Econometrica Supplementary Material

SUPPLEMENT TO "TAXING IDENTITY: THEORY AND EVIDENCE FROM EARLY ISLAM"

(*Econometrica*, Vol. 89, No. 4, July 2021, 1881–1919)

MOHAMED SALEH
Toulouse School of Economics, University of Toulouse Capitole

JEAN TIROLE
Toulouse School of Economics, University of Toulouse Capitole

APPENDIX A: THEORY

A.1. *Proof of Proposition 1*

LET US FIRST IMAGINE THAT THE LA'S revenue-collection constraint is not binding. The monotone-hazard-rate condition implies that the LA's optimal discriminatory tax in district i, τ_i^a, is a weakly increasing function of r_i and c_i, with pass-through rates between 0 and 1.

Next we identify the discriminatory tax rates that are implementable by the CA through a transfer requirement. For $c_i < 0$, these are exactly described by the interval $[\tau_i^a(c_i), \tau_i^m]$: For any $\tau_i < \tau_i^a(c_i)$, the LA would raise the discriminatory tax to $\tau_i^a(c_i)$, keep the extra revenue for itself, raise more revenue, and reach its optimal discriminatory taxation. For any $\tau_i > \tau_i^m$, the LA would lower the discriminatory tax to τ_i^m or below. Symmetrically, the implementable discriminatory tax rates for $c_i > 0$ are exactly described by the interval $[\tau_i^m, \tau_i^a(c_i)]$.

The upper bound on T_i for $c_i < 0$ is therefore equal to $\lambda_i + R_i^m$. Furthermore, setting $T_i = \lambda_i + R_i^m$ forces the LA to set discriminatory tax τ_i^m, which is as close to τ_i^c as the CA can get. Strict quasi-concavity of the latter's objective function then implies that this transfer requirement is optimal. For $c_i > 0$, the analysis is similar. For $c_i \geq c$, the CA can get its first best by setting $T_i = \lambda_i + R_i(\tau_i^c)$. The LA is then forced to moderate its discriminatory taxation so as to be able to raise enough revenue. Finally, for $c_i \in (0, c)$, the closest implementable tax rate (which is therefore optimal from strict quasi-concavity) is $\tau_i^a(c_i)$; the requested transfer is then $T_i = \lambda_i + R_i(\tau_i^a(c_i))$.

A.2. *Dynamics of Conversion*

Imagine first a world in which *both* rulers and agents are myopic ($\beta = 0$). Consider the tax that yields the CA's unconstrained static optimum under the budget constraint:

$$\tau^*(c_t, B_t - \lambda_t) \equiv \arg_{\{\tau_t[1-F(\tau_t)]\geq B_t-\lambda_t\}}(\tau_t - c_t)\big[1 - F(\tau_t)\big].$$

The τ^* is increasing in c_t and decreasing in $(B_t - \lambda_t)$. Being myopic, agent θ converts whenever he has not yet converted and $\tau_t > \theta$. Ruler t chooses

$$\tau_t = \max\big\{\tau^*(c_t, B_t - \lambda_t), \theta^*_{t-1}\big\}. \tag{A.1}$$

Mohamed Saleh: mohamed.saleh@tse-fr.eu
Jean Tirole: jean.tirole@tse-fr.eu

 https://doi.org/10.3982/ECTA17265

To understand (A.1), suppose first that $\tau^*(c_t, B_t - \lambda_t) \geq \theta^*_{t-1}$ (as is the case, for instance, if there have been few or no conversions yet). By definition, $\tau^*(c_t, B_t - \lambda_t)$ yields the static optimum and cutoff $\theta^*_t = \tau^*(c_t, B_t - \lambda_t)$. Next, suppose that $\tau^*(c_t, B_t - \lambda_t) < \theta^*_{t-1}$. In the range $\tau_t \in [0, \theta^*_{t-1}]$, the demand for conversion is inelastic and so the objective function, $\lambda_t + [\tau_t - c_t][1 - F(\theta^*_{t-1})]$, is strictly increasing in τ_t. In either case, $\theta^*_t = \tau_t$. It turns out that these strategies are still optimal when the players value the future.

