Appendix for Online Publication

A Modeling Selection

A Copt chooses consumption (c) and religious affiliation ($\kappa = 1$ if he remains Coptic Christian and $\kappa = 0$ if he converts to Islam) in order to maximize:

$$U = u(c) - (1-\kappa)r \tag{A.1}$$

$$\text{s.t.} \qquad c \leq y - \tau\kappa$$

where $u'(.)>0$ and $u''(.)<0$. It follows that a Copt converts if and only if:

$$u(y) - u(y-\tau) \geq r \tag{A.2}$$

Proof of Proposition 1 Let $y^*(\tau; r)$ be the threshold level of income at which a Copt is indifferent about conversion to Islam at a given level of religiosity, that is defined by $u(y^*)-u(y^*-\tau) = r$. Copts' population share is given by $pop_{copt} = \int_{y^*}^{\infty} f(y)dy = 1-F(y^*)$. It follows that:

$$\frac{\partial pop_{copt}}{\partial \tau} = -f(y^*) \times \frac{\partial y^*}{\partial \tau} = -f(y^*) \times \frac{-u'(y^*-\tau)}{u'(y^*) - u'(y^*-\tau)} < 0$$

because $u'(.)>0$ and $u''(.)<0$.

Proof of Proposition 2

$$\frac{\partial}{\partial \tau} E(y \mid Copt) = \frac{\partial}{\partial \tau} E(y \mid y>y^*) = \frac{\partial}{\partial \tau}\left(\frac{\int_{y^*}^{\infty} y f(y)dy}{1-F(y^*)}\right) = \frac{f(y^*)\frac{\partial y^*}{\partial \tau}}{1-F(y^*)}\left[E(y \mid y>y^*) - y^*\right] > 0$$

$$\frac{\partial}{\partial \tau} E(y \mid Muslim) = \frac{\partial}{\partial \tau} E(y \mid y \leq y^*) = \frac{\partial}{\partial \tau}\left(\frac{\int_{0}^{y^*} y f(y)dy}{F(y^*)}\right) = \frac{f(y^*)\frac{\partial y^*}{\partial \tau}}{F(y^*)}\left[y^* - E(y \mid y \leq y^*)\right] > 0$$

Define the Coptic-Muslim income gap as $\Delta \equiv E(y \mid Copt) - E(y \mid Muslim) = E(y \mid y>y^*) - E(y \mid y \leq y^*)$. It follows that:

$$\frac{\partial \Delta}{\partial \tau} = f(y^*)\frac{\partial y^*}{\partial \tau}\left[\frac{1}{1-F(y^*)}\left(E(y \mid y>y^*) - y^*\right) - \frac{1}{F(y^*)}\left(y^* - E(y \mid y \leq y^*)\right)\right]$$

which could be either positive or negative depending on the income distribution. The second statement of proposition 2 follows from Jewitt (2004) and $\frac{\partial y^*}{\partial \tau} > 0$.

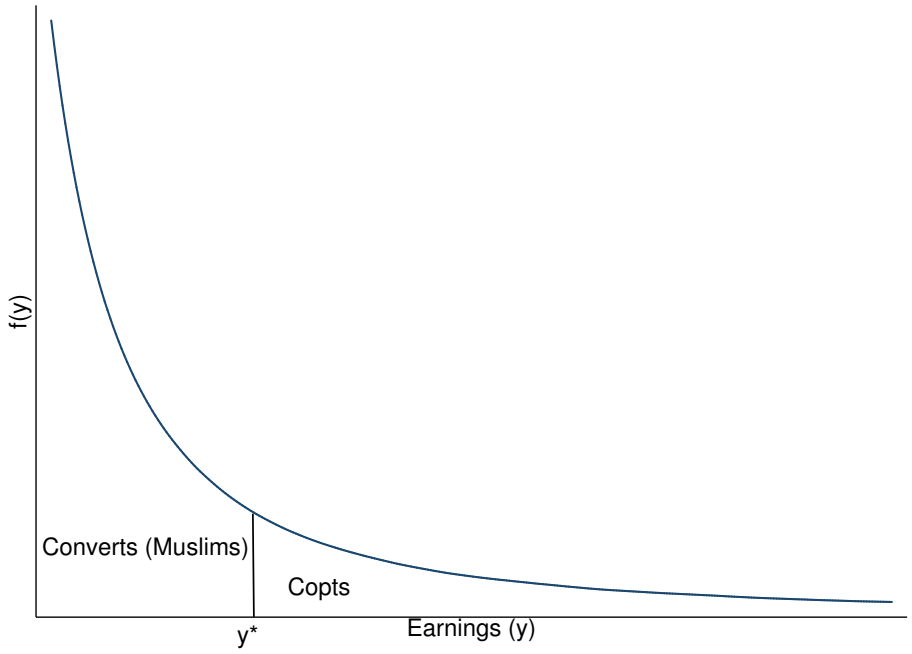

Figure A.1: **The Marginal Effects of the Poll Tax in a Static Environment**

Source: See text.

B Data Sources

B.1 The 1848 and 1868 Population Census Samples

The 1848 and 1868 population censuses are among the earliest pre-Colonial population censuses from any non-Western country to enumerate all household members, including females, children, and slaves. They contain information on a wide range of variables including place of residence, name, gender, age, relationship to household head, nationality, ethnicity, free/slave status, religion, place of origin, occupation, school enrollment (for male children in cities), and dwelling ownership and type (in cities only). I digitized a nationally representative sample of each census (about 80,000 individuals in each sample) from the original Arabic manuscripts at the National Archives of Egypt. The sampling rate is 8-10 percent in the two major cities of Cairo and Alexandria and 1 percent in all the other provinces. Details about the census samples are in Saleh (2013).

