Taxing Identity
Theory and Evidence from Early Islam

Mohamed Saleh* Jean Tirole*

March 6, 2021

Webpage Appendix

C Theory

C.1 Threat of rebellion in static model with district heterogeneity

Suppose that in the static analysis the CA is concerned about the possibility of a rebellion driven by high taxes. A successful rebellion overcomes the central and local authorities and eliminates all taxes. A rebellion can succeed only if rebels in the various districts unite. Each agent incurs cost ρ of rebelling; he is willing to rebel if and only if the gain from a successful rebellion exceeds the cost of rebellion:

$$G_i(\theta) \equiv \lambda_i + \min\{\max\{\theta, 0\}, \tau_i\} \geq \rho$$

Summing over all districts, the fraction of potential rebels is

$$\alpha \equiv \int_0^1 [1 - F(\rho - \lambda_i - r_i)] 1_{\{\tau_i > \rho - \lambda_i\}}] di$$

(so $1 - \alpha$ is the fraction of docile agents). We assume that the probability of rebellion H smoothly and strictly increases with the number of potential rebels α.[1]

*Toulouse School of Economics and Institute for Advanced Study in Toulouse, 1 Esplanade de l'Université, Toulouse 31080, France.

[1]This is consistent with many possible stories, with and without coordination failures. To give but one example, we could assume that, after taxes are set but before deciding whether to rebel, the agents learn what it takes for making the rebellion successful; namely, the rebellion will be successful if and only if $\alpha \geq \varepsilon$, where $\varepsilon \geq 0$ is the CA's capacity to counter the rebellion. The parameter ε is ex-ante distributed according to smooth cdf $H(\varepsilon)$. In the absence of coordination failure, a rebellion occurs whenever all those who gain from it have a mass exceeding ε. If there are more agents willing to rebel than is needed

Summing up, the CA's objective function W^r when more broadly rebellion is possible equals the previous expression times the probability $1-H$ that the tax system generates enough docile agents across the territory so as to prevent a rebellion:[2]

$$W^r \equiv [1-H(\alpha)]W$$

where, recall,

$$W = \int_0^1 [\lambda_i + R_i^c(\tau_i)]di.$$

Local authorities' objective functions do not reflect the threat of rebellion, as each district is infinitesimal and can free-ride, counting on the other districts to moderate their taxes so as to limit the threat of rebellion. Their objective function is unchanged.

Proposition 6 ***(cooptation)***
(i) Under the threat of rebellion, the discriminatory tax τ_i is still increasing in the LA's identity strength c_i.
(ii) Under a threat of rebellion the CA induces a reduction in the discriminatory tax rate of counterattitudinal LAs when the cost of rebellion is low, and a reduction in the discriminatory tax rate of soft or zealous LAs when the cost of rebellion is high. The discriminatory tax revenue (R_i) now is inverted-U shaped in the LA's identity strength (c_i), with a peak for a secular LA $(c_i = 0)$.
(iii) Assume that H is indexed by a parameter ν of threat of rebellion: $H(\alpha|\nu)$, with density $h(\alpha|\nu)$ satisfying the monotone hazard rate property (MLRP): $h(\alpha|\nu)/[1-H(\alpha|\nu)]$ is increasing in ν. Then a higher threat of rebellion (a higher ν) makes cooptation more likely.

Remark (agency benefits of counterattitudinal tax collectors). We have assumed that the tax collectors' identity is given by the available local competency pool (which as we discuss in the paper is a reasonable assumption for early Islam in Egypt). Nonetheless, it is useful to examine whether the CA would be willing to incur costs to replace existing tax collectors. Here the predictions are drastically different depending on whether there is a threat of rebellion. In the absence of such a threat, counterattitudinal tax collectors are a nuisance to the CA as they do not convert enough agents. In contrast, counterattitudinal tax collectors may help the CA avert a rebellion as their preferences make them committed to treating agents more leniently. The proof of Proposition 6 shows that the CA may not want to replace a tax collector with one whose preferences are more aligned with his objectives, even if it were costless to do so.

for the rebellion to be successful, an arbitrary selection mechanism will do. Furthermore, exactly the same analysis holds even if there are coordination failures, under the reasonable assumption that the probability of rebellion strictly increases with the number of potential rebels.

[2]This version generalizes that considered in the dynamic framework (see Supplemental Appendix A). There $H(\alpha) = 1$ iff $\alpha > 1 - F(\hat{\theta})$.

Proof of Proposition 6

The CA's welfare is $[1-H(\alpha)]W$, where $[1-H(\alpha)]$ is the probability of staying in power and W the welfare when in power. Let $w_i = \lambda_i + (\tau_i - c)[1-F(\tau_i - r_i)]$ denote the CA's welfare corresponding to district i in the absence of threat of rebellion (so τ_i is as in Proposition 1 and Figure 1), and $\widehat{w}_i(\widehat{\tau}_i) \equiv \lambda_i + (\widehat{\tau}_i - c)[1-F(\widehat{\tau}_i - r_i)]$ denote the CA's welfare when the transfer demand to district i (possibly) accounts for the threat of rebellion (that is, is chosen with an eye on maximizing $[1-H(\alpha)]W$ and not just W). Only values $\widehat{\tau}_i \leq \tau_i$ are relevant when adding a no-rebellion constraint.

If $\lambda_i + \tau_i \leq \rho$, agents in district i will not join a rebellion in the absence of policy change (such districts are "not rebellion-prone"), and so at the optimum $\widehat{\tau}_i = \tau_i$. A district is more likely not to be rebellion-prone, the lower λ_i, r_i, and c_i.

So, we will be interested only in districts such that $\lambda_i + \tau_i > \rho$. For these, either the CA induces $\lambda_i + \widehat{\tau}_i > \rho$; in this case the number of rebels in district i is an exogenous $1-F(\rho - \lambda_i - r_i)$ and the optimal discriminatory tax is still $\widehat{\tau}_i = \tau_i$. Or, by strict quasi-concavity and the property that the optimal rebellion-free tax satisfies $\tau_i \leq \tau_i^a(c)$ (Proposition 1), $\lambda_i + \widehat{\tau}_i = \rho$ and $\widehat{w}_i \equiv \widehat{w}_i(\rho - \lambda_i) = \lambda_i + (\rho - \lambda_i - c)[1-F(\rho - \lambda_i - r_i)] < w_i$. Let $x_i = 1$ if $\lambda_i + \widehat{\tau}_i > \rho$, and $x_i = 0$ otherwise. The probability of staying in power is then $1-H(\alpha)$, where

$$\alpha = \int_{i\in[0,1]} [1-F(\rho - \lambda_i - r_i)] 1_{\{\lambda_i + \tau_i > \rho\}} x_i di$$

The CA solves

$$\max_{\{x_i\in[0,1]\}_{i\in[0,1]}} \left[1-H\left(\int_{i\in[0,1]}[1-F(\rho-\lambda_i-r_i)]1_{\{\lambda_i+\tau_i>\rho\}}x_i di\right)\right] \times \left[\int_{i\in[0,1]}[w_i 1_{\{\lambda_i+r_i\leq\rho\}} + [x_i w_i + (1-x_i)\widehat{w}_i]1_{\{\lambda_i+\tau_i>\rho\}}]di\right]$$

Solving this program (for districts that are rebellion-prone), there exists a parameter $\xi \equiv \left[\frac{h(\alpha)}{1-H(\alpha)}W\right]$ (determined country-wide, i.e. independant of i) such that for districts such that $\lambda_i + \tau_i > \rho$, then $\widehat{\tau}_i = \tau_i$ if and only if the cost of detering rebels in district i relative to the fraction of discouraged rebels in that district exceeeds the country-wide threshold:

$$\xi \leq \frac{w_i - \widehat{w}_i}{1-F(\rho - \lambda_i - r_i)} \tag{C.1}$$

- Suppose, first, that $\lambda_i + \tau_i^m > \rho$. The implementability condition implies that district i cannot be made rebellion free if $c_i \geq 0$. Let $c^* < 0$ be defined by $\tau_i^a(c^*) = \rho - \lambda_i$. Because of the implementability constraint, only districts satisfying $c_i \leq c^*$ can be made rebellion free, so that for $c^* < c_i \leq 0$, the optimal policy remains that under no rebellion threat. For $c_i \leq c^*$, strict quasi-concavity implies that the optimal policy is either $\tau_i^a(c^*)$, or the no-rebellion-threat policy τ_i^m, and (C.1) implies that the choice between the two is the same for all $c_i \leq c^*$.

- Second, assume that $\lambda_i + \tau_i^m \leq \rho$. Let $c^* \geq 0$ be defined by $\tau_i^a(c^*) = \rho - \lambda_i$. If

$c^* > c$, then the optimum is either the same as in the absence of threat of rebellion, or the same for $c_i < c^*$ and $\tau_i^a(c^*)$ for all $c_i \geq c^*$. Next, suppose that $c^* \leq c$. Then, for $c_i \leq c^*$, the no-rebellion policy policy does not generate rebellion in district i ; so the discriminatory tax is unchanged. For $c_i > c^*$, the CA faces a choice between eliminating the threat of rebellion in the district (inducing $\widehat{\tau}_i = \tau_i^a(c^*) = \rho - \lambda_i$) and sticking to the no-rebellion policy τ_i. The CA optimally quells the rebellion in district i if $c_i \in [c^*, c^{**}]$ where $c^* < c^{**} \leq c$ or $c^{**} = +\infty$ (if not, $\widehat{\tau}_i = \tau_i$).

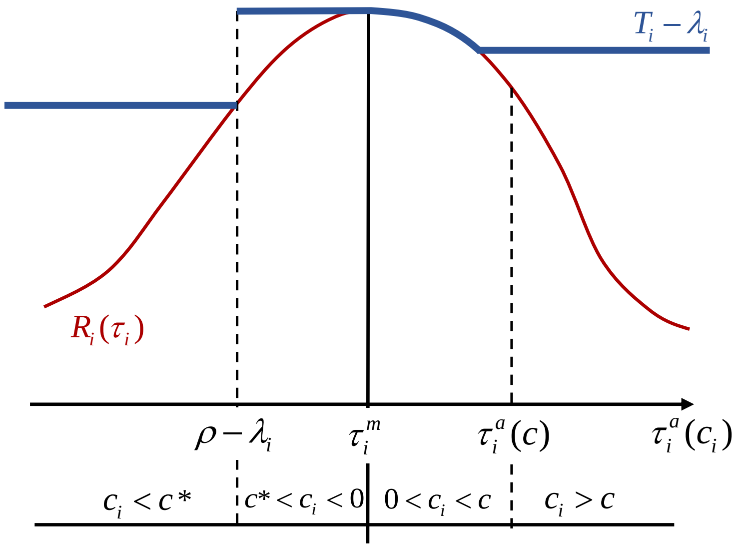

Figure C.1: **Revenue and LA identity strength under the threat of rebellion**

Finally, index the distribution $H(\alpha|\nu)$, where ν is an index of the threat of rebellion. Suppose that ξ does not increase as ν decreases. Then, from (C.1), the set of districts i such that $x_i = 1$ expands, increasing W. Furthermore $h(\alpha|\nu)/[1 - H(\alpha|\nu)]$ increases as well, and so ξ increases, a contradiction. This implies that part (iii) of Proposition 6 holds.