PROPOSITION 2—Dynamics of conversion: *For any $\beta \in [0, 1)$, there exists a Markov perfect equilibrium in which both the ruler and the agents behave as if they were myopic. The date-t tax and cutoff are $\tau_t = \theta^*_t = \max_{1\leq k\leq t} \tau^*(c_k, B_k - \lambda_k)$. This implies that the tax base shrinks and the discriminatory tax increases over time.*

(i) *If only c_t varies, then $\tau_t = \tau^*(\max_{1\leq k\leq t}\{c_k\}, B - \lambda)$.*
(ii) *If only $B_t - \lambda_t$ varies, then $\tau_t = \tau^*(c, \min_{1\leq k\leq t}\{B_k - \lambda_k\})$.*

The equilibrium can further be shown to be unique if the horizon is finite, and, under additional assumptions, under infinite horizon (the environment considered here) as well. The formal proof of Proposition 2 follows the lines in Tirole (2016).

It can further be checked that even if the CA does not set taxes itself, it can still, through a transfer demand T_t, induce aligned LAs to implement the policy described in Proposition 2.[77]

The apostasy assumption and its ratcheting corollary validate this "ever more religious tax base" argument, but also show that it is not a foregone conclusion. Indeed, the discriminatory tax and tax revenue are constant in a stationary economy for the identity-based model. They are also constant for a nonstationary economy under the extraction model: In the extraction model, the ruler maximizes $\lambda_t + R(\tau_t, \theta^*_t)$ subject to $R(\tau_t, \theta^*_t) \geq B_t - \lambda_t$, where $R(\tau_t) = \tau_t[1 - F(\theta^*_t)]$. From our assumption that the budgetary need can always be met, then $\tau_t = \tau^m$, the monopoly level that maximizes $\tau[1 - F(\tau)]$ for all t.

COROLLARY 2—Time-series comparison with the extraction model: *In the extraction model, the tax base and the discriminatory tax are constant over time.*

Time-Increasing Relevance of Extraction Model

Our claim that the extraction model gains in predictive power over time is based on Proposition 2, which states that $\theta^*_t = \max_{1\leq k\leq t} \tau^*(c_k, B_k - \lambda_k)$. Suppose that the joint distribution of the ruler's type c_t and of the net budgetary needs $B_t - \lambda_t$ is the same over time. This generates a distribution $H(\tau_t)$ on some interval $[\underline{\tau}, \overline{\tau}]$ for the date-t ruler's *desired* discriminatory tax $\tau^*(c_t, B_t - \lambda_t)$ (the actual one, as we showed, may be constrained by previous choices), and a cumulative distribution function $H^{t-1}(\max_{1\leq k\leq t-1} \tau^*(c_k, B_k - \lambda_k))$ for the highest discriminatory tax prior to date t. The expected number of conversions at date t is equal to $\int_{\underline{\tau}}^{\overline{\tau}} [\int_{\chi}^{\overline{\tau}} [F(\tau) - F(\chi)]\, dH(\tau)]\, dH^{t-1}(\chi)$, which after an integration by parts can be shown to be decreasing in t. Similarly, using the fact that the discriminatory tax is on the downward-sloping side of the Laffer curve, the expected reduction in discriminatory tax revenue from date $t-1$ to date t is $\int_{\underline{\tau}}^{\overline{\tau}} [\int_{\chi}^{\overline{\tau}} [R^c(\chi) - R^c(\tau)]\, dH(\tau)]\, dH^{t-1}(\chi)$ and is decreasing in t.

[77]Either $\theta^*_{t-1} > \tau^a(c_t)$ and then LA i's objective function is $\lambda_t + \tau_t[1 - F(\theta^*_{t-1})] - T_t$ for $\tau_t \leq \theta^*_{t-1}$ or the smaller $\lambda_t + \tau_t[1 - F(\tau_t)] - T_t$ for $\tau_t > \theta^*_{t-1}$. Strict quasi-concavity then implies that the LA's optimum is at $\tau_t = \theta^*_{t-1}$. For $\tau^a(c_t) \geq \theta^*_{t-1}$, the equilibrium policy can be decentralized by similarly setting a transfer demand $T_t = \max\{B_t, \lambda_t + \tau^a(c_t)[1 - F(\tau^a(c_t))]\}$.

