A Pre-Colonial Population Brought to Light: Digitization of the Nineteenth Century Egyptian Censuses[*]

Mohamed Saleh[†]

(September 10, 2012)

Abstract

Our knowledge of pre-colonial Middle Eastern populations has been limited by the lack of data. The 1848 and 1868 Egyptian censuses provide two snapshots of the Egyptian population in its early attempts to make the transition into a modern society. These censuses are perhaps the earliest in the Middle East and among the earliest in any non-Western country to include individual-level information on all segments of the population, including females, children, and slaves, on a wide range of demographic and socioeconomic variables. This paper describes the digitization of two nationally representative samples of the 1848 and 1868 censuses from the original manuscripts at the National Archives of Egypt. It then introduces an application of the samples in Egyptian economic history.

Keywords: Egypt historical census; historical demography; Muhammad Ali; Middle Eastern economic history; state industrialization

[*] I sincerely thank Dora Costa and Leah Boustan for their great support and advice. I gratefully acknowledge the financial support for the project that I received from IPUMS, Economic History Association, and University of Southern California. I also gratefully acknowledge the logistic support of the administration of the National Archives of Egypt (Saber Arab, Rifaat Hilal, and Emad Hilal), which was crucial to undertake the project. I thank Julie Iskander for the IT guidance and all the members of the data entry team for their great work, without which the project would not have come to fruition. I sincerely thank Jeffrey Nugent, Ragui Assaad, Naomi Lamoreaux, Steven Ruggles, Robert McCaa, Roger Owen, Khaled Fahmy, Kenneth Cuno, Terry Walz, Dimitris Pipinis, and the participants at the UCLA Economic History Proseminar, the International Commission of Historical Demography Conference 2010, and the Population Association of America 2011 for their valuable comments and suggestions. All errors are mine.

[†] Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST) (E-mail: mohamed.saleh@tse-fr.eu).

1. Introduction

> *It is therefore necessary for Us to enumerate exactly the people of Our country so that it may be a cause of its progress in civilization.*
>
> Muhammad Ali Pasha (1845)[1]

> *The conducting of the census is a matter of benefit for the homeland, and whoever understands its merit will strive body and soul for the sake of its execution.*
>
> Census Order (1847)[2]

In 1816, Muhammad Ali Pasha, the autonomous Ottoman viceroy of Egypt, embarked on one of the world's earliest state industrialization programs. The program, which failed in the end to transform Egypt into an industrial nation, was so intriguing to historians that it remained one of the most central topics in the social and economic history of modern Egypt and the Middle East at large. Nevertheless, despite the vast historical literature on the topic (Al-Gritli 1952; M. Fahmy 1954; Al-Hitta 1967; Marsot 1984; Owen 2002), little is known with respect to the size of the workforce in these early Egyptian manufactories, let alone the occupational, ethnic, age, or religious distributions of their workers. Such detailed picture, which is indispensable to evaluate the social impact of the experiment, is unattainable either because the data do not exist at all, or, even if they do exist, they are too sporadic to form a complete picture of the topic.

In this context, the censuses from 1848 and 1868 offer two extraordinary data sources on the Egyptian population in the nineteenth century. The censuses, conducted during the reigns of Muhammad Ali (1805-48) and Ismail (1863-79) respectively, are

[1] Cuno and Reimer (1997), p. 213.
[2] Cuno and Reimer (1997), p. 215.

preserved in 6,592 handwritten registers at the National Archives of Egypt (henceforth, NAE). They include information on a wide range of variables including location, name, ownership of dwelling, type of dwelling, relationship to household head, age, gender, religion, ethnicity, nationality, place of origin, occupation, and place of work, among other variables. The data are recorded systematically for every individual in the household including females, children, and slaves. Using these records, one is able to examine the social impact of Muhammad Ali's state industrialization, as well as many other central topics in Middle Eastern social and economic history that are currently understudied because of the lack of data.

The Egyptian censuses are perhaps the oldest modern censuses in the Middle East and among the earliest from any non-Western country. Several salient features make these censuses "modern" and distinguish them from both the Ottoman tax registers of the fifteenth and sixteenth centuries (*Tahrir Defterleri*) (Barkan 1957; Cosgel 2004; Hütteroth and Abdulfattah 1977) and the contemporary Ottoman censuses of the nineteenth century. While the Ottoman records enumerated only adult males (for taxation and military conscription purposes), the Egyptian censuses enumerated everyone including females, children, and slaves.[3] Also, in the same spirit of contemporary Western censuses, the Egyptian censuses kept standardized individual-level lists and not a mere "count" of heads. Finally, the Egyptian censuses recorded socioeconomic

[3] According to Karpat (1978), the first Ottoman census took place in certain parts of the empire in 1831-8, but was an enumeration of households rather than individuals. The second census took place in 1844 and was an enumeration of the adult male population. Finally, the third census (1866-73) was limited to the Danube province and was also an enumeration of adult males. In fact, the first census in the Ottoman Empire that contained an enumeration of females took place in 1881/82-1893 (Karpat 1978).

information (such as school enrollment and occupation), perhaps indicating the interest of the Egyptian government in the "human capital" of the population.

The digitization of these censuses is significant for researchers in various fields. It opens the door for cliometricians to employ quantitative methods in studying Middle Eastern history at the micro- or individual-level. So far, the study of the history of this region has been based on qualitative methods, with the application of quantitative methods being limited to the macro- or institutional-level. While qualitative research is indispensable in understanding various historical phenomena, the introduction of quantitative methods will enrich the historical research and will allow tackling a number of research questions that cannot otherwise be addressed.

In addition to this, little is known of the demographic and socioeconomic characteristics of pre-colonial populations in the Middle East. Such knowledge is hindered by the lack of unbiased sources of information that can provide *representative* snapshots of the population. The available sources (such as tax, cadastral, and court registers) are usually confined to specific segments of the population (perhaps the wealthier strata), and are hence of limited use to the historical demographer who is interested in inferring the overall characteristics of the population. By the inclusion of all segments of the population, the census records will allow the study of the basic demographics of this pre-colonial Middle Eastern population.

Finally, a significant feature of the census records is that they come from the pre-colonial period, unlike comparable historical censuses from other developing countries that were conducted under colonial administration (e.g. British India 1872). Thus, when

combined with later post-colonial censuses, this source will allow historians and social scientists to examine the impact of colonization on the colonized population.

This paper describes the digitization project of two nationally representative samples of the 1848 and 1868 censuses, which I undertook with the help of data entry assistants at the NAE. The project resulted in the creation of two datasets from the 1848 and 1868 censuses each with about 80,000 individual observations. The datasets are currently being integrated into the historical censuses of the North Atlantic Population Project (NAPP) and will be disseminated on the web, with free access to all, as soon as the integration is done. Upon their public dissemination, the two datasets will hopefully boost quantitative research on the history of the Middle East.

The rest of the paper is organized as follows. Section 2 describes the census records: their historical context, their format, and the enumeration methodology. Section 3 describes the sampling strategy. Section 4 discusses how the digitized samples can shed light on Muhammad Ali's program, as an example for the research questions that can be tackled using this new data source. Section 5 concludes.