C.2 Internal threats and time-decreasing resistance

Proof of Proposition 5

(i) For all t, the CA chooses $\{\lambda_t(\theta_{t-1}^*), \tau_t(\theta_{t-1}^*)\}$ so as to maximize:

$$W_t = \sum_{k=0}^{+\infty} \beta^k [\lambda_{t+k} + (\tau_{t+k} - c)[1 - F(\theta_{t+k}^*)].$$

$s.t:$

$$\lambda_t \leq \lambda \text{ (uniform tax capped at extractive capacity)}$$
$$\sum_{k=0}^{+\infty} \beta^k [\lambda_{t+k} + (\tau_{t+k} - \hat{\theta}) 1_{\{\hat{\theta} > \theta^*_{t+k}\}}] \leq \frac{\rho - \hat{\theta} 1_{\{\hat{\theta} > \theta^*_{t-1}\}}}{1-\beta} \text{ (no-rebellion constraint)}$$
$$\tau_t \in [\tau^m, \tau^c] \text{ (implementability).}$$

The agents' strategy can be described by $\theta^*_t(\tau_t, \lambda_t, \theta^*_{t-1}) \geq \theta^*_{t-1}$, the cutoff rule at date t (types $\theta \geq \theta^*_t$, and only them, keep their identity up to date t included). Type θ solves

$$U_t = \sum_{k=0}^{+\infty} \beta^k [-\lambda_{k+t} - (\tau_{k+t} - \theta) 1_{\{\theta > \theta^*_{t+k}\}}].$$

Lemma 1 *Suppose that type $\hat{\theta}$ converts at some date $T \in \{1, \ldots, +\infty\}$. From date $T+1$ on, $\lambda_t = \min(\lambda, \rho)$ and $\tau_t = \tau^c$.*

Proof of Lemma 1. Consider an equilibrium path $\{\tau_t, \lambda_t\}_{t \geq 1}$ such that type $\widehat{\theta}$ converts at some date $T \in \{1, \ldots, +\infty\}$. After that date, there is no threat of rebellion provided that for all $t \geq T+1$, $\sum_{k=0}^{+\infty} \beta^k \lambda_{t+k} \leq \frac{\rho}{1-\beta}$. Given that $\lambda_t \leq \lambda$ for all t, from date $T+1$ on, the CA optimally charges $\lambda_t = min(\lambda, \rho)$ and choose τ_t so as to maximize $R^c(\tau_t)$, so $\tau_t = \tau^c$.[3] ||

Lemma 2 *For $t \leq T$, $\theta^*_t = \tau_t$.*

Proof of Lemma 2. The path of conversions if described by a sequence of cutoffs $\{\theta_t\}_{t \in \{1, \ldots, +\infty\}}$, satisfying

$$\theta^*_1 \leq \theta^*_2 \leq \cdots \leq \theta^*_{T-1} < \hat{\theta} \leq \theta^*_T \leq \theta^*_{T+1} \cdots.$$

Suppose that, for some t, $\tau_t > \theta^*_t$, implying that type θ^*_t loses utility at date t from not converting. This utility must be recouped in the future, and so there exists $k \geq 1$ (possibility infinite) such that

$$(\theta^*_t - \tau_t) + \beta(\theta^*_t - \tau_{t+1}) + \ldots + \beta^k(\theta^*_t - \tau_{t+k}) \geq 0.$$

This implies in particular that $\theta^*_{t+1} = \theta^*_t$. In an MPE, this implies that $\theta^*_{t+\ell} = \theta^*_t$, $\tau_{t+\ell} = \tau_{t+1} < \theta^*_t$, $\lambda_{t+\ell} = \lambda_t$ for all $\ell \geq 0$. Suppose, first, that $\tau_t < \theta^*_{t-1}$. Then, for any $\tau'_t \in [\tau_t, \theta^*_{t-1})$, there is no new conversion at date t as any $\theta \geq \theta^*_{t-1}$ enjoys a current surplus, $\theta - \tau'_t > 0$, and keeps an option value. So for $\tau'_t \in [\tau_t, \theta^*_{t-1})$, $\theta^*_t = \theta^*_{t-1}$ and the Markov property implies that the continuation equilibrium remains the same. But with $\tau'_t > \tau_t$, the revenue is higher for the CA. Therefore $\tau_t \geq \theta^*_{t-1}$ for all t. Because θ^*_t cannot recoup the loss in the future, $\tau_t \leq \theta^*_t$.

[3]Recall that $\theta^*_t > \tau^c$ is not implementable for any t because implementability requires that discriminatory taxes be below τ^c in each period and so no type above τ^c would ever convert.

Next, let us show that $\tau_t = \theta_t^*$. Suppose, to the contrary, that $\tau_t < \theta_t^*$. Either $\theta_t^* > \theta_{t-1}^*$, but then type $\theta_t^* - \varepsilon$ should not convert, as $\theta_t^* - \varepsilon - \tau_t > 0$. Or, $\theta_t^* = \theta_{t-1}^*$ and then $\tau_t < \theta_{t-1}^*$, a contradiction. $\|$

Now define $\iota \equiv \arg\max_{\tau \in [\tau^m, \tau^c]}\{-\min\{\tau, \hat{\theta}\} + (\tau - c)(1 - F(\tau))\}$. Also define $\tau^{**} \equiv \max(\tau^m, \tilde{\tau})$, and $\theta^{**} \equiv R^c(\tau^c) - R^c(\tau^{**}) + \tau^{**}$. With simple algebra, we have $\tau^c \geq \theta^{**} \geq \tau^{**}$, and

Lemma 3 *$\iota = \tau^{**}$ if $\hat{\theta} > \theta^{**}$, and $\iota = \tau^c$ if $\hat{\theta} \leq \theta^{**}$.*

Proof of Lemma 3. Either $\tau < \hat{\theta}$ and then the maximand, $(\tau - c)[1 - F(\tau)] - \tau$, is maximized at $\tilde{\tau}$ in the absence of the implementability constraint; so, $\tau = \max\{\tau^m, \tilde{\tau}\} = \tau^{**}$, yielding maximand $-\tau^{**} + R(\tau^{**})$. Or $\tau \geq \hat{\theta}$ and then the maximand, $R^c(\tau) - \hat{\theta}$, is maximized at $\tau = \tau^c$ and then equal to $R^c(\tau^c) - \hat{\theta}$. To see that $\theta^{**} \in (\tau^{**}, \tau^c)$, it suffices to observe that $(R^c)' < 1$ for $\tau > \tilde{\tau}$. $\|$

To find an upper bound for W_1, we first ignore the constraint that $\lambda_t \leq \lambda$ for all λ. This constraint will be satisfied in two cases, and will need to be reintroduced in the third. Finally, consider the date-1 no-rebellion constraint. Rebelling at date 1 yields net cost $(\rho - \hat{\theta})/(1 - \beta)$ to the marginal rebel. Suppose that type $\hat{\theta}$ converts at some date $T \in \{1, \ldots, +\infty\}$. So, it must be the case that

$$\sum_{t=1}^{T-1} \beta^{t-1}[\lambda_t + \tau_t - \hat{\theta}] + \sum_{t \geq T} \beta^{t-1}\lambda_t \leq \frac{\rho - \hat{\theta}}{1 - \beta},$$

where for $t \leq T - 1$, $\tau_t \leq \theta_t^* < \hat{\theta}$ from Lemma 2. So, the CA's date-1 welfare can be bounded above by using, successively, Lemma 2 and the date-1 no-rebellion constraint:

$$\begin{aligned} W_1 &= \sum_{t \geq 1} \beta^{t-1}[\lambda_t + (\tau_t - c)[1 - F(\theta_t^*)]] \\ &= \sum_{t \geq 1} \beta^{t-1}[\lambda_t + R^c(\tau_t)] \\ &\leq \sum_{t=1}^{T-1} \beta^{t-1}[\rho + R^c(\tau_t) - \tau_t] + \sum_{t=T}^{+\infty} \beta^{t-1}[\rho + R^c(\tau^c) - \hat{\theta}]. \end{aligned}$$

Because $\tau_t \geq \tau^m$ and $R(\tau) - \tau$ is maximized at $\tilde{\tau}$, a new upper bound is $\frac{1}{1-\beta}\max\{R^c(\tau^{**}) - \tau^{**}, R^c(\tau^c) - \hat{\theta}\}$. And so:

$$W_1 \leq \begin{cases} \dfrac{\rho - \hat{\theta} + R^c(\tau^c)}{1 - \beta} & \text{if } \hat{\theta} \leq \theta^{**} \\ \dfrac{\rho - \tau^{**} + R^c(\tau^{**})}{1 - \beta} & \text{if } \hat{\theta} \geq \theta^{**}. \end{cases}$$

Next we show that this upper bound is reached for some MPE in the following cases. (a) Suppose that $\hat{\theta} \leq \theta^*$ and that the CA sets $\tau_t = \tau^c$ for all t, $\lambda_t = \min\{\lambda, \rho\}$ for $t \geq 2$ and $\lambda_1 < \min\{\lambda, \rho\}$ such that $\lambda_1 + \frac{\beta}{1-\beta}\min\{\lambda, \rho\} = \frac{\rho - \hat{\theta}}{1-\beta}$ (recall that the no-rebellion constraint is binding, so $\min\{\lambda, \rho\} > \rho - \hat{\theta}$). All agents $\theta \leq \tau^c$ convert at date 1, and no conversion occurs later on. These strategies yields CA welfare equal to

$[\rho - \hat{\theta} + R^c(\tau^c)]/(1-\beta)$ and therefore are optimal for the CA; they also are optimal for the CA from date 2 on, and so form an MPE. If $\hat{\theta} > \tau^c$, then it can be shown that the upper bound on W_1 is the same as for $\hat{\theta} \geq \theta^{**}$ (and the strategies implementing this upper bound are the same as well).

(b) Suppose that $\hat{\theta} \geq \theta^*$ and that the CA set $\tau_t = \tau^{**}$ and $\lambda_t = \rho - \tau^{**} \leq \lambda$ for all t and that $\hat{\theta} \geq \tau^{**}$. Then all conversions occur at date 1 and only types $\theta \leq \tau^{**}$ convert. And the strategies yield the upper bound for W_1 in each period if a fortiori $\hat{\theta} \geq \theta^{**}$.