Delegation in Time-Series Model

We assumed for expositional simplicity that LAs were congruent with the CA. However, even in the presence of (possibly district-specific) agency problems, the optimality of myopic behaviors and the ratchet property still hold. Consider, for example, the case in which, at date t, the CA has identity c_t and the LA has identity c_t^{LA} (again, this identity could be district specific, at the cost of heavier notation). The date-t CA must then account for the implementability constraint $\tau_t \in [\tau^a(c_t^{\text{LA}}), \tau^m]$ if $c_t^{\text{LA}} \leq 0$ and $\tau_t \in [\tau^m, \tau^a(c_t^{\text{LA}})]$ if $c_t^{\text{LA}} \geq 0$. Let (I) denote this implementability condition. Let $\tau^*(c_t, c_t^{\text{LA}}, B_t - \lambda_t) \equiv \arg\max_{\{(I), \tau_t[1-F(\tau_t)] \geq B_t - \lambda_t\}} (\tau_t - c_t)[1 - F(\tau_t)]$ denote the CA's desired static discriminatory tax, characterized in Proposition 1 and Figure 1. The equilibrium discriminatory tax in the time-series model is then $\tau_t = \max\{\tau^*(c_t, c_t^{\text{LA}}, B_t - \lambda_t), \theta^*_{t-1}\}$. This implies, for example, that even ignoring time-varying budget constraints, a strong-identity earlier ruler may not have had the opportunity to convert as many Copts as he desired because of the strong presence of Copts among the LAs, leaving scope for further conversions by subsequent rulers who were not necessarily more religious.

A.3. *External Threats*

In Appendices A.3 and A.4 (on external and internal threats, respectively), we make the following assumption.

ASSUMPTION 1: *In the rest of the section*, $c_t = c$, $\lambda_t = \lambda$, *and* $B_t = B$ *for all* t.

To capture external threats, we assume that there is probability $x_t \geq 0$ that the ruler is evicted at date t conditional on not having been evicted earlier. The sequence $\{x_t\}$ is (for simplicity) known and exogenous. As discussed in the text, the ruler cares about what happens when and when not in power; that is, he cares about his legacy. Except for the sequence $\{x_t\}_{t\geq 1}$, all parameters are invariant as stated in Assumption 1, and we suppose that the budget constraint is never binding (the analysis can be generalized if that is not the case).

PROPOSITION 3—Option value under external threats: . *Let* $\tau^c \equiv \arg\max_{\{\theta\}} (\theta - c)[1 - F(\theta)]$ *and* $K_t \equiv (1 + \frac{\beta}{1-\beta} x_{t+1})$. *In equilibrium,*[78] *the date-*t *discriminatory tax is* $\tau_t = K_t \tau^c$ *and the discriminatory tax revenue is* $R_t = K_t \tau^c[1 - F(\tau^c)]$. *In particular, if* x_t *is weakly decreasing* (*increasing*) *over time, so are* τ_t *and* R_t. *All conversions occur at date* 1.

COROLLARY 3—External threats: comparison with the extractive model: *The external-threats dynamics for the extractive model are identical with those of the identity model, except that the stable fraction of converts is* $F(\theta^m)$, *where* θ^m *solves* $\max\{\theta[1 - F(\theta)]\}$.

PROOF OF PROPOSITION 3: The agents' equilibrium strategy can be described by the cutoff rule at date t,

$$\theta_t^* = \max\left\{\theta_{t-1}^*, \frac{\tau_t}{K_t}\right\},$$

and the discriminatory tax obeys

$$\tau_t = K_t \max\{\theta^*, \theta_{t-1}^*\}.$$

[78] As in Proposition 2, equilibrium uniqueness requires further assumptions in the case of an infinite horizon.

To see that this is an equilibrium, note that the date-t cutoff, if interior ($\theta^*_t > \theta^*_{t-1}$) satisfies $(1+\frac{\beta}{1-\beta}x_{t+1})\theta^*_t = \tau_t$: Either the ruler is removed at date $(t+1)$ and then the cutoff type enjoys θ^*_t forever or the ruler remains in place and then type $\theta^*_t \leq \theta^*_{t+1}$ prefers (weakly or strongly) to pay the tax τ_{t+1}.

As for the ruler, note that the equilibrium behaviors deliver the upper bound on his intertemporal payoff that would correspond to the no-external-challenge environment ($x_t \equiv 0$ for all t):

$$W_t^{\max}(\theta^*_{t-1}) = \begin{cases} \dfrac{1}{1-\beta}\big[\lambda + (\theta^* - c)\big[1 - F(\theta^*)\big]\big] & \text{if } \theta^*_{t-1} \leq \theta^*, \\ \dfrac{1}{1-\beta}\big[\lambda + (\theta^*_{t-1} - c)\big[1 - F(\theta^*_{t-1})\big]\big] & \text{if } \theta^*_{t-1} > \theta^*. \end{cases}$$