For the purpose of this article, I pooled the census samples from both years and restricted the analysis to Egyptian local free Coptic and Muslim employed men of a rural district of origin who are at least 15 years of age with non-missing values on age, religion,

occupation, and district of origin. The sample restrictions aim at limiting the sample to the likely descendants of Egypt's pre-641 population who either converted to Islam or remained Coptic Christian:

1. I restricted the sample to "Egyptians," or individuals who are listed in the census manuscripts as *dakhil al-hukuma* (under the control of the Egyptian government). This excludes foreigners including Turks, Levantines, Armenians, Ottoman Greeks, North Africans, Yemenis, people of the Arab peninsula, Asians, Americans, and Europeans.

2. I further restricted the sample to "locals." This excludes "Egyptians" who are not originally from the Nile Delta or the Nile Valley (north of Nubia), namely, Bedouins, Nubians, Sudanese, and Abyssinians.[1]

3. I restricted the sample to individuals of a rural district of origin. District of origin in the census manuscripts refers to the origin of an individual's family (not necessarily the place of birth). This excludes individuals whose family's place of origin is either major cities (Cairo, Alexandria, Suez, Rosetta, and *Damietta*) or deserts (*Arish* in the Sinai peninsula, *Qusayr* on the Red Sea, and the five oases of the Western desert). This mitigates the potential cross-district movement of an individual's ancestors in 641-1868 under the presumption that most migration was directed towards cities.

In order to construct the occupational outcomes, I first manually translated and coded all the occupational titles in the census manuscripts following the coding of the Historical International Standard Classification of Occupations (HISCO) (Saleh, 2015). I then classified each occupational code into one of nine categories (the largest occupational titles within each category are mentioned in parentheses): (1) professionals (nurse, engineer, physician), (2) high-level bureaucrats (farm supervisor, government official), (3) mid-low bureaucrats (scribe, messenger, bookkeeper, stock clerk), (4) judiciary, military, and police (policeman, military officer, judge, legal delegate), (5) clergy and rural elites (village headman, minister of religion), (6) merchants (working proprietor, auctioneer), (7) artisans (weaver, carpenter, baker, grain miller, dyer, butcher), (8) farmers (farmer,

[1] Bedouins (*'orban*) in 1848-1868 are likely (descendants of) Arab tribes that arrived to Egypt during the eighteenth century.

farm worker, livestock worker, gardener, fisherman), and (9) unskilled workers. The first six categories comprise the three white-collar indicators. The full list of occupational titles that fall under each category is available upon request.

Table B.1 shows the summary statistics of the 1848-1868 sample for which I observe the poll tax (statistics on occupational outcomes in the 1848-1868 full sample are in Table 2 in the article). Muslims are more likely than Copts to be of a district of origin that faced a higher-than-average poll tax in 641-1100. These districts had a larger urban population in 300, were less likely to lie on the Holy Family legendary path, and were more likely to have an *autopract* estate and a Byzantine garrison in 600. They did not differ though in the share of villages that had at least one Coptic monastery in 1200.

Table B.1: **The 1848-1868 Population Census Samples: Summary Statistics**

	Individual-Level			District-Level
	Copts	Muslims	Total	Total
District's Copts' population share in 1848-1868	0.22	0.14	0.15	0.13
	(0.08)	(0.10)	(0.10)	(0.09)
District's average poll tax in 641-1100 (Dinars)	1.19	1.27	1.26	1.27
	(0.14)	(0.13)	(0.13)	(0.13)
=1 if average poll tax high in district in 641-1100	0.42	0.71	0.67	0.73
	(0.49)	(0.45)	(0.47)	(0.47)
Log(urban population) in 300	9.86	10.33	10.26	10.41
	(0.97)	(0.84)	(0.87)	(0.85)
=1 if district on Holy Family route in 400	0.23	0.20	0.21	0.27
	(0.42)	(0.40)	(0.41)	(0.47)
=1 if *Autopract* estates in district in 600	0.88	0.94	0.93	0.91
	(0.33)	(0.23)	(0.25)	(0.30)
=1 if Byzantine garrison in district in 600	0.44	0.56	0.55	0.55
	(0.50)	(0.50)	(0.50)	(0.52)
Observations	396	2286	2682	11

Note: Standard deviations are in parentheses.
Source: The 1848-1868 census samples combined with other data sources. The sample is restricted to individuals for whom I observe the poll tax in 641-1100 in their district of origin. See section B in the online appendix for details.

B.2 Christian Churches and Monasteries in 1200 and 1500

I constructed a village-level dataset on the number of Christian (Coptic and non-Coptic) churches and monasteries in 1200 and 1500 from two medieval sources, (1) Abul-Makarim (1200) who lists Christian churches and monasteries in Egypt at the end of the twelfth

century, and (2) Al-Maqrizi (1500) who provides a similar list in the fifteenth century.[2] Both sources list locations at the village level (smaller than districts). I matched villages in both sources to villages in 1848-1868 using Ramzi (1954).

I employ this dataset to construct three variables:

1. The share of villages in Egypt with at least one Christian church or monastery in 1200 and 1500, that I use as an estimate of Egypt's non-Muslims' population share in Figure 3. This holds under the presumption that a village with at least one Christian church or monastery is 100 percent non-Muslim and that a village without any Christian church or monastery is 100 percent Muslim.

2. A dummy variable that takes value 1 if a village had at least one Coptic church or monastery in 1200 and 1500, that I use as a proxy for Copts' population share, the dependent variable in equation (1).

In both measures, I restrict the universe of villages to those that existed as of 1477 using Ramzi (1954)'s list of pre-1477 and post-1477 villages; itself based on a cadastral survey in 1477 that was recorded in Ibn-Al-Jay'an (1477). This restriction aims at mitigating the concern about the possible emergence of new villages between 1200 (or 1500) and 1848-1868.