(c) If the maximum corresponds to τ^{**} and if $\tau^{**} < \rho - \lambda$, we will face a problem when implementing τ^{**} from 1 to $+\infty$, as the per-period uniform tax that would satisfy the no-rebellion constraint would exceed λ, which is impossible. Let us reintroduce the constraint that $\lambda_t \leq \lambda$ in a weaker form:

$$\sum_{t=1}^{+\infty} \beta^{t-1} \lambda_t \leq \frac{\lambda}{1-\beta}.$$

If this constraint is binding, the date-1 no-rebellion constraint becomes:

$$\sum_{t=1}^{T-1} \beta^{t-1} (\hat{\theta} - \tau_t) \geq \frac{\hat{\theta} + \lambda - \rho}{1-\beta}$$

(the RHS of this inequality is by assumption strictly positive). Letting $z_t \equiv 1_{\{\hat{\theta} > \theta_t^*\}} \in \{0,1\}$, and substituting the date-1 no-rebellion constraint,

$$\sum_{t=1}^{+\infty} \beta^{t-1} [\lambda_t + (\tau_t - \hat{\theta}) z_t] \leq \frac{\rho - \hat{\theta}}{1-\beta},$$

$$W_1 \leq \sum_{t=1}^{+\infty} \beta^{t-1} [R^c(\tau_t) - (\tau_t - \hat{\theta}) z_t + \rho - \hat{\theta}].$$

Maximize over z_t and τ_t the RHS of this inequality subject to the constraint coming from the upper bound on the uniform tax:

$$\sum_{t=1}^{+\infty} \beta^{t-1} (\hat{\theta} - \tau_t) z_t \geq \frac{\hat{\theta} + \lambda - \rho}{1-\beta} \qquad (\mu)$$

The period-by-period maximization amounts to solving

$$\max_{\{\tau_t, z_t\}} R^c(\tau_t) - (\tau_t - \hat{\theta})(1+\mu) z_t$$

and so τ_t and z_t are both constant over time (call these τ and z). Furthermore

$$z = 1 \iff (\hat{\theta} - \tau)(1+\mu) > R^c(\tau^c) - R^c(\tau).$$

When the constraint is non-binding ($\mu = 0$), then the solution is as in cases (a) and (b). When it is binding

$$\hat{\theta} - \tau = \hat{\theta} + \lambda - \rho \iff \tau = \rho - \lambda.$$

And so, letting $\tau^* \equiv \max\{\tau^m, \tilde{\tau}, \rho - \lambda\}$ and $\theta^* \equiv \tau^* + [R^c(\tau^c) - R^c(\tau^*)]$ the solution is the same as in Proposition 4, except for the sequencing of uniform taxes in case (a).

(ii) Suppose that $\hat{\theta} < \theta^{**}$. Could a coalition of size (at least) $1 - F(\widehat{\theta})$ coordinate and not convert at date 1, so that the rebellion constraint would remain at date 2? Could it do so repeatedly? Let the CA set $\tau_2 = \tau^c$ and $\lambda_2 = \lambda_1$ and continue doing so as long as type $\widehat{\theta}$ (and types below necessarily) has not converted[4]. From date 2 on, type $\widehat{\theta}$ in this deviation from equilibrium behavior obtains value function $V_1 \equiv \frac{\widehat{\theta}-\rho}{1-\beta}$, i.e. its date-1 value function. So, type $\widehat{\theta}$ does not want to deviate at date 1 if $-\lambda_1 - \tau^c + \beta V_1 \leq V_1$ or $\tau^c \geq \beta \frac{\widehat{\theta}+\lambda-\rho}{1-\beta}$. ∥

Context	Tax	Low threat of rebellion	High threat of rebellion	Intensity of threat
		τ^m τ^*	θ^* τ^c →	$\hat{\theta}$
Statics	Uniform	$\hat{\lambda} = \rho - \hat{\theta}$	$\hat{\lambda} = \rho - \tau^*$	
	Discriminatory	$\hat{\lambda} = \tau^c$ (high)	$\hat{\tau} = \tau^*$ (low)	
Myopic (β = 0)	Uniform	$\lambda_1 = \rho - \hat{\theta} < \lambda_t = \min\{\rho, \lambda\}$ for $t \geq 2$	$\lambda_t = \rho - \tau^*$	
	Discriminatory	$\tau_t = \tau^c$	$\tau_t = \tau^*$	
Far-sighted (β > 0)	Uniform	$\lambda_1 = \rho - \hat{\theta} - \dfrac{\beta\lambda}{1-\beta} < \lambda_t = \min\{\rho, \lambda\}$ for $t \geq 2$	$\lambda_t = \rho - \tau^*$	
	Discriminatory	$\tau_t = \tau^c$	$\tau_t = \tau^*$	

Figure C.2: **Threat of rebellion (summary of Propositions 4 through 5)**

Finally, we note that the extraction model, which is a special case of the identity-based model[5], exhibits the same pattern regarding decreasing resistance:

Corollary 1 ***(comparison with the extraction model).*** *Under Assumptions 1 and 2,*
(i) If $\widehat{\theta} < \tau^m$, the marginal rebel is a convert; the discriminatory tax is equal to its extractive level τ^m in all periods, while the uniform tax is raised over time from $\rho - \widehat{\theta}$ in the first period to max$\{\rho, \lambda\}$ thereafter.
(ii) If $\widehat{\theta} \geq \tau^m$, the marginal rebel is a non-convert; the discriminatory tax is equal to its extractive level τ^m in all periods, while the uniform tax is constant at level $\rho - \tau^m < \lambda$ over time.

[4]Observing the discriminatory tax volume supplies this information; indeed, we have assumed that types are not observable.

[5]It satisfies in particular $\tau^* = \tau^m$ when the threat of rebellion is binding.

C.3 Other extensions

(a) *Discrimination through access to public goods*. When direct discrimination is prohibited by the constitution or a higher-level polity (which was not the case for early Islam), we naturally observe more indirect forms of discrimination, such as neighborhood-based access to public goods, ethnicity-based patronage and incendiary rhetoric. Glaeser and Shleifer (2005) describe such forms of discrimination in 20$^{\text{th}}$-century US, staging an Irish-catholic/Anglo-Saxon-protestant conflict in Boston and a black/white conflict in Detroit. In both examples, the mayor induced over the years substantial emigration of the minority out of the city, reinforcing the incumbent's political power;[6] Glaeser and Shleifer call this the "Curley effect," after the name of a Boston mayor who was in power for most of the 1913-1951 period. A direct, ethnic or race-based, tax discrimination being prohibited by the federal government, the ruler's hostility toward the minority shifted to presumably less efficient forms of utility extraction. Their paper also documents Robert Mugabe's tactic in Zimbabwe, which led to substantial migration by white farmers. In either case, more discrimination involved a revenue cost, in terms of either migration or incompetent management (patronage). And it increased the probability of a rebellion.

The trade-off between loss in revenue and preference alignment through emigration of members with a dissonant identity also arises in modern democracies when a ruler may also want to increase the cohesiveness of the polity. Democratic regimes and organizations sometimes function more efficiently when their membership is more homogeneous. For example, Hansmann (1996) argues that congruence in objectives facilitate both the flow of information and the fluidity of decision making in cooperatives. Besley et al. (2017) argue that districts with single party majority yield more cohesive policies, presumably because this cohesion facilitates agreement on the use of tax revenue and thereby raises incentives to collect tax revenue. Relatedly, Alesina et al. (1999) have shown that the provision of local public goods is facilitated by religious or ethnic homogeneity. Without applying a value judgment to such objectives, we can capture the ruler's demand for cohesiveness within the function.

(b) *Discriminatory empathy.* Suppose that LA i puts (positive or negative) weight $w_i(\theta)$ on type-θ agent's utility, where $w_i'(\theta) \leq 0$ and $\int_{-\infty}^{+\infty} w_i(\theta) dF(\theta - r_i) \equiv \bar{w}_i < 1$.[7] The

[6] Migration then reduces resistance to the ruler over time because of the majoritarian electoral system. By contrast, our time-decreasing resistance in Section 4.1 will be based on a reduced stake for the converts.

[7] The LA need not observe individual agents' types to form such preferences (and actually it does not).

LA's objective function is then

$$V_i \equiv [\lambda_i + \tau_i[1-F(\tau_i - r_i)] - T_i] + \Big[\int_{-\infty}^{\tau_i} w_i(\theta)(-\lambda_i)dF(\theta - r_i) + \int_{\tau_i}^{+\infty} w_i(\theta)[-\lambda_i + (\theta - \tau_i)]dF(\theta - r_i)\Big] = (1-\bar{w}_i)\lambda_i + R_i^a(\tau_i) - T_i$$

where

$$R_i^a(\tau_i) \equiv \tau_i[1-F(\tau_i - r_i)] + \int_{\tau_i}^{\infty} w_i(\theta)(\theta - \tau_i)dF(\theta - r_i).$$

Note that

$$(R_i^a)'(\tau_i) = [1 - w_i^+(\tau_i)][1-F(\tau_i - r_i)] - \tau_i f(\tau_i - r_i)$$

where $w_i^+(\tau_i) \equiv E[w_i(\theta)|\theta \geq \tau_i] \leq \bar{w}_i$.

The difference with the model in the text is that $c_i \equiv \frac{-\tau_i w_i^+(\tau_i)}{1-w_i^+(\tau_i)}$ depends on the discriminatory tax, which itself depends on the extent of discriminatory empathy. Substantial hostility to the high-identity agents is then required to be on the wrong side of the Laffer curve. More generally, lower empathy (w_i^+ falls) implies a higher discriminatory tax.

(c) *Social incentives: norms and network externalities.* When contemplating changing his identity, an agent may take into account not only his own preferences (θ) and the material incentive (τ_i), but also the resulting perception of his choice within his community. Letting $F_i(\theta) \equiv F(\theta - r_i)$, suppose that the potential convert has image concerns $\mu M_i^+(\theta_i^*) = \mu E_{F_i}[\theta|\theta \geq \theta_i^*]$ if he does not convert and $\mu M_i^-(\theta_i^*) = \mu E_{F_i}[\theta|\theta \leq \theta_i^*]$ if he does, where θ_i^* is the threshold type in district i and $\mu \geq 0$ is a parameter of intensity of image concerns. $M_i^+(\theta_i^*)$ and $M_i^-(\theta_i^*)$ are the upward and downward truncated means, respectively (i.e. the expectations of θ conditional on θ being above or below θ_i^*). The cutoff θ_i^* (or alternatively the tax $\tau(\theta_i^*)$ that induces θ_i^*) is then given by

$$\theta_i^* - \tau_i + \mu\Delta(\theta_i^* - r_i) = 0,$$

where $\Delta(\theta^*) \equiv E_F[\theta|\theta \geq \theta^*] - E_F[\theta|\theta < \theta^*]$. The co-variation of the threshold and the discriminatory tax is no longer 1 for 1 if $\mu > 0$, and is given by:

$$\frac{d\tau_i}{d\theta_i^*} = 1 + \mu\Delta'(\theta_i^* - r_i).$$

Let us assume that image concerns are not too large, $1 + \mu\Delta'(\theta_i^* - r_i) > 0$, and so the equilibrium threshold is unique and $\tau(\theta_i^*)$ well-defined. The new revenue function is $\hat{R}_i(\tau_i) \equiv \tau_i[1 - F(\theta_i^*(\tau_i) - r_i)]$. The analysis is unchanged, except that now LA i's objective function is:

$$V_i = \lambda_i + (\tau_i - c_i)[1 - F(\theta_i^*(\tau_i) - r_i)] - T_i.$$

Introducing social pressure adds a few interesting additional insights, though. If the distribution $f(\cdot)$ is unimodal, the function $\Delta(\theta^*)$ is U-shaped. When conversions are rare, the reputational concern is driven mainly by the strong stigma attached to conversions (and so $\Delta'(\theta_i^* - r_i) < 0$). The discriminatory tax has a strong impact on the threshold because it not only provides a material incentive for conversion, but it also releases the social stigma attached to conversions. When in contrast there are few Copts remaining, reputational concerns are mainly driven by the social prestige attached to resistance (and so $\Delta'(\theta_i^* - r_i) > 0$); the discriminatory tax impact on the threshold is then less than 1 for 1.[8] The model can also easily be extended to allow for *network externalities*.