2. The Census Records: Historical Context, Format, and Enumeration Methodology

2.1 Historical Context

In 1845, Muhammad Ali Pasha ordered the undertaking of a nationwide census. The census operations started in rural provinces as early as in 1846 but were then extended to the urban ones in 1847 and were further repeated in 1848 in most of the rural units that were already enumerated in 1846-7. Although a number of later enumerations took place in the 1850s and the early 1860s, these remained limited to scattered geographic units and

did not constitute a national census. Egypt had to wait for almost twenty years to have its second wide-scale census in 1865-69 during the reign of Khedive Ismail (1863-79). That second census did not cover the entire country, however. The result of these two census operations was the creation of standardized lists of individual-level census records kept in 6,592 hand-written registers (the vast majority of which belong to the years 1848 and 1868) that were since then preserved at the NAE.

Before the discovery of these censuses, the series of Egyptian modern censuses was believed to have started in 1882 (a few months before the British occupation). The 1882 census was widely known because it resulted in a published census report at the village/urban quarter-level although the microdata did not survive.[4] The second published census was conducted in 1897, and was then followed by almost regular decennial censuses (1907, 1917, 1927, 1937, 1947, 1960, 1966, 1976, 1986, 1996, and 2006). However, the microdata for the censuses over the period 1882-1976 were destroyed either by choice or by chance. This resulted in the anomalous situation that the only available census microdata in Egypt are either the early ones from 1848 and 1868 or the most recent ones from 1986, 1996, and 2006.

But why did the Egyptian government conduct the 1848 and 1868 censuses? Answering this question is in no way definite. Perhaps, one incentive was to "control" people. The censuses, like earlier and contemporary Ottoman enumerations, provided the basis for tax collection, military conscription, and corvée work for public projects.

[4] There are a number of census registers at the NAE that date back to 1879-82. These registers include individual-level census records as well as preliminary tabulations and are probably remains of the mostly destroyed microdata of the 1882 census.

Nevertheless, enumerating females and children and including socioeconomic variables may not be motivated by "control." Here, one may speculate, following Alleaume and Fargues (1998), that the Egyptian students who were sent on educational missions to Europe in the first half of the nineteenth century may have played a role in introducing the European enumeration techniques to Egypt. Also, the European technicians, who were employed in the Egyptian administration at the time, may have also played a role.

2.2 Format of the Census Records

The census registers are organized by province. The registers of urban provinces are further classified by district, which is divided into urban quarters,[5] while those of rural provinces are classified into towns, villages, hamlets, and Arab tribes.

Enumeration in urban provinces is on a dwelling-by-dwelling basis with the heading of each page stating the urban quarter and the street name.[6] For each dwelling, the dwelling type (e.g. ruined hut, tenement house) is recorded, followed by the legal status of the dwelling (e.g. private ownership, religious endowment), the name of the property right holder, and the dwelling number in the street.[7] Households residing in the dwelling are then recorded: males followed by females, and, within each gender, free individuals followed by free servants and slaves.

[5] This applies to Cairo and Alexandria. The city of Rosetta is treated as one single district and is hence directly divided into urban quarters. Besides, there are two single-city provinces that have one register each: Al-Arish (in Sinai Peninsula) and Al-Qusayr (on the Red Sea coast).
[6] See Appendix A for two scanned pages of the registers.
[7] I found a government order from 1847, published in Sami (1928, II: 547-52), on naming Cairo's streets and numbering its dwellings. The order might have been related to the then-ongoing census operations.

In rural provinces, the records of provincial towns, villages, and hamlets are classified into: (i) locals, (ii) strangers, who are legal or illegal immigrants,[8] and (iii) deserters, those who escaped illegally from the geographic unit. Locals are recorded according to the section (*Hissa*) of the geographic unit they are residing in, with the household of the *Hissa*'s headman (*sheikh*) recorded first. Strangers and deserters, on the other hand, are classified geographically according to their place of origin (for strangers), or destination (for deserters).[9] Finally, the records of Arab tribes are classified by tribal sub-group (*gama'a*) and/or by location.[10]

For each individual, a systematic list of variables is recorded including:

1. *Name*: Egyptian name system uses the first name of the individual followed by the name of the father, the paternal grandfather, and so on. The names in the records are recorded up to the father's name. First names of free females in urban provinces are *not* recorded unless the female is the household head.

2. *Relationship to household head*: is recorded for all individuals. Confusion might arise in large rural households where the recorded relationship might be relating the individual either to the household head or to the previous household member.

3. *Age*: is recorded in full years (as of the next birthday) for every individual. Age is categorical (child or adult) for females in urban provinces. Age for infants is often recorded in months or days.

---

[8] To control peasants' flight, starting from 1829, immigration from one village to another required government permission. People were "illegally" deserting their villages in order to avoid taxation, military conscription, and corvée in public works (Cuno 1992, 121-4; K. Fahmy 1998, 99-103).

[9] In large provincial towns, the census records are classified by urban quarter and street name, and are hence similar to the recording in urban provinces.

[10] For Arab tribes, the recording starts with the household of the headman of the tribe.

4. *Occupation*: is recorded mainly for males. Occupational titles are detailed and reflect the highly specialized guild system at the time.[11] For male children, school enrollment is recorded as an occupation. Unemployment and retirement are recorded. Work establishment is recorded if the person is employed in governmental enterprises.

5. *Nationality*: This includes two categories: inside or outside government's control (i.e. Egyptian or foreigner). Nationals of other regions of the Ottoman Empire are recorded as "outside government's control," indicating that they are foreigners.

6. *Ethnicity*: is recorded (e.g. Turkish, Armenian, Black) for foreigners and slaves. For Egyptians, categories such as "son of an Arab" or "son of the country" are often used.

7. *Religion*: This includes Muslims, Christians, and Jews. The denomination within Christianity and Judaism is often recorded (e.g. Coptic Christian, Rabbi Jew), or can be inferred from ethnicity (e.g. Armenian Christian).

8. *Place of origin*: is recorded up to the village-level (rural origin) and the province-level (urban origin). However, it is *not* exactly the place of birth. In 1848 census, children inherit the place of origin of their father. The wife does *not* take the place of origin of her husband, and she takes instead that of her father.[12]

9. *Infirmities*: An infirmity is recorded if the individual has any (e.g. blind).

10. *Marital Status*: is recorded for female household heads.

However, randomness is also evident in the censuses. Information is often not standardized, with some variables omitted or mentioned irregularly. Religion and

[11] I found about 3,700 distinct occupational titles in each census. The titles "scribe" and "scribe at the customs department" are counted as distinct.

[12] In the 1868 census, the place of origin of the child is often different from the father, thus suggesting that the concept of this variable was getting closer to the "place of birth."

nationality are often omitted so long as the "default" category is assumed ("Muslim" and "inside government's control"). The census takers appear to have been confused over some variables. Age, which is a categorical variable for females in urban provinces, is sometimes recorded in full years. Also, there is confusion on recording occupation for male children.[13] Spelling mistakes are numerous, and scribes appear to have discretion in recording information, with some being keener on details than others. Finally, the records reflect the preliminary statistical knowledge of the census takers. Apart from tabulations on the age and occupational distributions at the street- or village section (*Hissa*)-level, the census operations in both years did not result in national census reports.