To see this, assume that $\theta^*_{t-1} \leq \theta^*$, say (the proof is the same in the opposite case, due to strict quasi-concavity of the adjusted tax revenue). Let the ruler charge $K_t\theta^*$. Then

$$\begin{aligned} W_t = {} & \big[\lambda + (K_t\theta^* - c)\big[1 - F(\theta^*)\big]\big] \\ & + \beta x_{x+1}\left[-\frac{c\big[1 - F(\theta^*)\big]}{1-\beta}\right] + \beta(1 - x_{t+1})W_{t+1}, \end{aligned}$$

so W_{t+1} is equal to $\frac{1}{1-\beta}[\lambda + (\theta^* - c)[1 - F(\theta^*)]$. Then W_t takes this value as well. The upper bound on the ruler's continuation payoff can be reached though a stationary policy $\theta^*_{t+k} = \max\{\theta^*, \theta^*_{t-1}\}$. So no deviation for any history can yield more that the equilibrium strategy. *Q.E.D.*

A.4. *Internal Threats and Time-Decreasing Resistance*

To facilitate the understanding of endogenously evolving internal challenges, we first gain intuition about the threat of rebellion by analyzing the static case (Proposition 4) and then state our main proposition (Proposition 5).

Static analysis of the rebellion threat. Let us first consider the (noiseless version of the) static case. Assume that it takes $[1 - F(\hat{\theta})]$ rebels to topple the CA and the individual cost of doing so is ρ. In the following discussion, we will say that the threat of rebellion is low (resp. high) if $\hat{\theta}$ is low (high), that is, if the number of required rebels is high (low); we could alternatively index the threat of rebellion by (minus) the cost ρ of rebelling. To avoid unnecessary notation, assume $\hat{\theta} \geq 0$. The no-rebellion constraint for taxes $\{\hat{\lambda} \leq \lambda, \hat{\tau}\}$ is that the rebellion cost exceeds the marginal rebel's gain $G(\hat{\theta})$ from a successful rebellion:[79]

$$\rho \geq \hat{\lambda} + \min\{\hat{\tau}, \hat{\theta}\} \equiv G(\hat{\theta}).$$

ASSUMPTION 2—Relevant rebellion threat: *We have* $\lambda + \min\{\hat{\theta}, \tau^c\} > \rho$.

Recall that in the absence of rebellion threat, the CA's first best is $\hat{\lambda} = \lambda$ and $\hat{\tau} = \tau^c$. Were Assumption 2 violated, the threat of rebellion would be irrelevant and the first-best

[79]Allowing for negative values of $\hat{\theta}$, this condition would be $\rho \geq \lambda + \min\{\max\{\hat{\theta}, 0\}, \tau\}$.

level of taxes $\{\lambda, \tau^c\}$ would prevail. We look at the optimal pair $\{\hat{\lambda} \le \lambda, \hat{\tau}\}$ of taxes that the CA would like to implement. Let $\tilde{\tau} < \tau^c$ be uniquely defined as $\arg\max\{R^c(\hat{\tau}) - \hat{\tau}\}$ or $\tilde{\tau} + \frac{F(\tilde{\tau})}{f(\tilde{\tau})} = c$ (this is the optimal discriminatory tax when an increase in that tax must be offset one-for-one by a decrease in the uniform tax). The CA picks the discriminatory tax rate that maximizes $\hat{\lambda} + (\hat{\tau} - c)[1 - F(\hat{\tau})]$ subject to $\hat{\lambda} \le \lambda$ (feasibility), $\hat{\lambda} + \min\{\hat{\tau}, \hat{\theta}\} \le \rho$ (no-rebellion constraint), and $\hat{\tau} \in [\tau^m, \tau^c]$ (implementability). For the sake of simplicity, we do not put any lower bound at 0 for $\hat{\lambda}$ (uniform subsidies are feasible).

Finally, let $\theta^* \in [\tau^m, \tau^c]$ be defined by $\theta^* \equiv R^c(\tau^c) - R^c(\tau^*) + \tau^*$, where $\tau^* \equiv \max\{\tau^m, \tilde{\tau}, \rho - \lambda\} \in [\tau^m, \tau^c]$.

PROPOSITION 4—Capping the uniform tax to thwart rebellion: the static case: *Under Assumptions 1 and 2, we have the following cases*:

(i) *For a low threat of rebellion ($\hat{\theta} < \theta^*$), the marginal rebel is a convert; the optimal policy for the CA is to reduce the uniform tax to $\hat{\lambda} = \rho - \hat{\theta} < \lambda$ and to keep the discriminatory tax at $\hat{\tau} = \tau^c$.*

(ii) *For a high threat of rebellion ($\hat{\theta} > \theta^*$), the optimal policy for the CA is to reduce both the uniform tax from λ to $\rho - \tau^*$ and the discriminatory tax from τ^c to τ^*. The marginal rebel is a nonconvert.*