(d) *Malthusian ruler.* Suppose now that agents care not only about consumption and identity, but also about the number of their children. We use a model à la Galor and Weil (2000) and enrich it through an identity decision. A district-i agent's utility is[9]

$$U(\theta) = \max_{z \in \{0,1\}} \frac{\rho^{1-\alpha}}{\alpha^\alpha (1-\alpha)^{1-\alpha}} a^\alpha n^{1-\alpha} + \theta z$$

s.t.

$$a + \rho n \leq y - \lambda_i - \tau_i z,$$

where z equals 1 if the agent maintains his identity and 0 if he converts, a is consumption, n the number of children, y the endowment, ρ the cost of a child's upbringing, and $\alpha \in (0,1)$. Hence

$$U(\theta) = y - \lambda_i + (\theta - \tau_i) z,$$

which yields, as in the model without fertility choice, cutoff

$$\theta^* = \tau_i.$$

Suppose now that LA i is motivated to reduce the number of non-converts:

$$V_i = \lambda_i + (\tau_i - c_i)[1 + \nu n_i][1 - F(\tau_i - r_i)] - T_i$$

where some weight $\nu > 0$ is put on the indirect conversions (of children). Let us show that n_i is a decreasing function of τ_i. A non-convert's number of children is given by $\rho n_i = (1-\alpha)(y - \lambda_i - \tau_i)$. So n_i is a decreasing function of τ_i. Note that the LA, when raising the poll tax, achieves double benefits: directly by inducing the adult generation to convert, and indirectly by making holdouts poorer and therefore reducing their reproductive rate. Webpage Appendix Section D.1 fails to find empirical support for this indirect mechanism in our historical context, but it might be relevant to other contexts.

D Empirics

D.1 Conversion or demographic Islamization?

An alternative theory of Egypt's, and the region's, Islamization traces the process to population replacement, in the sense that Arabs (Muslims) replaced the local non-Muslim populations of the region, rather than to conversions to Islam among the local

[8] One can go further in the elasticity analysis by assuming that $\Delta''(\theta_i^* - r_i) > 0$ (a hypothesis for which Jia and Persson (2017) find supporting evidence in a different context).

[9] In this version, the agent cares about his own identity or, alternatively, about the identity of his dynasty.

populations. In the absence of Copts' conversion to Islam, five demographic processes could have driven the decline in Egypt's non-Muslim population share between 641 and 1200, and subsequently through 1848 (Fargues, 2001):[10] Muslim immigration into Egypt, Coptic emigration, Muslims' higher fertility (net of child mortality), Muslims' lower adult mortality, and intermarriage between Coptic females and Muslim males (the opposite scenario is prohibited) without pre-marriage conversion, which results by law in a Muslim offspring.[11] These processes, we argue, are *not* the main causes of Islamization.

Muslim immigration. Arab immigration, the largest Muslim immigration wave in Egypt between 641 and 1200, was small compared to the Egyptian (Coptic) population. In 641, Egypt's population (2.7 million) was three times that of the Arab peninsula (1 million) (Russell, 1958, p. 89). Russell (1966) estimates the number of Arab immigrants in 650 at 100,000. Furthermore, Arab immigration subsided after 833 with the shift to recruiting slave armies and the stoppage of state stipends to Arabs, which led Arabs to lose their military aristocratic position to Turks. It is also important to note that if Arab immigration were the sole driver of the decline in Egypt's non-Muslim population share between 641 and 1200, we would normally expect Arabs (Muslims) to be better off, on average, than Copts, because Arabs dominated by law the top white-collar positions in the military, judiciary, police, and the high-level bureaucracy, and because Copts were subject to a higher tax. This prediction contradicts though the papyrological evidence in 641-969 that shows that Copts were better off than Muslims; they were over-represented among white-collar workers and artisans and under-represented among farmers and unskilled non-agricultural workers (Saleh, 2018).

Copt emigration. Copts rarely emigrated from Egypt, because of their unique Christian denomination that differed from both Catholics and Greek Orthodox Christians. Until today, Coptic Christianity has been considered a "heretical" "non-Chalcedonian" Oriental Orthodox Christian denomination, which split from the Catholic Church at the Council of Chalcedon in 451. Egypt's Chalcedonian Christians, who remained loyal to the Roman/Byzantine Church, formed a small minority called the *Melkites*.

Coptic-Muslim fertility difference. Even if Arab immigration was small compared to Egypt's population, Arabs (Muslims) could have gradually replaced Copts over time if they had more children.[12] While this alternative hypothesis (which rules out Copt

[10]This section draws on and expands the discussion in Saleh (2018, pp. 425-426).

[11]A marriage in which a Coptic male converts to Islam prior to marriage is excluded because the mechanism of converting the offspring in this case is paternal conversion, and not cross-marriage per se.

[12]Recall that Copts were about 2.7 million in 641, and that Arabs were about 100,000 in 650. In 1200, Egypt's population was 2.3 million, with Muslims constituting 84% (1.9 million) and Copts 16% (0.3

conversions to Islam) still does not explain why Copts were better off than Muslims as early as in 641-969, we attempt to test it directly using the 1848 and 1868 census samples which were digitized by Saleh (2013). Because these censuses predate Egypt's demographic transition, which started in the second half of the twentieth century, they provide a glimpse of the demographics of medieval (Malthusian) Egypt. They also allow us to measure the number of *surviving* children, which is arguably a better measure of the *desired* number of children than the number of children *ever born*, which we do not observe. Specifically, our measure is fertility net of child mortality: the number of surviving children below 10 years and below 1 year. Measuring fertility from the population censuses is subject to two caveats, though: (1) We only observe children who reside with their parent(s) at the time of the census. But this is less of a concern for children below 10, who are more likely to live with their parent(s). (2) We do not observe the father and mother of every individual in the censuses (except for children of the household head), but we inferred the (potential) father and mother from the relationship to the household head (the household structure). The findings in Table D.1 reveal that Muslim males do *not* have more surviving children than Coptic males, whether we count the number of surviving children below 10 years of age or below 1. This null finding holds within each occupational group: unskilled non-agricultural workers, farmers, artisans, and white-collar workers. Furthermore, Muslim females have *fewer* children under 10 than their Coptic counterparts, especially in households headed by farmers and white-collar workers, but the difference is statistically insignificant if we measure fertility by the number of surviving children under 1 (except for females in households headed by white-collar workers).

Coptic-Muslim adult mortality difference. Measuring adult mortality from the population censuses is more challenging, because we do not observe deaths. Saleh (2018) measures adult life expectancy among Copts and Muslims by comparing the age distribution between 1848 and 1868. The findings in Table D.2 (taken from the Online Appendix of Saleh (2018)) show that Muslims had lower adult mortality (higher life expectancy) at younger ages (10-29 or 10-39), but higher adult mortality (lower life expectancy) at older ages (30-79 or 40-79). However, the differences are small in mag-

million). Let's assume that (1) a generation lasts for 30 years, which is the typical life expectancy at birth in pre-industrial populations, and hence there were 18.66 generations in 641–1200, (2) one third of children died before adulthood, which is the typical pre-industrial child mortality, and hence a woman must give birth to 3 children, on average, in order to have a zero population growth rate per generation. So if the change in Egypt's ethno-religious composition in 641–1200 is due to the fertility mechanism alone, without any conversions to Islam or immigration, this implies that (1) Copts shrank at 10% per generation in 641–1200, whereas Arabs grew at 17% per generation during the same period, (2) Copts and Arabs must have had gross fertility rates of 2.7 and 3.5 children per woman, respectively.

Table D.1: **Coptic-Muslim fertility difference in 1848 and 1868**

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Children <10	Children <10	Children <1	Children <1	Children <10	Children <10	Children <1	Children <1
Copt	-0.050	-0.128	0.015	0.012	0.159	0.085	0.034	0.013
	(0.067)	(0.120)	(0.029)	(0.052)	(0.069)**	(0.075)	(0.023)	(0.022)
Farmer		0.067		0.050		0.243		0.073
		(0.062)		(0.015)***		(0.037)***		(0.012)***
Artisan		-0.070		-0.027		0.374		0.091
		(0.092)		(0.025)		(0.101)***		(0.029)***
White-collar		0.424		0.086		0.109		0.032
		(0.090)***		(0.030)***		(0.085)		(0.013)**
Copt * Farmer		0.261		-0.036		0.320		0.022
		(0.153)*		(0.048)		(0.121)***		(0.037)
Copt * Artisan		0.042		0.049		-0.223		-0.005
		(0.228)		(0.069)		(0.206)		(0.063)
Copt * White-collar		-0.118		0.012		0.373		0.147
		(0.188)		(0.066)		(0.213)*		(0.061)**
Constant	1.836	1.768	0.328	0.301	1.198	1.120	0.211	0.188
	(0.036)***	(0.059)***	(0.012)***	(0.011)***	(0.022)***	(0.027)***	(0.007)***	(0.006)***
Obs (individuals)	22119	22119	22119	22119	14780	14780	14780	14780
Clusters (districts)	106	106	106	106	98	98	98	98
R^2	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.01
Mean dep. var.	1.54	1.54	0.23	0.23	1.20	1.20	0.21	0.21

Notes: Robust standard errors clustered at the district level are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. The omitted group is unskilled non-agricultural Muslim workers.
Source: The 1848 and 1868 population census samples (Saleh, 2013) and an over-sample of non-Muslims in Cairo in 1848 and 1868. Census samples are pooled and restricted to Copts and Muslims aged 15 to 60 years. Regressions are weighted by sample design. Because almost all females have missing occupations, we assigned the household head's occupational title to all household members with missing occupations, including females. Number of children is inferred from the relationship to the household head, and includes only surviving children residing with their parent(s) at the time of the census.

nitude, and may be attributable to statistical caveats in the 1848 and 1868 censuses, namely, (1) the gap (20 years) that separates the two censuses is longer than ideal (5 or 10 years) as it increases the chance of population movement, and (2) age heaping (tendency to report age as a number ending in "0" or "5") and age exaggeration (for older individuals); since both phenomena are negatively correlated with socioeconomic status, they are less prevalent among Copts.

