, like Spanish and Italian, are phonetically based on vowels. [↑](#footnote-ref-2)
3. For the regressions using only the status of the parents, but taken at two different points in the lifecycle, we would only need that the status of both the father and the mother of the groom to be the same in all candidates. In this case there are only 4 combinations. [↑](#footnote-ref-3)
4. Although the *Catastro* began in 1749, the individual-level responses from Murcia we use date from 1756. [↑](#footnote-ref-4)
5. These classifications include, respectively, HISCLASSes 1 and 2, HISCLASSes 3, 4, and 5, HISCLASS 7, and HISCLASSES 9 and 11. [↑](#footnote-ref-5)