B.3 Religion and Occupations in 641-969

B.3.1 Data Construction

Arabic papyri, most of which were discovered since the late nineteenth century in Egypt's dry-climate Nile Valley, remain a mostly unpublished source of information on the micro-level characteristics of the medieval populations of the Middle East under the early Arab Caliphate in 641-969. An ongoing research project called the Arabic Papyrology Database (henceforth, APD) attempts at digitizing of all the *published* Arabic papyri.[3] There

[2]There are two versions of Abul-Makarim's book. The first is *The Churches and Monasteries of Egypt and Some Neighboring Countries*, edited by Evetts in 1895 and (wrongly) attributed to "Abu-Saleh, the Armenian." This version lists Christian churches and monasteries in the Nile Valley only. The second version that I use in the article is a two-volume book edited by Anba-Samuel in 1984. The first volume includes the missing part on the Nile Delta, whereas the second is a re-publication of Evetts' 1895 book on the Nile Valley. The book is now believed to belong to the twelfth-century Coptic chronicler, Abul-Makarim.

[3]Out of more than 150,000 Arabic papyri that were ever *discovered*, only 2,500 documents (less than 2 percent) were *published* since 1900. The APD, which was launched in 2004, has, as of April 2015, digitized 2,068 documents or about 83 percent of the published papyri.

are various types of documents in the APD, namely, (1) protocols, (2) legal texts (e.g. marriage and sale contracts), (3) administrative texts (e.g. official letters, lists, and accounts), (4) private texts (e.g. private and business letters), and (5) literary texts.[4]

I used all the papyri that has been digitized by the APD in order to construct an individual-level dataset on occupational titles and religious affiliation, where I inferred a worker's religion from his name (converts adopted an Arabic name upon conversion). I included in the sample every male mentioned in any APD document with a non-missing name and occupation.

A few notes on the APD sample are in order:

1. Because the occupational title of "landholding farmer" is almost never mentioned in the APD (a landless farmer, or *'agir*, is explicitly mentioned though), I infer if a male with a non-missing name in the APD land tax registers and receipts is a landholding farmer if he pays a positive land tax, or other money tax on land; specifically, on pasture, grazing land, palms, gardens, and other uses of land. The vast majority of farmers in Egypt were landholding farmers who held usufruct rights on land and paid a land tax.

2. Because my objective is to compare the occupational outcomes of Copts versus converts (i.e. Egyptian Muslims), ideally I want to exclude non-Egyptian Muslims (Arabs and Turks) from the APD sample. For this purpose, I exclude individuals in top government posts, namely, Caliphs, viceroys, and top government administrators, who were certainly Arabs in 641-833 and either Arabs or Turks in 833-969. I am not able though to identify all Arabs since converts adopted Arabic names.

3. I restrict the APD sample to the papyri dated between 641 and 969 because there are fewer papyri after 969 as paper increasingly replaced papyrus.

4. I infer religious affiliation from names only, and not from any other contextual information (e.g. occupation), in order to not introduce non-random measurement error. Furthermore, I only include males with names that are distinctly Muslim or Christian based on the names list in the 1848-1868 census samples and on papyrologists'

[4]Papyrus is a material of writing that was prevalent in Egypt until the administration increasingly shifted to using paper around 969. The APD documents are written on other writing materials besides papyrus, including paper, ostracon, woodtable, waxtable, stone, parchment, leather, bone, and textile. All these documents are included in the APD sample.

comments on the papyri.

5. Similarly, in order to not introduce non-random error measurement in observing occupations, I only use explicit information on occupational titles in the papyri instead of inferring occupations from contextual information (with the exception of inferring the occupational title of "landholding farmer" from paying a positive land tax).

These procedures result in a final sample of 402 males with religious affiliation and occupational title in 641-969. This is the APD sample that I use in Table 2.

B.3.2 Addressing the Concerns about the APD Sample

There are at least three concerns about the APD sample. First, the sample may not be nationally representative: (a) It is not random on location; 34 percent is from the Nile Valley, 65 percent is from unknown locations (in Egypt), and less than 1 percent is from the (more humid) Nile Delta. (b) It is (likely) not random on SES, because it may over-represent males of high SES. Second, there is measurement error in inferring religious identity from first names (Macdonald, 1999): (a) The same individual may have double names, "non-Muslim" and "Muslim," that are not recorded in the papyrus. (b) A Muslim person may have a "non-Muslim" name and vice versa. (c) I cannot tell if a non-Muslim name is Coptic, non-Coptic Christian, or Jewish. (d) I cannot tell if a Muslim name is (a descendant of) a convert or an Arab/Turkish settler. Third, there is measurement error in observing occupational titles, because I infer if an individual is a landholding farmer from the incidence of paying a non-zero land tax.

While it is not possible to completely rule out these concerns, there are counter-arguments that mitigate each of them:

1. On the first concern: (a) Even if the APD sample is mostly from the Nile Valley, the 1848-1868 population census samples reveal that the Nile Valley has a higher Copts' population share and a smaller, yet still positive and statistically significant, Coptic-Muslim SES difference compared to the Nile Delta. Hence, the APD sample, if anything, *overestimates* the true Copts' population share and *underestimates* the true Coptic-Muslim SES gap. (b) 72 percent of the APD sample is from administrative documents, namely, lists (54 percent) and receipts of payment (18 percent).

These are arguably *neutral* documents in which every individual, regardless of their SES, has an equal chance of appearance. The findings in Table 2 do not change if I exclude the remaining 28 percent of the sample that come from contracts (sale, lease, marriage, divorce, and written obligation contracts) and private and business letters. Even if the APD sample is not representative of the occupational distribution of each religious group, it may be still representative of the Coptic-Muslim *difference* in occupational outcomes under the assumption that selection-on-SES in the APD sample is the same within each group.