2.3 Enumeration Methodology

Enumeration follows the "de facto" principle: A person who is not present in his/her permanent place of residence at the time of enumeration is recorded in that place as "not enumerated," and is enumerated instead in his/her actual location.[14] Servants and slaves are recorded as part of the household they are residing with. However, there are a few exceptions to this principle: Foreign protégés are not enumerated with a note that they are to be enumerated by their consulates. Also, military personnel and students in military schools are not enumerated with a note that they are to be enumerated in the army's census. Finally, Cairo's notables in 1848 (but not in 1868) are not enumerated in their

[13] Occupation for males who are less than 6 years old is often left blank or is recorded as "unemployed." The words "child" and "infant" are also often recorded.

[14] The "not enumerated" note is either mentioned explicitly or by assigning a "zero" to the individual indicating that he/she is not added to the enumeration count of the page. A non-enumerated person has only his/her name recorded (and occupational title for military personnel and students in military schools).

place of residence and are recorded separately in a special register.[15] Their records only mentioned the notable's name and the count of males and females in the household.

The 1848 census operations are described in the census orders of 1845 and 1847 (Sami 1928, II: 535-6; Cuno and Reimer 1997, 213-6). The census was conducted by the headmen of the urban quarters and villages' sections, under the supervision of an officer from the army. There were threats of harsh punishments for the concealment of individuals. Each enumerated individual was issued a "ticket" indicating that he/she has been enumerated. For the 1868 census, the only order that I found is very brief and refers to a parliamentary decree (Sami 1928, III: 785). However, the 1868 census operations seem to have been similar to those of 1848 as far as one can infer from the format of the registers.

3. Sampling Strategy

The digitization project aimed at creating 1-percent samples of each of the 1848 and 1868 censuses (5 percent in the major cities) (about 80,000 individuals in each year). In what follows, I will first describe the target population and the sampling frame. Then, I will explain the sampling strategy and the creation of sampling weights.

3.1 Target Population and Sampling Frame

I define the target population as those who were residing in Egypt at some point in time (the census night) in 1848 or 1868. However, the sampling frame, defined as the

[15] According to the census order (1847) (Cuno and Reimer 1997, 213-6), each quarter's headman had to send to the Prefect a list of the notables in his urban quarter. The Prefect then was to request counts of males and females in the household from the household head, which were to be recorded in the notables' register. This way, the census takers and the quarters' headmen in charge of the census operations could not intervene in the enumeration of the notables or access their households. Similarly, the *Khedivial Diwan* was responsible for collecting information on foreigners from the consulates.

available census records in each year, may be different. The discrepancy between the two concepts is due to: (i) under-enumeration and (ii) multiple enumerations of specific geographic units in 1846-48 (for the 1848 census) or in 1865-69 (for the 1868 census).

3.1.1 Under-Enumeration[16]

I use the list of geographic units from the published 1882 census as the comparison point in order to evaluate the completeness of the 1848 and 1868 enumerations. A few cities are entirely missing in the early censuses.[17] Table 1 shows the comparison in rural provinces. All the 1882- rural provinces have at least one register in 1848, yet two provinces have remarkably low enumeration rates at the village-level (7 and 14 percent respectively). In 1868, the problem is more serious, since 6 out of 14 provinces are completely missing, and three provinces have extremely low enumeration rates.

What about enumeration at the individual-level? I again use the 1882 census as the comparison point under the assumption that under-enumeration was less of a problem in that year. In particular, I estimate the "true" population of each province in 1848 (or 1868) based on its population share in the 1882 census.[18] [19] I then compare the "true"

[16] Under-enumeration could mean that some individuals were indeed not enumerated during the census operations, or that they were enumerated but their records were lost afterwards. In the absence of secondary information, there is no way to distinguish between the two possibilities and I thus treat them as one.

[17] Compared to the 1882 census' list of cities, Damietta, Suez, and Al-Qusayr (in 1868) are entirely missing. Enumerated cities are Cairo, Alexandria, Rosetta, Al-Arish, and Al-Qusayr (in 1848).

[18] This is based on two assumptions: (i) that the boundaries between provinces did not change between 1848 (or 1868) and 1882, and (ii) that the population distribution across provinces in 1848 (or 1868) remained the same in 1882. The first assumption is justified since only 2.5 percent of all geographic units in 1848 (less than 1 percent in 1868) belonged to a different province in 1882. The second assumption is perhaps reasonable across rural provinces, but might be less so between urban and rural provinces with presumably an increasing population share of urban provinces over time.

[19] A priori information on the population size of the entire country in 1848 (and 1868) is needed here. For 1848, I used the figure of 4,476,439 (Alleaume and Fargues 1998). I then assumed a constant annual growth rate between 1848 and 1882 in order to estimate the population size in 1868.

population to the actual size that has been enumerated.[20] Table 2 shows this comparison. In general, individual-level enumeration rates are similar to those in table 1.

I addressed under-enumeration in the pre-sampling phase as follows: If a province is entirely missing in 1848 (or 1868) or has less than 5 percent of its villages enumerated,[21] I excluded it from the sampling frame; otherwise, the province is included and sampled using its available records (i.e. its enumerated population). Thus, the sampling frame is *representative* of the target population under the assumption that non-enumerated individuals are randomly distributed. In the post-sampling phase, I used sampling weights to adjust for non-enumeration within provinces that are included in the sampling frame (Subsection 3.3).

3.1.2 Multiple Enumerations

Many geographic units were enumerated more than once in the 1848 (1868) census, and hence have more than one register. In this case, only the census register(s) from the *latest* enumeration year of each geographic unit was included in the sampling frame. So for example, if a village was enumerated in 1846, 1847, and 1848, I included only the register(s) from 1848 in the sampling frame.[22] [23] Table 1.3 shows the number of units

[20] I obtained the enumerated population of each province in the 1848 census by summing up the enumeration counts of all its census registers. Logistic problems, however, did not allow me to do the same for the 1868 census. Instead, I estimated the enumerated population of each province in that year using the total number of pages (adjusted for blank pages and tabulations) and the average number of individuals recorded per page in the sample taken from that province.

[21] Besides the entirely missing provinces, I excluded the registers of Al-Minya and Girga in 1868. However, I decided to include Al-Sharqiya in 1868 although only 1 percent of its villages were enumerated because it has 20 registers, which is a relatively large number.

[22] This assumes that the register(s) of the earlier year(s) enumerate exactly the same population as the latest year, and, that, if there is more than one register from the latest year, they must enumerate *different* segments of the population.

[23] I discovered the existence of duplicate registers for a few units after their inclusion in the sampling frame of the 1848 census, and hence these units had higher chances of selection. To correct for this, I adjusted the

with multiple enumerations by province and year. The problem exists in only a few provinces, and is more prevalent in the 1848 census.

3.2 Sampling Strategy

The sample is stratified by province. Stratification is motivated by the desire to make the geographic distribution of the sample reflect that of the target population at the province-level. I estimated the population size of each stratum (province) using its population share in the 1882 census and the a priori information on the population size of the whole country in 1848 and 1868. I then calculated the targeted sample size in each province, by setting the targeted sampling rate at 5 percent in the major cities and 1 percent in all other provinces. Table 4 shows the geographic distribution of the two samples along with the targeted and actual sampling rates.