Cross-marriages without pre-marriage conversion. Another way of replacing the Coptic population is by Arab (Muslim) males marrying (possibly more than one) Coptic females, as the off-spring in this case will be Muslim. Cross-marriages between Muslim

Table D.2: **Estimating adult life expectancy from the 1848 and 1868 population 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. We applied the two methods to the census samples of 1848 and 1868, in order to estimate 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) we 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 socioeconomic status 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, we defined age groups in intervals of 10 years instead of 5 years.
Source: The 1848 and 1868 population census samples. This table is reproduced from the Online Appendix of Saleh (2018).

males and Coptic females were rare as suggested by the dearth of cross-marriage contracts in the papyri in 641-969. The 1848 and 1868 population census samples record only two cross-marriages.

D.2 Additional figures and tables

Descriptive figures. Figure D.1 shows examples of the secondary sources, Morimoto (1981, pp. 67-79, 85-87) and the Arabic Papyrology Database that we used to construct our dataset on the annual poll tax payments at the individual taxpayer level in 641–1100. Figure D.2 shows the histogram of the annual poll tax payments by *kura* in our dataset. Figure D.3 shows the evolution of Caliph-level piety (non-drinking alcohol) and governor-level hostility in 641–1170.

Table I Register of Gold-Tax Assessment for "Five Fields",
The 3rd Indiction (704/05 : A.H. 85/86)

Taxpayers	Location of fields	Taxes		Taxes		Total sol.	Corn tax art.
		Land sol.	Corn art.	Land sol.	Poll sol.		
Mēnas Apollōs	Belekau			½	3	3½	½
Kaumas Antheria	Sarseltōh			2½	2½	5	3
Psoios Andreas	Pkathakē Pkarou	1 ⅙	1 ⅓	1⅙	1½	2⅔	1⅓
Horsenuphios Hermaōs	Ammōniu Pankul & others Piah Alau	8½ 1½ ½	10 1½ 1	10½	4	14½	12½
Abraham Theodosios	Piah Boōn Piah Kam Hagiu Biktōr	½ 1 1	½ 1½ 1	2½	4½	7	3
Bethanias Pkaloos	Pkarou			⅓	0	⅓	⅓
Taam, Johannes Th[ ]-liaie & Eudoxia	Pkarou & Belekau			2⅓	0	2⅓	2⅓

P.Cair.Arab. 202 : List of poll-tax payers. (unknown (Egypt); 1. 1. 801 - 31. 12. 900 CE)

Inv. No.: **P.Cair.EgLib.inv. 293 recto and verso** Material: **papyrus** Height: **19 cm.** Width: **12 cm.** Language(s): **Arabic** Kind: **List, account** Edition(s): **Grohmann, P.Cair.Arab. 202** Translation(s): **Grohmann, P.Cair.Arab. 202**

recto 1 *** سدس وثمن شنوده القرّاش نصف وربع
recto 2 *** وثمن بسنّة كرسدودره ثلثة دنانير ونصف وثلثمن
recto 3 *** وسدس كيل بسنّه ثلثي وسدس ثمن
recto 4 *** دينر أدّى سويرس أبلّوا ثمن دينر

Figure D.1: **Examples of the secondary sources of the poll tax sample**

Top: Morimoto (1981, p. 67): Register of "Five Fields" in *Aphrodito* in 704/05 CE. Bottom: Arabic Papyrology Database: List of poll-tax payers in 801-900.

Robustness checks: Full results tables. Appendix Tables D.3-D.18 show the full results of the robustness checks that are summarized in the Supplemental Appendix Tables.

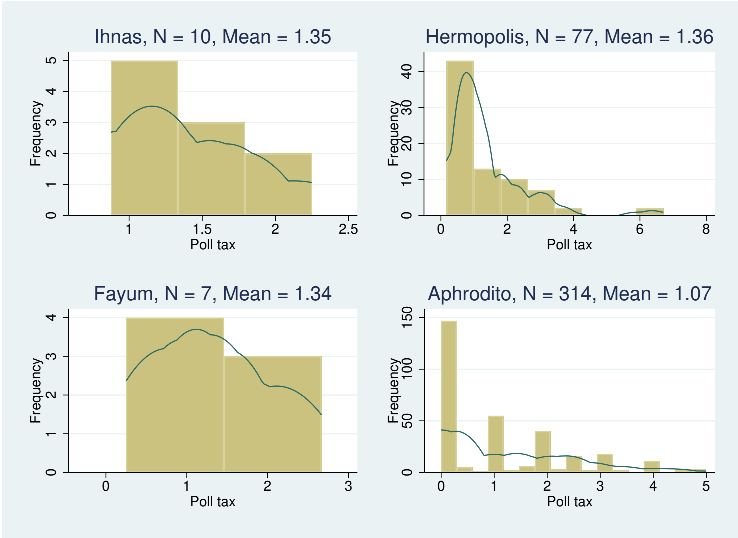

Figure D.2: **Histogram of individual poll tax payments by *kura* in 641–1100**

Notes: Arab settlement =1 in *Ihnas*, *Hermopolis*, and *Fayum*, and =0 in *Aphrodito*. Sources: Greek papyri in Morimoto (1981, pp. 67-79, 85-87) and Arabic papyri in the Arabic Papyrology Database. Sample is restricted to papyri from a known location.

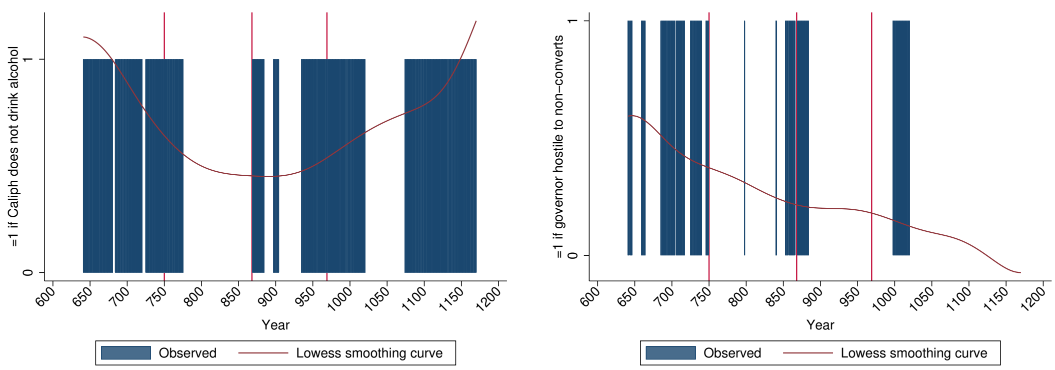

Figure D.3: **Central authority's identity strength in 641–1170**

The vertical red lines at years 750, 868, and 969, indicate major dynastic changes. 641–750: Rashidun and Umayyads; 750–868: First Abbasid Period; 868–969: Tulunids, Second Abbasid Period, Ikhshidids; 969–1170: Fatimids. Source: See the Supplemental Appendix.

Table D.3: **Local determinants of conversions to Islam in 641–1200: Amélineau's Byzantine-period villages only**

Dependent variable: =1 if no Coptic church or monastery in village j circa 1200

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.04		0.06	0.08	0.12	0.01	0.00	-0.01	0.07
	(0.09)		(0.08)	(0.08)	(0.07)*	(0.15)	(0.13)	(0.14)	(0.12)
=1 if HF visit (r_{ji})		-0.60	-0.61	-0.61	-0.58		-0.60	-0.60	-0.58
		(0.04)***	(0.05)***	(0.05)***	(0.07)***		(0.05)***	(0.05)***	(0.06)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	163	163	163	163	157	163	163	163	157
Clusters (*kuras*)	39	39	39	39	37	39	39	39	37
R^2	0.00	0.08	0.08	0.09	0.14	0.00	0.08	0.08	0.14
Mean dep. var. in control	0.54	0.60	0.54	0.54	0.54	0.54	0.54	0.54	0.54
KP Wald F-stat						12.98	13.33	19.00	15.71

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.10		0.13	0.15	0.16	0.05	0.07	0.11	0.08
	(0.09)		(0.08)	(0.08)*	(0.08)**	(0.20)	(0.17)	(0.17)	(0.15)
=1 if HF visit (r_{ji})		-0.61	-0.64	-0.63	-0.60		-0.63	-0.63	-0.59
		(0.06)***	(0.07)***	(0.07)***	(0.08)***		(0.06)***	(0.06)***	(0.07)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	163	163	163	163	157	163	163	163	157
Clusters (*kuras*)	39	39	39	39	37	39	39	39	37
R^2	0.02	0.09	0.11	0.11	0.15	0.02	0.10	0.11	0.14
Mean dep. var. in control	0.54	0.60	0.54	0.54	0.54	0.54	0.54	0.54	0.54
KP Wald F-stat						8.42	8.13	9.86	16.33

Notes: Robust standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A constant is included in all regressions.
Source: See the Supplemental Appendix.

Table D.4: **Local determinants of conversions to Islam: Alternative measure of F_{ji} in 1848–1868**

Dependent variable: =1 if Coptic Christian in 1848–1868

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.07		-0.07	-0.07	-0.03	-0.12	-0.11	-0.12	-0.05
	(0.04)*		(0.04)*	(0.04)*	(0.02)	(0.04)***	(0.04)***	(0.04)***	(0.02)***
=1 if HF visit (r_{ji})		0.18	0.18	0.18	0.14		0.17	0.16	0.13
		(0.07)***	(0.06)***	(0.06)***	(0.05)***		(0.06)***	(0.05)***	(0.05)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (individuals)	16641	16641	16641	16641	16195	16641	16641	16641	16195
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.02	0.01	0.03	0.03	0.07	0.01	0.02	0.02	0.07
Mean dep. var. in control	0.12	0.06	0.12	0.12	0.12	0.12	0.12	0.12	0.12
KP Wald F-stat						8.22	8.22	7.66	5.98

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.03		-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.05
	(0.02)**		(0.02)*	(0.02)*	(0.01)***	(0.02)	(0.02)	(0.02)	(0.02)**
=1 if HF visit (r_{ji})		0.09	0.10	0.11	0.12		0.10	0.11	0.13
		(0.06)	(0.06)	(0.06)*	(0.04)***		(0.06)*	(0.06)**	(0.04)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (individuals)	16641	16641	16641	16641	16195	16641	16641	16641	16195
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.08	0.08	0.08	0.08	0.09	0.08	0.08	0.08	0.09
Mean dep. var. in control	0.12	0.06	0.12	0.12	0.12	0.12	0.12	0.12	0.12
KP Wald F-stat						5.75	5.73	5.27	5.33

Notes: Robust standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions.