2. On the second concern: (a) I do not observe in the APD sample any worker with double first names. This is not to say that the phenomenon did not exist but it suggests that it was rare. (b) A papyrological list of converts in 700-900 indicates that converts adopted Arabic names and became clients of Arab patrons upon conversion (Morimoto, 1981, p. 131). Of course, some converts may have kept their non-Muslim names but it is plausible that this was rare. This is actually consistent with the evidence in Rapoport (2004) on converts in *Fayum*. (c) "Non-Muslims" in the APD sample are mostly Copts because Copts constituted 96 percent of Christians in 1200 (based on the dataset on Christian churches and monasteries), and non-Coptic Christians and Jews rarely settled in the Nile Valley, where most of the APD sample is from.[5] (d) Even I misidentify Arab/Turkish settlers as (Egyptian) converts, this will likely *overestimate* the share of high-SES workers among converts since Arab/Turkish settlers in 641-969 were more likely to be in high-SES jobs (as the ruling elite). This would underestimate the true Coptic-Muslim SES differential.

3. On the third concern, the share of workers in the APD sample whose occupation is imputed as a landholding farmer within each religious group is similar to the share of landholding farmers in the 1848-1868 population census samples.

[5] According to Mikhail (2004, p. 134), there is no evidence on the presence of non-Coptic Christians (*Melkites*') in the Nile Valley in the post-Conquest period. In 1200, 91 percent of non-Coptic Christian (*Melkite* and Armenian) churches and monasteries were in the Nile Delta and major cities (Cairo, Alexandria, and *Damietta*).

Table B.2: **Individual-Level Annual Poll Tax Payments (Dinars) in 641-1100 Summary Statistics by *Kura***

Name	Period	N	% No Tax	Median	Mean	Std. Dev.	Min	Max
Ihnas	701-900	10	0	1.38	1.35	0.5	0.88	2.25
Ashmunayn	731-1100	77	0	0.96	1.36	1.18	0.17	6.71
Fayum	641-1005	7	0	0.99	1.34	0.81	0.25	2.67
Qahqawa	703-733	314	46.5	1	1.07	1.27	0	5

Source: Greek and Arabic poll tax registers and receipts in 641-1100. See section B in the online appendix.

B.4 Poll Tax Papyri in 641-1100

I constructed an individual-level dataset on annual nominal poll tax payments (in dinars) between 641 and 1100 ($N = 552$) from Greek and Arabic papyri. The location (*kura*) of the tax papyri is only available for a smaller sample ($N = 408$) in four *kuras* in the Nile Valley. Finally, Greek poll tax registers from three sub-districts in the *kura* of *Qahqawa* in 703-733 contain information for each landholder on both the poll and total land tax payments, which is the restricted sample ($N = 230$) that I use in Table D.1. The papyri come from two sources:

1. Fragments of Greek poll tax registers of the *kura* of *Qahqawa* (pre-641, *Aphrodito*) in 703-733 in Morimoto (1981, pp. 67-79).

2. Fragments of Arabic poll tax registers and receipts in the APD for the *kuras* of *Fayum* (pre-641, *Arsinoe*) in 641-1005, *Ihnas* (pre-641, *Herakleopolis*) in 701-900, and *Ashmunayn* (pre-641, *Hermopolis*) in 731-1100.[6]

Table B.2 shows the sample size, the chronological range, and the summary statistics of the individual-level annual poll tax payments in the dataset by *kura*. I also depict the distribution of poll tax payments by *kura* in Figure B.1.

B.5 Arab Settlement in 700-969 and Control Variables

1. Data on Arab settlement in 700-969 are constructed from Al-Barri (1992) who traces the destination of the Arab tribes that settled in Egypt between 700 and 969 using information from Arabic medieval narratives. Destination of each tribe is

[6]There are other Greek poll tax registers and Coptic poll tax receipts in 641-800 that I did not employ in the analysis because they are not digitized.

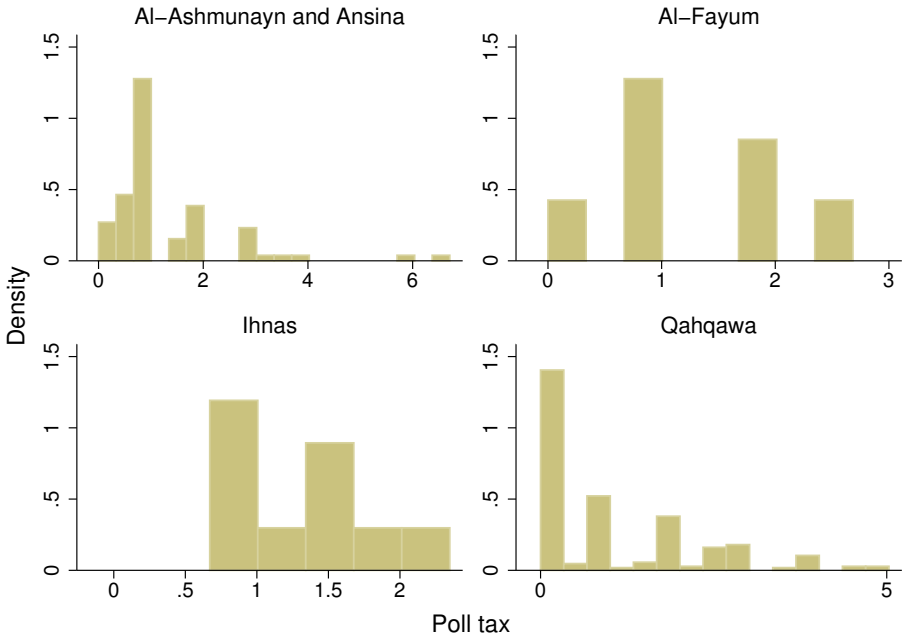

Figure B.1: **Distribution of Poll Tax Payments (Dinar) by *Kura***

Source: Greek and Arabic poll tax registers and receipts in 641-1100.

usually recorded at the *kura* level. I restrict the analysis to permanent settlement by excluding seasonal migration (*irtiba'*) between 641 and 700.[7]

2. Data on urban population in 300 are constructed from Wilson (2011, pp. 185-187). These are estimates of the population of Greek cities (metropolis) and the capital of each *nome* (Egypt's administrative units in the Roman period) around 300. Wilson assigns a fixed population for the capitals of all other *nomes* that are not included in his estimates.