I applied systematic sampling by page within each stratum: A page is chosen randomly in the beginning of each stratum, and successive pages are taken into the sample according to a pre-calculated interval.[24] A sampled page is accepted if at least 75 percent of it has individual-level records. If a page is rejected, the closest preceding or succeeding page, which satisfies the 75-percent rule, is taken on an alternating basis. If a page is accepted, all the households in the page are recorded in the sample, with both dwelling-level information (including the number of households and individuals in the dwelling), and reference information (register code and page number).[25] Overall,

---

sampling weights of individuals in these units by multiplying their original weights (subsection 3.3) by the reciprocal of the number of duplicates that were included in the sampling frame.

[24] The interval for a given province is calculated based on the targeted sample size, the total number of pages of the registers of the province, and the average number of individual records per page.

[25] Specifically, every household that *starts* in the page is included in the sample. I defined the starting point of a household as the line including the information of the first member in the household. Hence, if a

systematic sampling is less costly and easier to apply than pure random sampling. It also ensures the geographic spread of the sample within each stratum.

A few notes on sampling are in order. First, group quarters are dwellings where unrelated individuals reside (e.g. monastery, jail). I treat a group quarter as a single household. Thus, if a group quarter starts in a sampled page, it is included in the sample in its entirety.[26] Second, fragments, meaning individuals enumerated separately from their place of residence, are included in the sample if they fall within the sampled page with the note that they are fragments. Finally, in a few registers in 1848 Alexandria, breaks between households are not marked. Instead, individuals residing in a dwelling are classified by nationality, place of origin, and gender, *regardless of* the household they belong to. In these cases, I decided to enter the dwelling in its entirety.

Overall, the sampling strategy is similar to both the 1850 U.S. census sample constructed by IPUMS (Ruggles and Menard 1995), and the unreleased and incomplete Egyptian 1848 census sample described in Alleaume and Fargues (1998) and Fargues (2003). Examples of the digital samples are shown in Appendix B.

3.3 Sampling Weights

Sampling weights are required when estimating population means, totals, and proportions. They adjust for (i) different actual sampling rates across provinces (table 4),

---

household starts in a previous page and continues on the sampled page it is *not* included in the sample. But, if a household starts on the sample page and continues in the following page it is included in its entirety. This ensures that all households have an equal chance of appearing in the sample regardless of their size.

[26] The IPUMS 1850 U.S. census sample applies individual-level sampling to a group quarter instead of recording it in its entirety as a single household. The drawback of my approach is the resulting increase in the standard error because observations in a group quarter are more likely to be correlated. Nevertheless, the very small number of group quarters that I found in the census registers mitigates such concerns.

and (ii) non-enumeration within provinces included in the sampling frame. The final weight for any observation is the product of these two weight adjustments:

(i) Sampling Weight Adjustment for Different Sampling Rates: The probability of selection of an observation in a given stratum can be defined as the sample size divided by the population of the stratum. One can see from table 4 that the actual sampling rate (or the probability of selection) varies across provinces. The sampling weight adjustment (W_A) is defined as the reciprocal of the probability of selection.

(ii) Sampling Weight Adjustment for Non-Enumeration: Enumeration within urban provinces that are included in the sampling frame is close to complete. Hence, I focus here on non-enumeration within rural provinces. I estimate the probability of enumeration of a geographic unit using a Probit model, where the dependent variable is a dummy for having at least one census register in 1848 (or 1868).[27] I regress this variable on the characteristics of the geographic unit that are known for both enumerated and non-enumerated units. These are: location (I try three specifications with region, province, and district fixed effects), type (village, hamlet/Arab tribe, and provincial town), and population. The results are shown in table 5. More populated units have higher probability of enumeration. Provincial towns have lower probability of enumeration than villages, while hamlets/Arab tribes have higher probability of enumeration. Units in Upper Egypt are more likely to be enumerated than units in Lower Egypt. The sampling weight adjustment (W_B) is defined as the reciprocal of the estimated probability of enumeration.

[27] I am grateful to Ragui Assaad for this suggestion.

4. Application to Research Questions

In this section, I will first provide descriptive statistics from the digital samples in order to give a general idea of the information contained in this data source. Then, I will discuss how this data source can provide useful insights on evaluating the social impact of the state industrialization program carried out by Muhammad Ali Pasha, as an example for the research questions that can be answered using the census samples.

4.1 Descriptive Statistics

Table 6 shows the highest frequencies of a set of variables from the 1848 and 1868 samples. I show the statistics for the two major cities (Cairo and Alexandria) separately from other (mainly rural) provinces. 73 percent of the population of Cairo and Alexandria in both 1848 and 1868 lived in houses, unspecified, or unknown dwelling types, 16 percent lived in low-status dwellings (e.g. ruined huts, courtyards, and single rooms), 7 percent lived in multiple-household dwellings (e.g. tenement houses), and 3 percent lived in production sites or workplaces (e.g. coffee shops and bathhouses). There is no variation in the dwelling type in other provinces.

Dwelling legal status is almost only found in Cairo and Alexandria. Private or public ownership is the most dominant legal status with about 64 percent in 1848 (71 percent in 1868) residing in privately- or publicly-owned dwellings, followed by *Waqf* dwellings, that are owned by the religious endowments of individuals or entities.

Gender distribution is balanced and does not indicate a gender bias in the census. Muslims constitute the vast majority in both the 1848 and 1868 samples. Christians are the largest religious minority followed by Jews. About 7 percent of the population of

Cairo and Alexandria are foreigners (outside government's control), as opposed to only 1 percent in other provinces. The slave population is concentrated in Cairo and Alexandria, constituting 2 to 3 percent of the population. Turks represent the largest ethnic minority in 1848, followed by blacks. The latter, however, overtook the Turks in 1868.

4.2 The Egyptian Nineteenth Century State Industrialization Program

Historians have long debated the intriguing manufacturing program carried out by Muhammad Ali Pasha, who is widely regarded as the founder of modern Egypt. Apart from the debate on the *economic* evaluation of the program with respect to its efficiency in production, which requires gathering data on the costs of production and market prices, the census records can provide useful insight into the *social* impact of the program on the labor force. This social face is a central, yet, an often ignored, aspect of the manufacturing experiment. As K. Fahmy (2009, 126) points out, it is important when evaluating Muhammad Ali's experiment to take into account "*the millions of Egyptian people whose lot, as a direct result of what Mehmed Ali actually did achieve, was that of hardship and suffering*" (italics mine). In this subsection, I will discuss the type of information that one can extract from the census samples in order to shed light on this social side. I will focus on two questions: First, to what extent did the manufacturing program lead to the disappearance of the traditional guilds? Second, did the state manufactories provide routes for upward mobility for the labor force? I want to emphasize, however, that I am not trying to *answer* these questions in this limited space. I am only giving examples to the sort of information that the census samples can provide with respect to these questions.