Source: Individual-level data on religious affiliation in 1848–1868 are from the 1848 and 1868 individual-level population census samples restricted to Egyptian local free Coptic and Muslim employed men of a rural district of origin who are at least 15 years of age and with non-missing information on age, religion, occupation, and district of origin.

Table D.5: **Local determinants of conversions to Islam: Alternative measure of F_{ji} circa 1500**

Dependent variable: =1 if no Coptic church or monastery circa 1500

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.03		0.03	0.04	0.04	0.03	0.03	0.03	0.04
	(0.03)		(0.02)	(0.02)	(0.01)***	(0.02)	(0.02)	(0.02)	(0.01)***
=1 if HF visit (r_{ji})		-0.28	-0.28	-0.28	-0.28		-0.28	-0.28	-0.28
		(0.07)***	(0.07)***	(0.07)***	(0.07)***		(0.07)***	(0.07)***	(0.07)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1782	1782	1782	1782	1751	1782	1782	1782	1751
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.01	0.04	0.05	0.05	0.12	0.01	0.05	0.05	0.12
Mean dep. var. in control	0.95	0.98	0.95	0.95	0.95	0.95	0.95	0.95	0.95
KP Wald F-stat						17.23	17.33	15.94	17.09

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.03		0.03	0.03	0.04	0.01	0.01	0.01	0.03
	(0.02)*		(0.01)*	(0.01)*	(0.01)***	(0.01)	(0.01)	(0.01)	(0.01)***
=1 if HF visit (r_{ji})		-0.24	-0.24	-0.24	-0.27		-0.24	-0.24	-0.27
		(0.07)***	(0.07)***	(0.07)***	(0.08)***		(0.07)***	(0.07)***	(0.07)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1782	1782	1782	1782	1751	1782	1782	1782	1751
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.07	0.10	0.10	0.11	0.13	0.07	0.10	0.10	0.12
Mean dep. var. in control	0.95	0.98	0.95	0.95	0.95	0.95	0.95	0.95	0.95
KP Wald F-stat						16.94	16.93	15.82	17.64

Notes: Robust standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Source: Village-level data on Coptic churches and monasteries in 1500 constructed from al-Maqrizi (2002).

Table D.6: **Local determinants of conversions to Islam in 641–1200: Alternative measure of r_{ji}**

Dependent variable: =1 if no Coptic church or monastery in village j circa 1200

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.08		0.08	0.08	0.07	0.13	0.12	0.12	0.11
	(0.03)**		(0.03)**	(0.03)**	(0.03)**	(0.06)**	(0.06)**	(0.06)**	(0.05)**
=1 if saint-martyr (r_{ji})		-0.50	-0.50	-0.50	-0.52		-0.50	-0.50	-0.52
		(0.09)***	(0.10)***	(0.10)***	(0.09)***		(0.10)***	(0.10)***	(0.09)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1782	1778	1778	1778	1748	1782	1778	1778	1748
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.01	0.03	0.04	0.04	0.05	0.01	0.04	0.04	0.05
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat						17.23	17.43	16.47	16.65

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.08		0.08	0.08	0.07	0.13	0.12	0.12	0.08
	(0.04)**		(0.03)**	(0.03)***	(0.03)*	(0.07)*	(0.06)**	(0.06)**	(0.05)*
=1 if saint-martyr (r_{ji})		-0.50	-0.51	-0.50	-0.51		-0.51	-0.50	-0.51
		(0.09)***	(0.09)***	(0.09)***	(0.09)***		(0.09)***	(0.09)***	(0.09)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1782	1778	1778	1778	1748	1782	1778	1778	1748
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.01	0.03	0.04	0.04	0.05	0.01	0.04	0.04	0.05
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat						16.94	17.21	15.58	19.38

Notes: Standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.7: **Local determinants of conversions to Islam in 641–1200: Interaction of c_i and r_{ji}**

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
=1 if Arab settlement (c_i)	0.08	0.08	0.08	0.13	0.13	0.12
	(0.03)**	(0.03)**	(0.03)**	(0.06)**	(0.06)**	(0.05)**
=1 if HF visit (r_{ji})	-0.62	-0.63	-0.63	-0.44	-0.46	-0.50
	(0.17)***	(0.18)***	(0.18)***	(0.28)	(0.29)	(0.28)*
$c_i \times r_{ji}$	0.05	0.05	0.00	-0.19	-0.17	-0.17
	(0.20)	(0.20)	(0.21)	(0.35)	(0.36)	(0.35)
Byzantine controls?	No	Yes	Yes	No	Yes	Yes
Geographic controls?	No	No	Yes	No	No	Yes
Obs (villages)	1782	1782	1751	1782	1782	1751
Clusters (*kuras*)	42	42	42	42	42	42
R^2	0.04	0.04	0.05	0.04	0.04	0.05
Mean dep. var. in control	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat				.	20.38	20.12

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
=1 if Arab settlement (c_i)	0.08	0.08	0.07	0.13	0.13	0.09
	(0.04)**	(0.03)**	(0.04)*	(0.07)*	(0.06)**	(0.05)*
=1 if HF visit (r_{ji})	-0.62	-0.63	-0.64	-0.46	-0.48	-0.50
	(0.18)***	(0.18)***	(0.18)***	(0.28)	(0.28)*	(0.28)*
$c_i \times r_{ji}$	0.05	0.06	0.03	-0.17	-0.15	-0.17
	(0.20)	(0.20)	(0.21)	(0.36)	(0.36)	(0.35)
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	Yes	Yes	No	Yes	Yes
Geographic controls?	No	No	Yes	No	No	Yes
Obs (villages)	1782	1782	1751	1782	1782	1751
Clusters (*kuras*)	42	42	42	42	42	42
R^2	0.04	0.04	0.06	0.04	0.04	0.06
Mean dep. var. in control	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat				18.25	12.49	10.99

Notes: Standard errors clustered at the *kura* level are in parentheses. In columns (4)–(6), the excluded instruments in the first-stage regressions for Arab settlement (c_i) and its interaction with the HF visit status ($c_i \times r_{ji}$) are: (1) *kura*'s distance to *Arish* ($DistancetoArish_i$) (2) =1 if *kura* borders desert ($BorderDesert_i$), (3) $DistancetoArish_i \times BorderDesert_i$, (4) $r_{ji} \times DistancetoArish_i$, (5) $r_{ji} \times BorderDesert_i$, (6) $r_{ji} \times DistancetoArish_i \times BorderDesert_i$. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura* i circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura* i circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.8: **Local determinants of conversions to Islam in 641–1200: Control for pre-641 land inequality**

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.11		0.10	0.11	0.10	0.13	0.08	0.10	0.07
	(0.03)***		(0.03)***	(0.04)***	(0.07)	(0.09)	(0.10)	(0.06)*	(0.19)
=1 if HF visit (r_{ji})		-0.57	-0.56	-0.58	-0.64		-0.57	-0.58	-0.64
		(0.10)***	(0.09)***	(0.09)***	(0.11)***		(0.09)***	(0.09)***	(0.11)***
=1 if *autopract* c. 600	-0.06	-0.02	-0.04	-0.03	-0.01	-0.06	-0.04	-0.03	-0.01
	(0.04)	(0.06)	(0.04)	(0.04)	(0.07)	(0.04)	(0.04)	(0.04)	(0.07)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	589	589	589	589	575	589	589	589	575
Clusters (*kuras*)	21	21	21	21	21	21	21	21	21
R^2	0.01	0.06	0.07	0.07	0.09	0.01	0.07	0.07	0.09
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat						10.24	10.61	25.51	6.90

Notes: Standard errors clustered at the *kura* level are in parentheses. In columns (6)–(9), the excluded instrument in the first-stage regression for Arab settlement (c_i) is *kura*'s distance to *Arish* ($DistancetoArish_i$). Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.9: **Local determinants of conversions to Islam in 641–1200: Spatial-autoregressive model with spatial-autoregressive errors**

Dependent variable: =1 if no Coptic church or monastery in village j circa 1200

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if no church in 1200									
=1 if Arab settlement (c_i)	0.08		0.08	0.08	0.07	0.13	0.13	0.13	0.11
	(0.02)***		(0.02)***	(0.02)***	(0.02)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***
=1 if HF visit (r_{ji})		-0.63	-0.62	-0.62	-0.63		-0.62	-0.62	-0.63
		(0.08)***	(0.08)***	(0.08)***	(0.08)***		(0.08)***	(0.08)***	(0.08)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1730	1730	1730	1730	1730	1730	1730	1730	1730
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if no church in 1200									
=1 if Arab settlement (c_i)	0.08		0.08	0.08	0.06	0.15	0.14	0.15	0.10
	(0.02)***		(0.02)***	(0.02)***	(0.02)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***
=1 if HF visit (r_{ji})		-0.61	-0.61	-0.61	-0.62		-0.60	-0.60	-0.62
		(0.08)***	(0.08)***	(0.08)***	(0.08)***		(0.08)***	(0.08)***	(0.08)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1730	1730	1730	1730	1730	1730	1730	1730	1730
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78

Notes: Columns (1)-(5) report the results of estimating a spatial autoregressive model with spatial standard errors with inverse distance weighting matrix estimated using generalized spatial two-stage least squares (GS2SLS) (STATA command spreg). Columns (6)-(9) report the results of estimating a spatial-autoregressive model with spatial-autoregressive errors and additional endogenous variables (STATA command spivreg). Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A constant is included in all regressions.
Source: See the Supplemental Appendix.

Table D.10: **Local determinants of the total tax transfer: Measuring total tax transfer in 1477**

Dependent variable: Tax transfer (*'ibra*) in army dinars per unit of land in 1477

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.05		-0.05	-0.20	-0.14	-0.31	-0.31	-0.41	-0.20
	(0.26)		(0.26)	(0.24)	(0.19)	(0.31)	(0.31)	(0.30)	(0.30)
=1 if HF visit (r_{ji})		0.47	0.47	0.41	0.38		0.48	0.41	0.38
		(0.44)	(0.44)	(0.48)	(0.59)		(0.45)	(0.48)	(0.58)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1511	1511	1511	1511	1485	1511	1511	1511	1485
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.00	0.00	0.00	0.00	0.04	-0.00	-0.00	0.00	0.04
Mean dep. var. in control	2.89	2.84	2.89	2.89	2.89	2.89	2.89	2.89	2.89
KP Wald F-stat						16.32	16.42	16.17	14.65

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.38		-0.38	-0.43	-0.39	-0.59	-0.59	-0.71	-0.65
	(0.23)		(0.23)	(0.22)*	(0.16)**	(0.30)*	(0.30)*	(0.29)**	(0.26)**
=1 if HF visit (r_{ji})		0.62	0.62	0.64	0.68		0.62	0.64	0.69
		(0.43)	(0.45)	(0.45)	(0.53)		(0.45)	(0.45)	(0.53)
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1511	1511	1511	1511	1485	1511	1511	1511	1485
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.03	0.03	0.03	0.03	0.07	0.03	0.03	0.03	0.07
Mean dep. var. in control	2.89	2.84	2.89	2.89	2.89	2.89	2.89	2.89	2.89
KP Wald F-stat						17.32	17.30	14.40	19.94

Notes: Standard errors clustered at the *kura* level are in parentheses. In columns (4)–(6), the excluded instruments in the first-stage regressions for Arab settlement (c_i) and its interaction with the HF visit status ($c_i \times r_{ji}$) are: (1) *kura*'s distance to *Arish* ($DistancetoArish_i$) (2) =1 if *kura* borders desert ($BorderDesert_i$), (3) $DistancetoArish_i \times BorderDesert_i$, (4) $r_{ji} \times DistancetoArish_i$, (5) $r_{ji} \times BorderDesert_i$, (6) $r_{ji} \times DistancetoArish_i \times BorderDesert_i$. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is the population in 1897 ÷ land area in 1315. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.