3. Data on the legendary route of the Holy Family are constructed from Anba-Bishoy (1999) and Gabra (2001), where I recorded information at the village level. The route is based on a book attributed to Theophilus, Alexandria's Patriarch in 385-412.[8] The date of the book is debated though with some scholars dating it to the fifteenth century, which raises the concern that the path may reflect post-641 Coptic traditions. There are two points that mitigate this concern: (a) There is evidence

[7] *Irtiba'* is Arabs' policy of seasonal migration during the Spring to the Egyptian countryside for grazing their animals. The policy was enforced between 641 and 700 where Egyptians were required to provide Arab tribes with food and shelter (Al-Barri, 1992, pp. 56-60).

[8] The legendary flight of the Holy Family to Egypt is mentioned in Matthew 13.

that local beliefs on the journey of the Holy Family, although not the path itself, emerged as early as the Roman period, and (b) I use the Holy Family legendary route as a control variable only and I show the results both with and without it.

4. Data on *autopract* estates in 600 are constructed from Hardy (1931), where I recorded information at the *nome* level. I restrict the data to *nomes* in the Nile Valley, since the papyri were less likely to survive in the Nile Delta.[9]

5. Data on Byzantine garrisons in 600 are constructed from Maspero (1912), where locations are recorded at the village level.

B.6 Mapping *Nomes* and *Kuras* into Districts in 1848-1868

Nomes were Egypt's administrative units before 641. Egypt was divided into about 40 *nomes*, 20 in the Nile Valley and 20 in the Nile Delta. After the Arab Conquest, the country was divided into about 42 *kuras*. In 1848-1868, there were 14 provinces in the Nile Valley and Nile Delta with 76 districts. *Nomes* and *kuras* are smaller in surface area than provinces but larger than districts.

I mapped *nomes* and *kuras* into districts in 1848-1868 using the 1897 census administrative division, since the 1848 and 1868 censuses do not contain an official list of villages under each district. In the absence of information on the boundaries of *nomes*, I mapped *nomes* into districts in 1848-1868 as follows. I first compiled the full list of *nomes* from Trismegistos; a specialized portal in Egyptian papyri documents between 800 BC and 800 CE. Second, I located the major towns of each *nome* using the maps of *nomes* in Trismegistos. Finally, I matched each district in 1848-1868 to the closest *nome* based on the proximity of the district's chef-lieu to the *nome*'s major towns according to Google maps. Similarly, as we do not know the boundaries of *kuras*, I mapped *kuras* into districts in 1848-1868 using a similar routine. I first compile the full list of *kuras* from Tousson (1926). Second, I located the major towns of each *kura* using Ramzi (1954) and Tousson (1926). Finally, I matched each district in 1848-1868 to the closest *kura* based on the proximity of the district's chef-lieu to the *kura*'s major towns according to Google maps.

[9]Later scholarship has questioned the *autopragia* status of certain *nomes* such as *Aphrodito*.

C Concerns about the Econometric Evidence

C.1 Representativeness of the Papyrological Poll Tax Records

The tax papyri may not be representative of Egypt. To mitigate this issue, I examine the impact of Arab settlement (that is observed for all 42 *kuras*, 76 districts) on Copts' population share and the Coptic-Muslim SES gap. The rationale is that Arab settlement should have an effect on outcomes via its impact on the enforcement of the poll tax. Using the full sample with the Arab settlement variable generates similar results to the main findings (see Tables C.1 and C.2).

C.2 Error in Observing the Poll Tax

The poll tax in the papyri may be observed with error. I address this concern as follows (the results are in Tables C.3 and C.4): (1) Instead of using the average tax, I use an indicator variable that takes the value of one if the median poll tax was high. (2) I drop two *kuras* where the number of taxpayers in the papyri is less than 10. (3) I drop the *kura* of *Qahqawa*, where the papyri come from an earlier period. (4) Instead of using a dummy variable indicating districts with a higher-than-average poll tax, I use the actual average and median poll taxes in dinars. The results are qualitatively similar.

C.3 Tax-Induced Migration across Districts in 641-1868

A potential source of endogeneity of the poll tax is people's movement across districts in response to cross-district differences in the poll tax. Historical evidence suggests that this is unlikely though. The state imposed restrictions on migration in rural Egypt since (at least) 641, and these restrictions lasted until 1857. Individual obligations such as land cultivation, taxes, and corvée in public works, were decided based on village of residence, and hence people were not allowed to leave their villages without a state permit. Morimoto (1981, pp. 113-24) describes the "fugitives" between 641 and 969, those who fled their villages in order to escape their land assignment and tax quota, and the state policies to identify fugitives and send them back to their villages. Historians of Ottoman Egypt document the same phenomenon (Mahmoud, 2009; Cuno, 1992).