The first question triggered a long debate among historians. M. Fahmy (1954, 73-6) argues that state industrialization led to the abolition of restrictions on labor that existed under the guild system, and thus to the complete liberation of the labor market. Baer (1964, 127-49) criticizes this argument and refers to the strong qualitative evidence that shows that the guild system survived until the early twentieth century, and that the state manufactories did not give the final blow to the medieval guilds, although it might have indeed hurt specific guilds. In support of Baer's thesis, Owen (2002, 76) points out that even textile workers who were allegedly hurt the most did survive in the aftermath of the closure of many of the manufactories. In the same direction, Marsot (1984, 181-3) states that the guild system remained largely unaffected except in the occupations that were in direct competition with the manufactories, although she also notices that the factory system opened possibilities for children to work and to be promoted "by training rather than by heredity" in contrast to the guild rules. Nonetheless, a slightly different viewpoint is provided by Ghazaleh (1999, 116-9), who criticizes Baer's thesis on the grounds that although the guilds did indeed survive until the early twentieth century they were largely weakened by the manufacturing program and gradually came under state control. In response to Baer's claim that the artisanal guilds, which were affected the most by the program, did not constitute more than one third of the total size of guild membership, Ghazaleh points to the difficulty of estimating the total number of guild members let alone their breakdown.

The census samples can provide useful insights into answering this question. First, by having the occupational title recorded for every male, along with the name of the work

establishment if the person is employed by the government, one is able to estimate the percentage of labor force that was employed in the state manufactories (henceforth, modern sector) in 1848 under Muhammad Ali and twenty years later in 1868 under a second wave of state industrialization that focused on transportation. Second, equally importantly, one is able to estimate the size and breakdown of the traditional guilds outside the manufactories (henceforth, traditional sector). An important caveat here is that the census records include information on *occupations* rather than guild membership. Although the two variables are not the same, one may argue that the occupational title is a reasonable proxy for guild membership, given the fact that there was a high degree of specialization within the guild system, with "the extreme splitting of occupations into guilds of specialized branches" (Baer 1964, 25). At any rate, holding this caveat in mind, the census records allow us to estimate the extent to which the traditional occupations/guilds disappeared due to the emergence of the modern sector, and also, to determine the specific occupations/guilds that were hurt the most. Table 7 shows that the modern sector employed 11 percent of Cairo's adult active male population (at least 15 years old) in 1848, but the percentage went down to 3 percent in 1868. In Alexandria, the second largest city, the percentage of workers in the modern sector was about 2 percent in 1848 and went up slightly to 4 percent in 1868. The other provinces had very low share of the population working in the modern sector in both years but the share was higher in 1848. Overall, it seems that state manufactories in 1848 were not large enough to crowd out the traditional occupations/guilds. Moreover, the percentage of workers employed in the modern sector fell sharply in 1868. One can also extract further information on which

guilds were hurt the most by state industrialization, by going through the detailed occupational distribution in both years.

The second question has to do with the routes of occupational mobility that were provided by the modern sector in 1848 and 1868. While historians agree that workers suffered from conscription into state manufactories (M. Fahmy 1954, 73-6; Marsot 1984, 184; Owen 2002, 76), little is known regarding the occupational distribution of workers in the modern sector, and how it compares to that in the traditional sector. This knowledge is crucial in order to evaluate the degree to which jobs in the modern sector compared (whether favorably or not) to those in the traditional sector, and hence, whether they provided routes for upward or downward occupational mobility. To the best of my knowledge, only Ghazaleh (1999, 131-3) provides a list of workers broken down by occupation, their working days, and their wages in *Al-Khurunfish*, the oldest textile manufactory in Egypt that was constructed in 1816. Table 8 shows the occupational distribution of workers in the modern sector in 1848 and 1868 in Cairo and Alexandria, and how it compares to that in the traditional sector. In 1848, the vast majority of the workers, about 69 percent, worked as "construction workers, carpenters, building painters, transportation workers, and laborers." About 9 percent worked as "wood treaters, gunpowder makers, textile workers, millers, and food processors." White-collar workers (e.g. scribes, stores clerks) constituted only 9 percent. In 1868, however, the distribution shifted slightly towards white-collar workers (17 percent) and away from "construction workers, carpenters, building painters, transportation workers, and laborers" (57 percent). This may suggest that the modern sector was actually a route for

downward rather than upward mobility in 1848, but may have provided a route for upward mobility to white-collar jobs in 1868.[28]

5. Conclusion

The 1848 and 1868 newly digitized census samples provide a rich source of information on the Egyptian population in the nineteenth century, which can benefit researchers in various disciplines. Historical demographers may use the samples to examine patterns of fertility, mortality, marriage, immigration, and multigenerational households in this population. Moreover, given the availability of historical census records from other countries in the NAPP database, international demographic comparisons will be feasible. Cliometricians and historians of the Middle East may benefit from the digitized samples in examining slavery, modernization, and education, among other topics, and should be able to make quantitative statements about each of these phenomena. Genealogists may make use of the samples in exploring history of families in Egypt. Furthermore, spatial data on location in the digitized samples may allow researchers in spatial sciences and urban history of Middle Eastern cities to reconstruct a detailed historical map of nineteenth century Egyptian cities. Overall, the 1848 and 1868 census samples may open entirely new possibilities for quantitative research in studying the social and economic history of Egypt and the Middle East at large.

[28] This may have been the case because of the nature of modern transportation projects in 1868 such as railways, telegraph, and steam navigation, which required more administrative/clerical jobs than Muhammad Ali's manufactories in 1848.

References

Alleaume, G. and P. Fargues. 1998. La naissance d'une statistique d'état. *Histoire & Mesure* 13: 147-193.

Baer, G. 1964. *Egyptian Guilds in Modern Times*. Jerusalem: The Israel Oriental Society.

Barkan, Ö. 1957. Essai sur les données statistiques des registres de recensement dans l'empire Ottoman aux XVe et XVIe siècles. *Journal of the Economic and Social History of the Orient* 1: 9-36.

Cosgel, M. M. 2004. Ottoman tax registers (*Tahrir Defterleri*). *Historical Methods* 37: 87-100.

Cuno, K. M. 1992. *The Pasha's peasants: Land, society, and economy in Lower Egypt, 1740-1858*. Cambridge: Cambridge University Press.

Cuno, K.M. and M. J. Reimer. 1997. The census registers of nineteenth-century Egypt: A new source for social historians. *British Journal of Middle Eastern Studies* 24: 193-216.

El-Gritli, A. 1952. *Tarikh al-sina'a fi Misr fil nisf al-awwal min al-qarn al tasi' 'ashar* (History of industry in Egypt in the first half of the nineteenth century). Cairo: Dar El-Ma'aref.

El-Hitta, A. A. 1967. *Tarikh Misr al-iqtisadi fil qarn al tasi' 'ashar* (Economic history of Egypt in the nineteenth century). Cairo: Al-Masry Publishing.

Fahmy, K. 1998. *All the Pasha's men: Mehmed Ali, his army, and the making of modern Egypt*. Cambridge: Cambridge University Press.

Fahmy, K. 2009. *Mehmed Ali: From Ottoman governor to ruler of Egypt*. Oxford: One World.

Fahmy, M. 1954. *La revolution de l'industrie en Egypte et ses consequences sociales au 19^{e}siecle (1800-1850).* Leiden: E. J. Brill.

Fargues, P. 2003. Family and household in mid-nineteenth-century Cairo. In *Family history in the Middle East*, edited by B. Doumani, New York: State University of New York Press, 23-50.

Ghazaleh, P. 1999. *Masters of the trade: Crafts and crafts people in Cairo, 1750-1850*. Cairo: The American University in Cairo Press.