Sources: See the Supplemental Appendix.

Table D.11: **Local determinants of the total tax transfer in 1375: Alternative measure of r_{ji}**

Dependent variable: Tax transfer (*'ibra*) in army dinars per unit of land in 1375

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.13		-0.12	-0.29	-0.23	-0.46	-0.44	-0.56	-0.36
	(0.30)		(0.29)	(0.27)	(0.21)	(0.35)	(0.34)	(0.33)*	(0.30)
=1 if saint-martyr (r_{ji})		0.39	0.38	0.31	0.45		0.36	0.29	0.43
		(0.69)	(0.69)	(0.68)	(0.67)		(0.69)	(0.68)	(0.65)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1511	1507	1507	1507	1482	1511	1507	1507	1482
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.00	0.00	0.00	0.01	0.04	-0.00	-0.00	0.00	0.04
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						16.32	16.51	16.18	14.67

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.51		-0.50	-0.55	-0.45	-0.73	-0.71	-0.83	-0.69
	(0.26)*		(0.26)*	(0.24)**	(0.18)**	(0.34)**	(0.33)**	(0.29)***	(0.28)**
=1 if saint-martyr (r_{ji})		0.62	0.64	0.60	0.54		0.65	0.60	0.54
		(0.64)	(0.62)	(0.65)	(0.65)		(0.61)	(0.63)	(0.64)
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1511	1507	1507	1507	1482	1511	1507	1507	1482
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.03	0.03	0.03	0.03	0.06	0.03	0.03	0.03	0.06
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						17.32	17.56	14.48	20.07

Notes: Tax transfer (*'ibra*) is in army dinars (≈13.3/20 dinars) per *feddan* (= 6,368 square meters) of land. Standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is the population in 1897 ÷ land area in 1315. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.12: **Local determinants of the total tax transfer in 1375: Province fixed effects**

Dependent variable: Tax transfer (*‘ibra*) in army dinars per unit of land in 1375

(a) One-way clustering by *kura*

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.52		-0.53	-0.57	-0.46	-0.39	-0.39	-0.45	-0.33
	(0.17)***		(0.17)***	(0.17)***	(0.10)***	(0.28)	(0.28)	(0.26)*	(0.17)*
=1 if HF visit (r_{ji})		0.86	0.91	0.92	0.84		0.90	0.91	0.83
		(0.44)*	(0.44)**	(0.44)**	(0.52)		(0.44)**	(0.44)**	(0.51)
Province FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1492	1492	1492	1492	1467	1492	1492	1492	1467
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.07	0.07	0.07	0.07	0.10	0.07	0.07	0.07	0.10
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						11.44	11.51	11.54	12.09

(b) Two-way clustering by both *kura* and province

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.52		-0.53	-0.57	-0.46	-0.39	-0.39	-0.45	-0.33
	(0.26)*		(0.26)*	(0.25)**	(0.10)***	(0.25)	(0.25)	(0.28)	(0.12)***
=1 if HF visit (r_{ji})		0.86	0.91	0.92	0.84		0.90	0.91	0.83
		(0.38)**	(0.38)**	(0.42)**	(0.45)*		(0.36)**	(0.37)**	(0.38)**
Province FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1492	1492	1492	1492	1467	1492	1492	1492	1467
Cluster 1 (*kuras*)	40	40	40	40	40	40	40	40	40
Cluster 2 (provinces)	19	19	19	19	19	19	19	19	19
R^2	0.07	0.07	0.07	0.07	0.10	0.07	0.07	0.07	0.10
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						49.13	49.94	54.77	58.59

Notes: Tax transfer (*‘ibra*) is in army dinars (≈13.3/20 dinars) per *feddan* (= 6,368 square meters). Standard errors are clustered at the *kura* level in panel (a) and at both the *kura* and province level (STATA commands reghdfe and ivreg2) in panel (b). Provinces are defined according to the administrative division in the 1315/1375 cadastre. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is (8) the population in 1897 ÷ land area in 1315. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.13: **Local determinants of the total tax transfer in 1375: Interaction of c_i and r_{ji}**

Dependent variable: Tax transfer (*‘ibra*) in army dinars per unit of land in 1375

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
=1 if Arab settlement (c_i)	-0.13	-0.30	-0.27	-0.45	-0.56	-0.45
	(0.30)	(0.28)	(0.21)	(0.35)	(0.34)*	(0.30)
=1 if HF visit (r_{ji})	0.90	0.78	0.66	0.88	0.61	0.81
	(1.06)	(1.15)	(1.18)	(1.27)	(1.37)	(1.37)
$c_i \times r_{ji}$	0.08	0.10	0.18	0.11	0.33	-0.01
	(1.15)	(1.21)	(1.25)	(1.48)	(1.58)	(1.49)
Byzantine controls?	No	Yes	Yes	No	Yes	Yes
Geographic controls?	No	No	Yes	No	No	Yes
Obs (villages)	1511	1511	1486	1511	1511	1486
Clusters (*kuras*)	40	40	40	40	40	40
R^2	0.00	0.01	0.02	-0.00	0.01	0.02
Mean dep. var. in control	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat				.	14.13	10.32

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
=1 if Arab settlement (c_i)	-0.51	-0.56	-0.50	-0.72	-0.85	-0.81
	(0.27)*	(0.25)**	(0.20)**	(0.34)**	(0.30)***	(0.33)**
=1 if HF visit (r_{ji})	0.67	0.77	0.90	0.47	0.53	1.04
	(1.10)	(1.10)	(1.08)	(1.38)	(1.38)	(1.30)
$c_i \times r_{ji}$	0.38	0.28	0.19	0.65	0.61	0.01
	(1.18)	(1.16)	(1.19)	(1.60)	(1.60)	(1.42)
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	Yes	Yes	No	Yes	Yes
Geographic controls?	No	No	Yes	No	No	Yes
Obs (villages)	1511	1511	1486	1511	1511	1486
Clusters (*kuras*)	40	40	40	40	40	40
R^2	0.03	0.03	0.04	0.03	0.03	0.04
Mean dep. var. in control	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat				7.18	11.01	10.11

Notes: Standard errors clustered at the *kura* level are in parentheses. In columns (4)–(6), the excluded instruments in the first-stage regressions for Arab settlement (c_i) and its interaction with the HF visit status ($c_i \times r_{ji}$) are: (1) *kura*'s distance to *Arish* ($DistancetoArish_i$) (2) =1 if *kura* borders desert ($BorderDesert_i$), (3) $DistancetoArish_i \times BorderDesert_i$, (4) $r_{ji} \times DistancetoArish_i$, (5) $r_{ji} \times BorderDesert_i$, (6) $r_{ji} \times DistancetoArish_i \times BorderDesert_i$. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is the population in 1897 ÷ land area in 1315. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A constant is included in all regressions.

Sources: See the Supplemental Appendix.

Table D.14: **Local determinants of the total tax transfer in 1375: Control for Mamluk LAs**

Dependent variable: Tax transfer (*'ibra*) in army dinars per unit of land in 1375

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.18		-0.18	-0.30	-0.26	-0.52	-0.51	-0.54	-0.38
	(0.29)		(0.29)	(0.27)	(0.21)	(0.31)*	(0.31)	(0.32)*	(0.30)
=1 if HF visit (r_{ji})		0.77	0.77	0.71	0.56		0.77	0.70	0.55
		(0.38)**	(0.39)*	(0.41)*	(0.48)		(0.40)*	(0.41)*	(0.48)
=1 if LA in 1375 Mamluk	0.64	0.62	0.63	0.55	0.72	0.65	0.64	0.55	0.73
	(0.17)***	(0.17)***	(0.18)***	(0.20)***	(0.20)***	(0.17)***	(0.17)***	(0.19)***	(0.19)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1485	1485	1485	1485	1460	1485	1485	1485	1460
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.01	0.01	0.01	0.01	0.06	0.01	0.01	0.01	0.06
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						16.59	16.70	16.72	14.80

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.55		-0.55	-0.59	-0.45	-0.80	-0.80	-0.91	-0.68
	(0.25)**		(0.25)**	(0.24)**	(0.18)**	(0.31)**	(0.31)**	(0.29)***	(0.27)**
=1 if HF visit (r_{ji})		0.85	0.84	0.85	0.82		0.83	0.85	0.82
		(0.39)**	(0.41)**	(0.41)**	(0.45)*		(0.41)**	(0.42)**	(0.45)*
=1 if LA in 1375 Mamluk	0.60	0.57	0.59	0.58	0.70	0.60	0.60	0.59	0.70
	(0.20)***	(0.20)***	(0.20)***	(0.20)***	(0.18)***	(0.19)***	(0.20)***	(0.19)***	(0.18)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1485	1485	1485	1485	1460	1485	1485	1485	1460
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.04	0.03	0.04	0.04	0.08	0.04	0.04	0.04	0.08
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						17.47	17.46	14.42	20.06

Notes: Standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is the population in 1897 ÷ land area in 1315. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions. Sources: See the Supplemental Appendix.