Table C.1: **Arab Settlement in 700-969 and Copts' Population Share in 1200, 1500, and 1848-1868**

Dependent Variable: 1200 and 1500: = 1 if At Least One Coptic Church or Monastery in Village
1848-1868: =1 if Individual is Coptic Christian

(a) Ordinary Least Squares

	1200			1500			1848-1868		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if Arab settlement in district in 700-969	-0.082**	-0.074**	-0.107**	-0.034*	-0.037**	-0.163***	-0.071**	-0.074***	-0.067*
	(0.031)	(0.030)	(0.045)	(0.018)	(0.018)	(0.047)	(0.029)	(0.028)	(0.035)
Controls urbanization and religiosity?	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Other controls?	No	No	Yes	No	No	Yes	No	No	Yes
Clusters (districts)	75	75	39	75	75	39	76	76	35
Obs (villages/individuals)	1817	1817	560	1817	1817	560	16641	16641	6792
R^2	0.009	0.010	0.015	0.008	0.009	0.079	0.017	0.017	0.019

(b) Two-Stage Least Squares

	(1)	(2)	(3)	(4)	(5)
	1200	1500	1848-1868	1200-1500 First Stage	1848-1868 First Stage
=1 if Arab settlement in district in 700-969	-0.636	-2.194	-0.234***		
	(1.543)	(5.433)	(0.071)		
District's distance to *Arish* (km)				-0.000	-0.001***
				(0.001)	(0.000)
Clusters (districts)	75	75	76	75	76
Obs (villages/individuals)	1817	1817	16641	1817	16641
KP Wald F-stat				0.154	9.972

Notes: Robust standard errors clustered at the district level are in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$. A constant is included in all regressions.
Source: 1200 and 1500: Dataset on Christian churches and monasteries in 1200 and 1500; 1848-1868: The population census samples. Data on regressors are from multiple data sources. See section B for details.

Table C.2: **Arab Settlement in 700-969 and the Coptic-Muslim Occupational Differences in 1848-1868**

Dependent Variable Indicated on Top of Each Column

(a) Ordinary Least Squares

	=1 if White-Collar3		=1 if Artisan		=1 if Farmer		=1 if Unskilled	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Copt	0.056*	-0.108	0.212***	-0.457*	-0.158***	0.201	-0.110***	0.365**
	(0.032)	(0.339)	(0.033)	(0.249)	(0.049)	(0.459)	(0.040)	(0.177)
Copt * Arab settlement	0.151**	0.045	-0.059	-0.136**	-0.037	0.134	-0.055	-0.043
	(0.070)	(0.103)	(0.058)	(0.060)	(0.093)	(0.135)	(0.046)	(0.048)
Copt * Controls?	No	Yes	No	Yes	No	Yes	No	Yes
District FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters (districts)	76	35	76	35	76	35	76	35
Obs (individuals)	16641	6792	16641	6792	16641	6792	16641	6792
R^2	0.033	0.027	0.044	0.044	0.109	0.114	0.065	0.065

(b) Two-Stage Least Squares

	(1) =1 if White-Collar3	(2) =1 if Artisan	(3) =1 if Farmer	(4) =1 if Unskilled	(5) First-Stage
Copt	-0.104	0.137**	0.036	-0.070	1.554***
	(0.093)	(0.061)	(0.119)	(0.051)	(0.207)
Copt * Arab settlement	0.473***	0.091	-0.429***	-0.135	
	(0.134)	(0.132)	(0.160)	(0.087)	
Copt * Distance to *Arish*					-0.002***
					(0.001)
Clusters (districts)	76	76	76	76	76
Obs (individuals)	16641	16641	16641	16641	16641
KP Wald F-stat					19.707

Notes: Standard errors clustered at the district of origin level are in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.
Source: The 1848-1868 population census samples combined with multiple data sources. See section B in the online appendix for details.

Table C.3: **The Poll Tax and Copts' Population Share in 1848-1868**
Measurement Error in Poll Tax - OLS
Dependent Variable =1 if Coptic Christian

	Full Sample		Excluding *Kuras* of *Ihnas* and *Fayum*		Excluding *Kura* of *Qahqawa*		Full Sample		Full Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
=1 if med. poll tax high	-0.128**	-0.096*			-0.069	-0.057				
	(0.051)	(0.049)			(0.046)	(0.033)				
=1 if av. poll tax high			-0.070	-0.128						
			(0.058)	(0.127)						
Med. poll tax (dinars)									-0.315**	-0.241*
									(0.127)	(0.121)
Av. poll tax (dinars)							-0.573***	-2.694**		
							(0.168)	(1.074)		
Control for urbanization?	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Other controls?	No	Yes	No	No	No	No	No	Yes	No	Yes
Clusters (districts)	11	11	5	5	8	8	11	11	11	11
Obs (individuals)	2682	2682	1300	1300	1790	1790	2682	2682	2682	2682
R^2	0.021	0.064	0.006	0.006	0.012	0.033	0.045	0.061	0.019	0.064

Notes: Standard errors clustered at the district of origin level are in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$. A constant is included in all regressions.
Source: The 1848-1868 population census samples combined with multiple data sources. See section B for details.

Table C.4: **The Poll Tax and the Coptic-Muslim Occupational Differences in 1848-1868**
Measurement Error in Poll Tax - OLS
Dependent Variable =1 if White-Collar3

	=1 if Median Tax High		=1 if Average Tax High: Excluding *Kuras* of *Ihnas* and *Fayum*	=1 if Median Tax High: Excluding *Kura* of *Qahqawa*	Average Tax in Dinars		Median Tax in Dinars	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Copt	0.040	-0.951***	0.010	0.082**	-0.349*	-0.334	-0.430	-1.366***
	(0.025)	(0.256)	(0.026)	(0.030)	(0.183)	(0.376)	(0.270)	(0.218)
Copt * Poll tax	0.200*	0.153***	0.042	0.158	0.337*	2.352***	0.477	0.374***
	(0.106)	(0.038)	(0.030)	(0.109)	(0.157)	(0.578)	(0.266)	(0.097)
District FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Copt * Controls?	No	Yes	No	No	No	Yes	No	Yes
Clusters (districts)	11	11	5	8	11	11	11	11
Obs (individuals)	2682	2682	1300	1790	2682	2682	2682	2682
R^2	0.021	0.025	0.014	0.027	0.021	0.024	0.021	0.025

Notes: Standard errors clustered at the district of origin level are in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.
Source: The 1848-1868 population census samples combined with multiple data sources. See section B for details.