Hütteroth, W. and K. Abdulfattah. 1977. *Historical Geography of Palestine, Transjordan and Southern Syria in the Late 16th Century*. Erlangen: Selbstverlag der Fränkischen Geographischen Gesellschaft.

Karpat, K. H. 1978. Ottoman population records and the census of 1881/82-1893. *International Journal of Middle East Studies* 9: 237-74.

Marsot, A. L. A. 1984. *Egypt in the reign of Muhammad Ali*. Cambridge: Cambridge University Press.

Owen, R. 2002. *The Middle East in the world economy 1800-1914*. UK: I. B. Tauris & Co Ltd.

Ruggles, S. and R. R. Menard. 1995. *Public use microdata sample of the 1850 United States census of population: User's guide and technical documentation*, 7-10. Minneapolis: Social History Research Laboratory.

Sami, A. 1928. *Taqwim Al-Neel* (Chronicles of the Nile). Vol. II-III. Cairo: Dar Al-Kutub.

Table 1: Estimating Enumeration in Rural Provinces at the District- and Village-Level in the 1848 and 1868 Censuses

Province	Districts			Villages		
	Number of Districts in 1882	% Enumerated in 1848	% Enumerated in 1868	Number of Villages in 1882	% Enumerated in 1848	% Enumerated in 1868
Al-Daqahliya	6	100	100	443	74	32
Al-Sharqiya	6	100	33	433	74	1
Al-Qalyubiya	3	100	0	163	84	0
Al-Gharbiya	10	80	0	547	7	0
Al-Minufiya	5	100	0	331	62	0
Al-Buhayra	6	100	0	304	48	0
Al-Giza	4	100	100	167	93	90
Bani Suwayf	3	100	100	168	83	31
Al-Fayyum	4	100	50	91	97	91
Al-Minya	4	100	25	267	89	1
Asyut	10	100	100	322	55	57
Girga	5	100	20	190	67	1
Qina	4	100	0	109	62	0
Isna	6	33	0	107	14	0

Provinces and districts are defined according to the 1882 census administrative borders. A district or a village is considered enumerated if it has at least one register in the 1848 (1868) census.

Table 2: Estimating Enumeration in All Provinces at the Individual-Level in the 1848 and 1868 Censuses

Province	1848			1868		
	True Population	Enumerated Population	% Enumerated	True Population	Enumerated Population	% Enumerated
Cairo	255,978	237,000	93	320,360	258,898	81
Alexandria	158,021	119,788	76	197,766	154,876	78
Rosetta	13,233	18,187	137	16,562	2,257	14
Al-Arish	2,005	2,311	115	2,509	1,820	73
Al-Qusayr	1,659	3,024	182	-	-	-
Al-Daqahliya	397,720	336,402	85	497,752	144,489	29
Al-Sharqiya	259,766	279,918	108	325,100	20,384	6
Al-Qalyubiya	173,804	178,232	103	-	-	-
Al-Gharbiya	627,764	42,287	7	-	-	-
Al-Minufiya	439,745	126,287	29	-	-	-
Al-Buhayra	270,526	144,583	53	-	-	-
Al-Giza	187,672	217,513	116	234,874	257,038	109
Bani Suwayf	133,099	190,176	143	166,576	44,424	27
Al-Fayyum	145,444	173,931	120	182,025	121,448	67
Al-Minya	201,507	238,457	118	-	-	-
Asyut	390,438	247,104	63	488,638	424,512	87
Girga	350,928	192,465	55	-	-	-
Qina and Isna	413,234	219,320	53	-	-	-

Provinces in this table are defined according to: (i) the 1882 census administrative borders for geographic units that were found in the 1882 list, and (ii) the 1848 (1868) census borders for units that were not found in the 1882 list. See text for definitions of true and enumerated populations.

Table 3: Number of Geographic Units with Multiple Enumerations (Duplicates) by Province

Province	Duplicates in 1848	Duplicates in 1868
Cairo	0	0
Alexandria	0	0
Rosetta	0	0
Al-Arish	0	0
Al-Qusayr	0	-
Al-Daqahliya	185	0
Al-Sharqiya	118	0
Al-Qalyubiya	0	-
Al-Gharbiya	15	-
Al-Minufiya	7	-
Al-Buhayra	2	-
Al-Giza	1	2
Bani Suwayf	1	9
Al-Fayyum	1	0
Al-Minya	9	-
Asyut	0	21
Girga	0	-
Qina and Isna	0	-

Provinces in this table are defined according to: (i) the 1882 census administrative borders for geographic units that were found in the 1882 list, and (ii) the 1848 (1868) census borders for units that were not found in the 1882 list. A duplicate is defined as a geographic unit that has more than one census register belonging to *different* years in 1846-48 (for the 1848 census) or 1865-69 (for the 1868 census). Under this definition, a unit that has more than one register that all belong to the same year is *not* a duplicate.

Table 4: Sampling in 1848 and 1868: Geographic Distribution by Province, and Targeted and Actual Sampling Rates

1848/1868 Province	1848				1868			
	Target Pop.	Actual Sample Size	Target Sample Rate	Actual Sample Rate	Target Pop.	Actual Sample Size	Target Sample Rate	Actual Sample Rate
Cairo	255,978	20,635	0.05	0.08	320,360	33,285	0.05	0.10
Alexandria	158,021	16,061	0.05	0.10	197,766	23,617	0.05	0.12
Rosetta	13,233	448	0.01	0.03	16,562	513	0.01	0.03
Al-Arish	2,005	51	0.01	0.03	2,509	26	0.01	0.01
Al-Qusayr	1,659	175	0.01	0.11	-	-	-	-
Al-Daqahliya	397,720	6,374	0.01	0.02	497,752	5,039	0.01	0.01
Al-Sharqiya	259,766	3,012	0.01	0.01	325,100	3,257	0.01	0.01
Al-Qalyubiya	173,804	3,908	0.01	0.02	-	-	-	-
Al-Gharbiya	627,764	7,369	0.01	0.01	-	-	-	-
Al-Minufiya	439,745	5,661	0.01	0.01	-	-	-	-
Al-Buhayra	270,526	3,135	0.01	0.01	-	-	-	-
Al-Giza	187,672	3,509	0.01	0.02	234,874	2,590	0.01	0.01
Bani Suwayf	133,099	1,456	0.01	0.01	166,576	1,568	0.01	0.01
Al-Fayyum	145,444	1,489	0.01	0.01	182,025	2,403	0.01	0.01
Al-Minya	201,507	2,264	0.01	0.01	-	-	-	-
Asyut	390,438	4,309	0.01	0.01	488,638	6,117	0.01	0.01
Girga	350,928	3,540	0.01	0.01	-	-	-	-
Qena & Isna	413,234	4,212	0.01	0.01	-	-	-	-

Provinces are defined according to the 1848/68 censuses administrative borders and division. Target population is calculated based on the province's population share in the 1882 census. Actual sample size is the number of individuals that were actually included in the sample in each province according to the province's definition given above. Target sampling rate is the *a priori* planned sampling rate. Actual sampling rate is the ratio of the actual sample size to the target population in each province.