Table D.15: **Local determinants of the total tax transfer in 1375: Spatial-autoregressive model with spatial-autoregressive errors**

Dependent variable: Tax transfer (*‘ibra*) in army dinars per unit of land in 1375

(a) No region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
Tax transfer in 1375									
=1 if Arab settlement (c_i)	-0.14		-0.12	-0.23	-0.05	0.00	0.01	-0.31	0.05
	(0.18)		(0.18)	(0.19)	(0.20)	(0.24)	(0.24)	(0.26)	(0.25)
=1 if HF visit (r_{ji})		1.00	0.99	0.88	0.64		1.01	0.88	0.65
		(0.71)	(0.72)	(0.71)	(0.70)		(0.71)	(0.71)	(0.70)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1456	1456	1456	1456	1456	1456	1456	1456	1456
Mean dep. var. in control	3.43	3.30	3.43	3.43	3.43	3.43	3.43	3.43	3.43

(b) Region fixed effects

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
Tax transfer in 1375									
=1 if Arab settlement (c_i)	-0.29		-0.29	-0.29	-0.20	-0.44	-0.48	-0.50	-0.14
	(0.19)		(0.19)	(0.21)	(0.21)	(0.26)*	(0.26)*	(0.28)*	(0.26)
=1 if HF visit (r_{ji})		0.90	0.89	0.89	0.86		0.89	0.90	0.86
		(0.70)	(0.70)	(0.70)	(0.70)		(0.70)	(0.71)	(0.70)
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1456	1456	1456	1456	1456	1456	1456	1456	1456
Mean dep. var. in control	3.43	3.30	3.43	3.43	3.43	3.43	3.43	3.43	3.43

Notes: Tax transfer (*‘ibra*) is in army dinars (≈13.3/20 dinars) per *feddan* (= 6,368 square meters) of land. Columns (1)-(5) report the results of estimating a spatial autoregressive model with spatial standard errors with inverse distance weighting matrix estimated using generalized spatial two-stage least squares (GS2SLS) (STATA command spreg). Columns (6)-(9) report the results of estimating a spatial-autoregressive model with spatial-autoregressive errors and additional endogenous variables (STATA command spivreg). Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is the population in 1897 ÷ land area in 1315. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A constant is included in all regressions.

Sources: See the Supplemental Appendix.

Table D.16: **Time-series determinants of poll tax hikes ($\Delta\tau_t$) and conversion waves (ΔF_t) in 641–1170: Governor-level regressions**

(a) Dependent variable =1 if a poll tax hike mentioned during the reign of governor t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if governor hostile ($\hat{c}_t$)	0.54	0.55	0.54				0.53	0.53	0.50
	(0.11)***	(0.14)***	(0.13)***				(0.15)***	(0.14)***	(0.13)***
No. previous hostile gov. (n^c_{t-1})				-0.02	-0.04	-0.07	-0.00	-0.02	-0.05
				(0.01)**	(0.02)**	(0.04)	(0.01)	(0.02)	(0.04)
Governor's start year		0.04	0.08		0.37	0.86		0.38	0.78
		(0.21)	(0.30)		(0.34)	(0.87)		(0.35)	(0.76)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (governors)	121	121	121	122	122	122	121	121	121
Years	526	526	526	530	530	530	526	526	526
R^2	0.29	0.29	0.30	0.07	0.08	0.13	0.29	0.30	0.33
p-value (Breusch–Godfrey test)	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Mean dep. var.	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13

(b) Dependent variable =1 if a conversion wave mentioned during the reign of governor t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if governor hostile ($\hat{c}_t$)	0.47	0.48	0.45				0.47	0.47	0.46
	(0.16)***	(0.19)**	(0.19)**				(0.19)**	(0.19)**	(0.19)**
No. previous hostile gov. (n^c_{t-1})				-0.02	-0.02	-0.02	0.00	-0.00	0.01
				(0.01)***	(0.01)	(0.03)	(0.01)	(0.01)	(0.02)
Governor's start year		0.03	-0.03		0.07	-0.04		0.05	-0.12
		(0.11)	(0.17)		(0.21)	(0.57)		(0.20)	(0.44)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (governors)	121	121	121	122	122	122	121	121	121
Years	526	526	526	530	530	530	526	526	526
R^2	0.33	0.33	0.38	0.06	0.06	0.15	0.33	0.33	0.38
p-value (Breusch–Godfrey test)	0.00	0.00	0.00	0.03	0.03	0.02	0.00	0.00	0.00
Mean dep. var.	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07

Notes: There are 122 governors. Governors and Caliphs are identical in 868–905 and 935–1170, when Egypt was independent. We omit governor 1, because for every governor $t \geq 2$, $\hat{c}_t = 1$ is interpreted as equal to the maximum $\hat{c}_t$ of previous governors ($\hat{c}_1 = 1$). Newey-West standard errors, assuming that the error structure is both heteroskedastic and autocorrelated up to 15 lags, are in parentheses. Controls are (1) =1 if at least one foreign attack occurred under governor t, (2) =1 if at least one adverse Nile shock occurred under governor t. Governor's start year is normalized $\in [0,1]$ by subtracting 641 and dividing it by the maximum. Regressions are weighted by the length of governor's tenure. H_0 for the Breusch-Godfrey test is that there is no serial correlation up to 15 lags. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.17: **Time-series determinants of poll tax hikes ($\Delta\tau_t$) and conversion waves (ΔF_t) in 641–1170: Caliph-level regressions controlling for the number of previous poll tax hikes**

(a) Dependent variable =1 if a poll tax hike mentioned during the reign of Caliph t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if Caliph pious ($\hat{c}_t$)	0.25	0.26	0.29				0.21	0.16	0.17
	(0.12)**	(0.10)***	(0.14)**				(0.10)**	(0.12)	(0.14)
No. previous poll tax hikes				-0.04	-0.09	-0.12	-0.04	-0.07	-0.09
				(0.01)***	(0.03)***	(0.06)**	(0.02)**	(0.04)*	(0.07)
Caliph's start year		-0.46	-0.40		0.76	1.26		0.50	0.96
		(0.18)**	(0.30)		(0.39)*	(0.91)		(0.45)	(0.96)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (Caliphs)	64	64	64	65	65	65	64	64	64
Years	526	526	526	530	530	530	526	526	526
R^2	0.07	0.16	0.18	0.17	0.22	0.26	0.21	0.23	0.27
p-value (Breusch–Godfrey test)	0.04	0.05	0.02	0.04	0.02	0.01	0.04	0.02	0.01
Mean dep. var.	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

(b) Dependent variable =1 if a conversion wave mentioned during the reign of Caliph t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if Caliph pious ($\hat{c}_t$)	0.23	0.25	0.26				0.19	0.26	0.31
	(0.15)	(0.13)*	(0.14)*				(0.13)	(0.19)	(0.20)
No. previous poll tax hikes				-0.04	-0.02	-0.01	-0.04	0.01	0.04
				(0.01)***	(0.03)	(0.06)	(0.01)***	(0.05)	(0.07)
Caliph's start year		-0.65	-0.68		-0.36	-0.59		-0.81	-1.21
		(0.17)***	(0.20)***		(0.50)	(0.94)		(0.72)	(1.07)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (Caliphs)	64	64	64	65	65	65	64	64	64
Years	526	526	526	530	530	530	526	526	526
R^2	0.07	0.25	0.29	0.18	0.19	0.21	0.21	0.26	0.30
p-value (Breusch–Godfrey test)	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Mean dep. var.	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18

Notes: There are 65 Caliphs. We omit Caliph 1, because for every Caliph $t \geq 2$, $\hat{c}_t = 1$ is interpreted as equal to the maximum $\hat{c}_t$ of previous Caliphs ($\hat{c}_1 = 1$). Newey-West standard errors, assuming that the error structure is both heteroskedastic and autocorrelated up to 11 lags, are in parentheses. Controls are (1) =1 if at least one foreign attack occurred under Caliph t, (2) =1 if at least one adverse Nile shock occurred under Caliph t. Caliph's start year is normalized $\in [0,1]$ by subtracting 641 and dividing it by the maximum. Regressions are weighted by the length of Caliph's tenure. H_0 for the Breusch-Godfrey test is that there is no serial correlation up to 11 lags. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

Table D.18: **Time-series determinants of poll tax hikes ($\Delta\tau_t$) and conversion waves (ΔF_t) in 641–1170: Interaction of c_t and n^c_{t-1}**

	=1 if poll tax hike			=1 if conversion wave		
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if Caliph pious ($\hat{c}_t$)	0.15	0.20	-0.19	0.32	0.64	0.29
	(0.22)	(0.27)	(0.44)	(0.33)	(0.41)	(0.67)
No. previous pious Caliphs (n^c_{t-1})	-0.02	-0.01	-0.03	-0.01	0.05	0.03
	(0.01)***	(0.03)	(0.03)	(0.01)	(0.03)*	(0.04)
$\hat{c}_t \times n^c_{t-1}$	0.01	0.00	0.03	-0.00	-0.02	-0.00
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)
Caliph's start year		-0.27	-0.11		-1.47	-1.36
		(0.60)	(0.66)		(0.48)***	(0.57)**
Controls?	No	No	Yes	No	No	Yes
Obs (Caliphs)	64	64	64	64	64	64
Years	526	526	526	526	526	526
R^2	0.16	0.16	0.20	0.19	0.29	0.31
p-value (Breusch–Godfrey test)	0.06	0.04	0.01	0.01	0.01	0.00
Mean dep. var.	0.25	0.25	0.25	0.18	0.18	0.18

Notes: There are 65 Caliphs. We omit Caliph 1, because for every Caliph $t \geq 2$, $\hat{c}_t = 1$ is interpreted as equal to the maximum $\hat{c}_t$ of previous Caliphs ($\hat{c}_1 = 1$). Newey-West standard errors, assuming that the error structure is both heteroskedastic and autocorrelated up to 11 lags, are in parentheses. Controls are (1) =1 if at least one foreign attack occurred under Caliph t, (2) =1 if at least one adverse Nile shock occurred under Caliph t. Caliph's start year is normalized $\in [0,1]$ by subtracting 641 and dividing it by the maximum. Regressions are weighted by the length of Caliph's tenure. H_0 for the Breusch-Godfrey test is that there is no serial correlation up to 11 lags. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See the Supplemental Appendix.

References

al-Maqrizi, T. (2002). *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.).

Alesina, A., Baqir, R., and Easterly, W. (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics*, 114:1243–1284.

Besley, T., Jensen, A., and Persson, T. (2017). Norms, Enforcement, and Tax Evasion. mimeo.

Fargues, P. (2001). Demographic Islamization: Non-Muslims in Muslim Countries. *SAIS Review*, 21(2):103–116.

Galor, O. and Weil, D. (2000). Population, Technology and Growth: From the Malthusian Regime to the Demographic Transition. *American Economic Review*, 90:806–828.

Glaeser, E. and Shleifer, A. (2005). The Curley Effect. *Journal of Law, Economics, and Organization*, 21:1–19.

Hansmann, H. (1996). *The Ownership of Enterprise*. Belknap Press, Cambridge, MA.

Jia, R. and Persson, T. (2017). Ethnicity in Children and Mixed Marriages: Theory and Evidence from China. mimeo.

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

Russell, J. C. (1958). *Late Ancient and Medieval Population*. The American Philosophical Society, Philadelphia.

Russell, J. C. (1966). The Population of Medieval Egypt. *Journal of the American Research Center in Egypt*, 5:69–82.

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. (2018). On the Road to Heaven: Taxation, Conversions, and the Coptic-Muslim Socioeconomic Gap in Medieval Egypt. *Journal of Economic History*, 78(2):394–434.

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