D Additional Results

Table D.1: **Poll Tax per Dinar among Landholders in 703-733**
Dependent Variable: Annual Poll Tax Per Person ÷ Annual Land Tax Per Person

	(1)	(2)
Land tax	-0.041***	-0.036***
	(0.015)	(0.013)
Sub-district FE?	No	Yes
Observations	230	230
Adjusted R^2	0.014	0.112

Notes: Robust standard errors are in parentheses. $*p < 0.1, **p < 0.05, ***p < 0.01$. A constant term is included in the first regression.
Source: Greek papyrological tax registers of three sub-districts in the *kura* of *Qahqawa* (pre-641, *Aphrodito*) in 703-733 (Morimoto, 1981, pp. 67-79, 85-87). The sample is restricted to landholders.

Table D.2: **Arab Settlement and Poll Tax Rates**
Dependent Variable: Poll Tax Rate per Person (Dinars) in 641-1100

	OLS		IV Second Stage	IV First Stage
	(1)	(2)	(3)	(4)
=1 if Arab settlement in *kura* in 700-969	0.290***	0.214***	0.285***	
	(0.004)	(0.000)	(0.005)	
Kura's Distance to *Arish* (km)				-0.009*
				(0.003)
Controls urbanization and religiosity?	No	Yes	No	No
Obs (individuals)	408	408	408	408
Clusters (*kuras*)	4	4	4	4
KP Wald F-stat				8.532

Notes: Robust standard errors clustered at the *kura* level are in parentheses. A constant term is included in all regressions. $*p < 0.10, **p < 0.05, ***p < 0.01$.
Source: Poll tax registers and receipts in 641-1100 merged with multiple data sources. See section B in the online appendix for details.

Table D.3: **Exogeneity of the Distance to *'Arish***

	(1) Log (urban population) in district in 300	(2) =1 if district on Holy Family route in 400	(3) =1 if *Autopract* estates in district in 600	(4) =1 if Byzantine garrison in district in 600
Distance to *Arish* (km)	0.0003	-0.0005	-0.0006	-0.0001
	(0.0005)	(0.0003)	(0.0005)	(0.0003)
Obs (districts)	76	76	35	76
R^2	0.005	0.028	0.033	0.001

Notes: Standard errors are in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$. A constant term is included in all regressions.
Source: Multiple data sources. See section B for details.

Table D.4: **Estimating Adult Mortality from the 1848-1868 Census Samples**

Age Group	Copts				Muslims			
	Estimated Size in 1848	Estimated Size in 1868	Estimated Life Expectancy (Method 1)	Estimated Life Expectancy (Method 2)	Estimated Size in 1848	Estimated Size in 1868	Estimated Life Expectancy (Method 1)	Estimated Life Expectancy (Method 2)
0-9	90,740	117,801	NA	NA	1,148,827	1,458,614	NA	NA
10-19	32,981	51,600	41.45	42.9	377,685	603,264	43.44	44.82
20-29	33,290	52,466	44.59	44.59	406,293	622,071	49.08	48.73
30-39	40,100	36,657	30.44	32.2	457,208	481,535	32.97	32.65
40-49	27,031	26,187	25.46	24.72	348,101	360,926	25.9	23.79
50-59	15,325	25,345	22.61	21.02	243,063	288,588	21.83	19.98
60-69	11,406	12,595	17.67	16.1	171,180	195,387	16.88	13.53
70-79	7,849	10,899	11.52	9.03	99,442	111,561	12.26	8.68
80+	7,094	5,107	NA	NA	125,336	78,559	NA	NA

Notes: The handbook of the United Nations Population Division (2002, pp. 5-20) outlines a methodology for estimating adult mortality from any two consecutive censuses that are separated by an interval of x years, where x is a multiple of 5. The methodology uses the relative sizes of age cohorts, defined in groups of 5-year intervals, in the two censuses in order to estimate the probability of survival to an age $y+x$, conditional on being of age y in the first census. A slightly different methodology, the synthetic survival ratio, calculates the growth rate of each age cohort in order to make the methodology applicable to any census interval, i.e. not necessarily a multiple of 5. I applied the two methods to the census samples of 1848 and 1868, in order to estimate the adult mortality by religious group. A few caveats arise though: (a) the time interval separating the two Egyptian censuses (20 years) is too long to apply the two methodologies; ideally, the interval should be around 5 or 10 years, (b) I do not have 100-percent samples of the two censuses and so there is a sampling error in estimating the size of each age cohort, and (c) there is a problem of age misreporting; in particular, age heaping and age exaggeration, which is typical in historical censuses and even contemporary censuses in developing countries. Age misreporting is likely correlated with SES and may thus vary in a non-random way across religious groups, where Muslims are more likely than Copts to misreport their true age. In order to mitigate age misreporting, I defined age groups in intervals of 10 years instead of 5 years.
Source: The full 1848 and 1868 population census samples.