Table 5: Probit Model for Probability of Enumeration in Rural Provinces in the 1848 and 1868 Censuses

	1848			1868		
	Dependent Variable: Dummy variable for having at least one register in the 1848/1868 census					
log(population)	0.138*** (0.021)	0.314*** (0.025)	0.353*** (0.029)	0.046* (0.027)	0.144*** (0.040)	0.166*** (0.049)
Provincial Town	-0.132 (0.165)	-0.160 (0.195)	-0.222 (0.219)	-0.184 (0.196)	-0.029 (0.289)	0.115 (0.347)
Hamlet or Arab Tribe	1.261*** (0.143)	1.652*** (0.172)	-	1.319*** (0.201)	0.974*** (0.219)	-
Lower Egypt	-0.400*** (0.044)	-	-	-1.037*** (0.052)	-	-
Province Fixed Effects	No	Yes	No	No	Yes	No
District Fixed Effects	No	No	Yes	No	No	Yes
N	3,959	3,959	3,336	3,671	2,115	1,440
Pseudo R^2	0.046	0.297	0.283	0.139	0.424	0.408
Log Likelihood	-2,475.95	-1,823.27	-1,547.17	-1,516.80	-765.83	-575.57
LR Chi-Squared	236.36 (*p=0.000*)	1541.74 (*p=0.000*)	1218.45 (*p=0.000*)	487.94 (*p=0.000*)	1126.29 (*p=0.000*)	791.89 (*p=0.000*)

Standard errors are in parentheses. The observations are the combined list of geographical units of rural provinces in the 1848 (or 1868) census and the 1882 census. Dependent variable takes one if the geographic unit has at least one register in 1848 (1868). Regressors include: (i) log(population) measured (a) using the 1882 census population shares for the units that existed in the 1882 census, and (b) using the enumeration counts for the units of the 1848 (1868) census that were not matched in the 1882 census. (ii) Type: "Village" is the base category. The other two types are provincial town and hamlet or Arab tribe. The type is also measured using the 1882 census type for the units that existed in the 1882 census, or using the 1848 (1868) census type for the units that were not matched in the 1882 census. (iii) Lower Egypt region, province, or district fixed effects. The regression is estimated for each year separately.

Table 6: Descriptive Statistics from the 1848 and 1868 Samples

	1848		1868	
	Cairo and Alexandria	**Other Provinces**	**Cairo and Alexandria**	**Other Provinces**
Dwelling Type				
House/unspecified/unknown	73%	100%	73%	100%
Low-Status dwellings	16%	0%	17%	0%
Multiple-Household dwellings	7%	0%	6%	0%
Production sites	3%	0%	3%	0%
Dwelling Legal Status				
Unspecified	25%	99%	15%	100%
Ownership (public or private)	64%	1%	71%	0%
Waqf (religious endowment)	11%	0%	14%	0%
Gender				
Male	49%	50%	51%	50%
Religion				
Muslim	89%	92%	86%	91%
Christian	6%	6%	4%	7%
Jew	1%	0%	1%	0%
Unspecified	4%	2%	9%	2%
Nationality				
Outside government control	7%	1%	7%	1%
Legal Status				
Slave	2%	0%	3%	1%
Ethnicity				
Local	87%	97%	87%	97%
Turkish	3%	1%	2%	0%
Black	1%	0%	5%	2%
European	1%	0%	1%	0%
Levantine	1%	0%	1%	0%
Nubian	1%	0%	1%	0%
Sample Size	**36,509**	**43,519**	**56,902**	**21,513**

The highest frequencies for each variable are recorded.

Table 7: Sectoral Distribution of Adult Active Male Workers in 1848 and 1868 Egypt

	1848			1868		
	Cairo	**Alexandria**	**Other Provinces**	**Cairo**	**Alexandria**	**Other Provinces**
% Adult active males in the modern sector	9.92	2.2	1.64	2.79	3.61	0.32
N	**6,086**	**4,735**	**9,034**	**9,986**	**6,446**	**4,018**

Sample is restricted to active males who are at least 15 years old with non-missing occupational title.

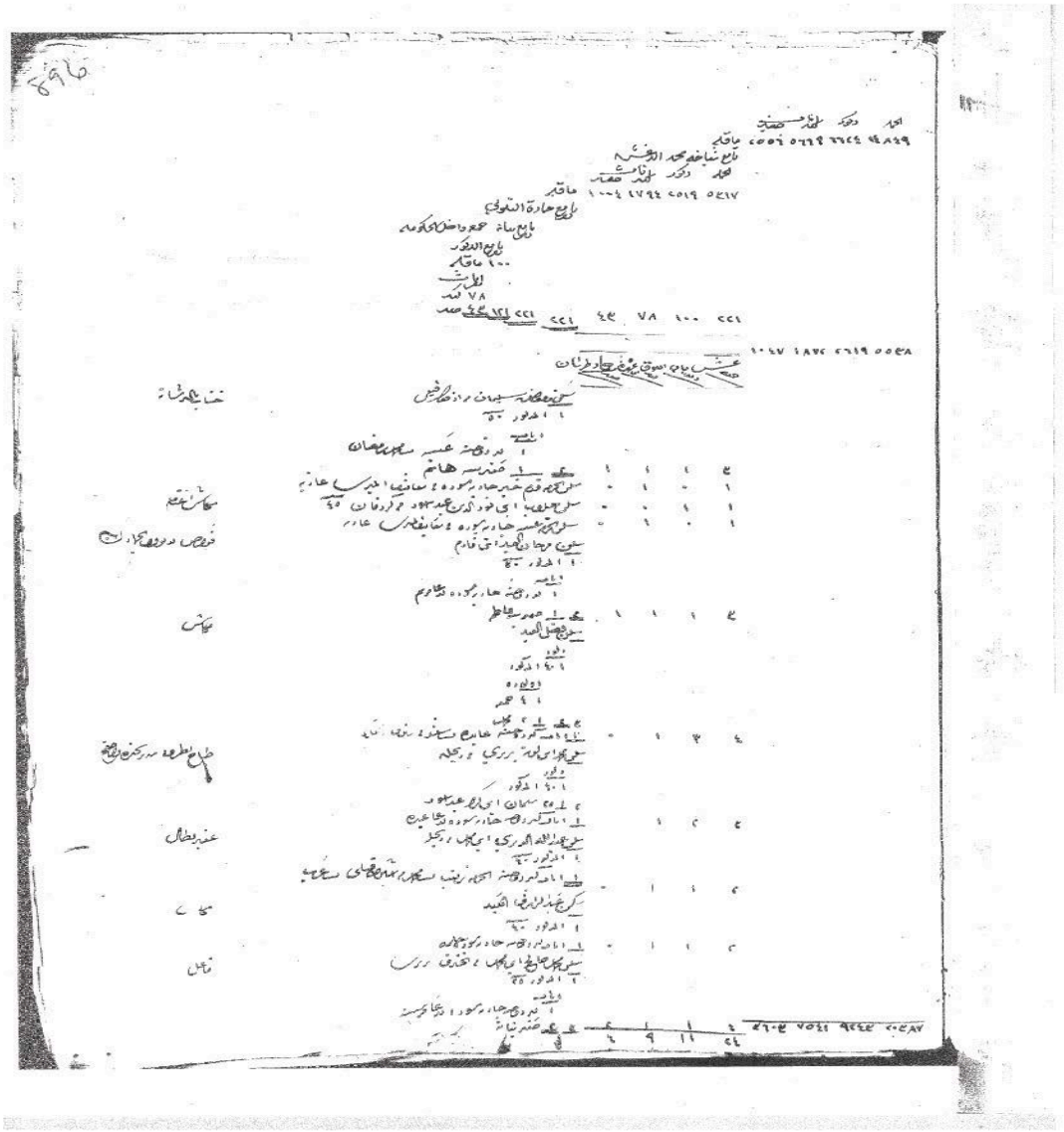
Table 8: Occupational Distribution of Workers in the Modern Sector in 1848 and 1868

	1848		1868	
	Modern	**Traditional**	**Modern**	**Traditional**
Engineers, physicians, pharmacists, ships' masters	1%	1%	4%	1%
Judges, agents, teachers, religious workers, artists	0%	5%	0%	6%
Administrative and managerial workers	2%	0%	2%	0%
Scribes, financiers, stores clerks, customs clerks, post clerks	11%	5%	20%	5%
Merchants, street sellers, auctioneers, water porters, slaves traders	1%	8%	0%	13%
Cooks, servants, slaves, policemen, military, assistants	2%	15%	8%	18%
Farmers, animal husbandry workers, fishermen	0%	35%	0%	14%
Wood treaters, gunpowder makers, textile workers, millers, food processors	10%	11%	5%	14%
Shoe makers, blacksmiths, goldsmiths, silversmiths, potters, stone cutters	5%	4%	6%	6%
Construction workers, building painters, transportation workers, laborers	68%	17%	55%	22%
N (Males at least 15 years old with non-missing occupational title)	**856**	**18,999**	**525**	**19,925**

Sample is restricted to active males who are at least 15 years old with non-missing occupational title.

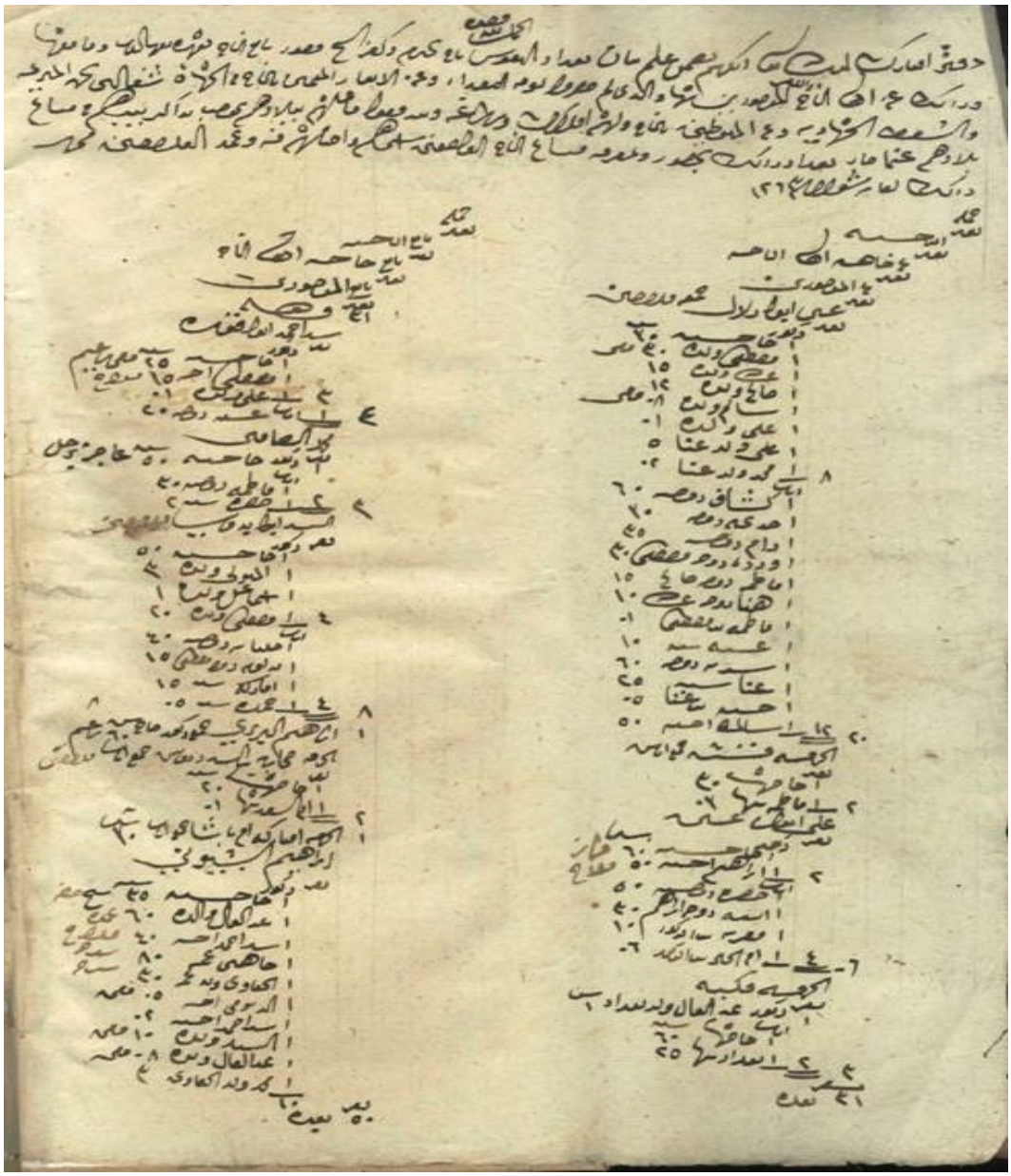
Appendix A
Scanned Pages of the Census Registers

A.1 Page from an Urban Province



Page 896 of Register of *Abdin* District, Cairo Province, 1848

A.2 Page from a Rural Province



Page 1 of Register of the Village of "*Bigirim wa Kafr al-Sheikh Mansour*," *Al-Gharbiya* Province, 1847

Appendix B
Examples of the Digitized Census Samples

1. (Al-Mu'allem) Antonios Luqa: Male, Free, Able-bodied, HH Head, 40 years, Inside the government's control, Coptic Christian, From Abu-Tig (in Abu-Tig, Asyut), Scribe at the Customs Department in Bulaq.

Address: 15 Harat Al-Izba (From Darb Al-Geneina), Shyakhat (Quarter of) Youssef Alam, (District of) Al-Azbakiya, (Province of) Cairo.

Type of Dwelling: House Waqf (Religious Endowment) of Copts.

2. Salman Abdel-Rehim: Male, Free, Able-Bodied, HH Head, 45 years, Inside the government's control, Muslim, From Tukh (in Qena, Qena), Farmer.

Address: House of Salman Abdel-Rehim, Hissat Naser Eleiwa, (Village of) Tukh, (District of) Qena, (Province of) Qena.

3. Mardukh Youssef: Male, Free, Able-Bodied, HH Head, 25 years, Inside the government's control, Jew, From Cairo, Goldsmith.

Address: 35 Harat Al-Yahud Al-Qarra'een (Karaite Jews), Shyakhat (Quarter of) Khidr Ibrahim, (District of) Bab El-Shiriyya, (Province of) Cairo.

Type of Dwelling: House owned by the heirs of Ibrahim Khidr.