Table D.5: **Copts' Tax Revolts in 726-768**

Year	Region	Reasons Cited	Parties Revolting
726	Nile Delta	Tightening state control over the tax system	Copts
740	Nile Valley	Stricter tax enforcement, higher tax rate, uniform tax regardless of income	Copts
750	Nile Delta	Heavy taxation and general suffering	Copts; Arabs revolted to overthrow the Umayyads
753	Nile Delta	Reorganizing the tax system	Copts
768	Nile Delta	Abbasids' fiscal reforms	Copts

Notes: I excluded ten tax revolts that erupted in 783-866 (nine of which were in the Nile Delta) because both Muslims and Copts participated in these revolts and, thus, they were likely motivated by other reasons apart from the poll tax.
Sources: Morimoto (1981, pp. 145-72) and Mikhail (2004, pp. 195-211).

Table D.6: **"Protégés" by Religious Group in 1848 and 1868**

Muslims		Copts		Non-Coptic Christians		Jews	
1848	1868	1848	1868	1848	1868	1848	1868
0.003%	0.076%	0.016%	0.069%	17.080%	25.705%	7.756%	10.446%

Notes: Non-Coptic Christians are primarily Greeks, Levantines, and Armenians.
Source: The full 1848 and 1868 population census samples.

Table D.7: **Rulers' Preference for Recruiting Copts in the Bureaucracy- OLS**
Dependent variable is indicated on the top of each column

	Share of Copts within mid-low bureaucrats			Number of Coptic bureaucrats per Muslim bureaucrat	
	(1)	(2)	(3)	(4)	(5)
Copts' population share	2.445***	2.277***	2.282***	30.046**	27.325**
	(0.379)	(0.369)	(0.317)	(12.446)	(10.711)
Number of bureaucrats		0.012***	0.013***		
		(0.002)	(0.002)		
Controls?	No	No	Yes	No	Yes
Observations	59	59	59	50	50

Notes: Regressions are weighted by district's population. A constant is included in all regressions. $*p < 0.10, **p < 0.05, ***p < 0.01$. Standard errors are in parentheses. Controls include log (urban population) in 300 and the Holy Family legendary route indicator.
Source: The 1848-1868 census samples aggregated to the district level and combined with multiple data sources. See section B for details.

References

Abul-Makarim (1984 [1200]). *tarikh al-kana'is wal adyera (History of Churches and Monasteries)*. Dar-Al-Na'am, Cairo. Anba-Samuel (Ed.).

Al-Barri, A. K. (1992). *al-qaba'il al-'arabiya fi misr fil qurun al-thalatha al-ula lil hijra (Arab Tribes in Egypt in the First Three Hijri Centuries)*. Egyptian General Book Authority, Cairo.

Al-Maqrizi (2002 [1500]). *al-mawa'iz wal i'tibar fi zhikr al-khitat wal athar (Sermons and Considerations in Examining Plans and Monuments)*, volume 4.1. Al-Furqan Islamic Heritage Foundation, London. A. F. Sayyid (Ed.).

Anba-Bishoy, P. (1999). *The Flight of the Holy Family to Egypt*. Nubar Printing House, Cairo.

Cuno, K. M. (1992). *The Pasha's Peasants: Land, Society, and Economy in Lower Egypt 1740-1858*. Cambridge University Press, Cambridge.

Gabra, G. (2001). *Be Thou There: The Holy Family's Journey in Egypt*. American University in Cairo Press, Cairo.

Hardy, E. R. (1931). *The Large Estates of Byzantine Egypt*. Columbia University Press, New York.

Ibn-Al-Jay'an (1898 [1477],). *al-tuhfa al-sanneya bi asmaa' al-bilad al-misriya (The Treasure Book of Names of Egyptian Localities)*. Al-Matba'a Al-Ahliya, Cairo.

Jewitt, I. (2004). Notes on the Shape of Distributions. Mimeo, Oxford University.

Macdonald, M. C. A. (1999). Personal Names in the Nabataean Realm: A Review Article. *Journal of Semitic Studies*, 44(2):251–289.

Mahmoud, A. A. (2009). *al-jizya fi misr 1713-1856 (The Poll Tax in Egypt in 1713-1856)*. Supreme Council of Culture, Cairo.

Maspero, J. (1912). *Organisation Militaire de l'Égypte Byzantine*. Bibliothèque de l'École des Hautes-Études, Paris.

Mikhail, M. S. A. (2004). *Egypt from Late Antiquity to Early Islam: Copts, Melkites, and Muslims Shaping a New Society*. PhD thesis, University of California Los Angeles, Los Angeles.

Morimoto, K. (1981). *The Fiscal Administration of Egypt in the Early Islamic Period*. Dohosha, Kyoto.

Ramzi, M. (1994 [1954]). *al-qamus al-gughrafi lil bilad al-misriya min 'ahd qudama' al-misriyin ila sanat 1945 (Geographic Dictionary of Egyptian Localities from the Time of Ancient Egyptians to 1945)*. Egyptian General Book Authority, Cairo.

Rapoport, Y. (2004). Invisible Peasants, Marauding Nomads: Taxation, Tribalism and Revolt in Mamluk Egypt. *Mamlūk Studies Review*, 8(2):1–22.

Saleh, M. (2013). A Pre-Colonial Population Brought to Light: Digitization of the Nineteenth-Century Egyptian Censuses. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(1):5–18.

Saleh, M. (2015). The Reluctant Transformation: State Industrialization, Religion, and Human Capital in Nineteenth-Century Egypt. *Journal of Economic History*, 75(1):65–94.

Tousson, O. (1926). *La Géographie de l'Egypte à l'Epoque Arabe*. L'Institut Français de l'Archéologie Orientale, Cairo.

United Nations Population Division (2002). Methods for estimating adult mortality. Technical report, United Nations, New York.

Wilson, A. (2011). City Sizes and Urbanization in the Roman Empire. In Bowman, A. and Wilson, A., editors, *Settlement, Urbanization, and Population*, pages 161–195. Oxford University Press, Oxford.