PROOF: Assume that $\lambda + \min\{\hat{\theta}, \tau^c\} > \rho$, so there is a real threat of rebellion. The CA's optimization program is

$$\max_{\{\hat{\tau}, \hat{\lambda}\}} \hat{\lambda} + (\hat{\tau} - c)\big[1 - F(\hat{\tau})\big]$$

subject to

$$\begin{cases} \hat{\lambda} + \min\{\hat{\theta}, \hat{\tau}\} \le \rho & \text{(no rebellion)}, \\ \hat{\lambda} \le \lambda & \text{(uniform tax cannot exceed its extractive level)}, \\ \hat{\tau} \in \big[\tau^m, \tau^c\big]. \end{cases}$$

Suppose that the CA chooses $\{\hat{\lambda}, \hat{\tau}\}$ such that $\hat{\lambda} + \hat{\tau} \le \rho$ (that is, $\hat{\tau} \le \hat{\theta}$ and so the marginal rebel is a nonconvert). Then the CA has welfare $\hat{\lambda} + (\hat{\tau} - c)[1 - F(\hat{\tau})] = \rho - \hat{\tau} + R^c(\hat{\tau})$, which is decreasing in $\hat{\tau}$ for $\hat{\tau} \ge \tilde{\tau}$, where

$$\tilde{\tau} + \frac{F(\tilde{\tau})}{f(\tilde{\tau})} = c.$$

Let us restrict the consideration set for the discriminatory tax. First, $\hat{\tau} < \tau^m$ is not implementable. Next, $\hat{\tau} < \tilde{\tau}$ is always weakly dominated. Consider a small change $\delta\hat{\tau} = +\epsilon$ and $\delta\hat{\lambda} = -\epsilon$; then the no-rebellion constraint, $\hat{\lambda} + \min\{\hat{\tau}, \hat{\theta}\} \le \rho$, remains satisfied and $\delta(\hat{\lambda} + R^c(\hat{\tau})) = \epsilon((R^c)' - 1) > 0$ for $\hat{\tau} < \tilde{\tau}$. Finally, $\tau < \rho - \lambda$ is not feasible unless $\hat{\theta}$ is a convert, that is, $\lambda + \hat{\theta} = \rho$ and $\hat{\theta} < \tau$.

Let $\tau^* \equiv \max\{\tau^m, \tilde{\tau}, \rho - \lambda\}$. Because we are interested only in the case of a rebellion threat ($\lambda + \tau^c > \rho$), $\tau^* < \tau^c$.

We distinguish three regions.

Region 1: $\hat{\theta} < \tau^m$. Then $\hat{\theta}$ is a convert, $\lambda + \hat{\theta} = \rho$, and $\hat{\tau} = \tau^c$. Welfare is

$$W^1 \equiv \rho - \hat{\theta} + R^c(\tau^c).$$

Region 2: $\hat{\theta} > \tau^c$. Type $\hat{\theta}$ is then necessarily a nonconvert and

$$W^2 \equiv \rho - \tau^* + R^c(\tau^*).$$

Region 3: $\tau^m \leq \hat{\theta} \leq \tau^c$. We have $\hat{\tau} > \hat{\theta}$ (the marginal rebel is a convert) and then at the optimum $\hat{\tau} = \tau^c$. Welfare is then $W^3 = W^1$. Furthermore, welfare W^1 can be obtained for any $\hat{\theta} \in [\tau^m, \tau^c]$.

Alternatively, $\hat{\tau} \leq \hat{\theta}$ (the marginal rebel is a nonconvert). Then $\hat{\tau} = \tau^*$, yielding welfare $W^3 = W^2$. But unlike for W^1, W^2 is not feasible for any $\hat{\theta} \in [\tau^m, \tau^c]$: It requires that $\tau^* \leq \hat{\theta}$.

Optimal welfare is, therefore, W^1 for $\hat{\theta} \in [\tau^m, \tau^*]$. On $[\tau^*, \tau^c]$, note that $dW^1/d\hat{\theta} = -1$ while $dW^2/d\hat{\theta} = 0$. Furthermore,

$$\begin{aligned} W^1(\tau^*) - W^2(\tau^*) &= R^c(\tau^c) - R^c(\tau^*) \\ &> 0 > W^1(\tau^c) - W^2(\tau^c) = -(\tau^c - \tau^*). \end{aligned}$$

Therefore, in this interval $W^3 = W^2$ if and only if $\hat{\theta} \geq \theta^*$, where

$$\theta^* \equiv R^c(\tau^c) - R^c(\tau^*) + \tau^*.$$

Putting all three regions together yields two alternatives:

(i) For $\hat{\theta} < \theta^*$, $W = W^1$, $\hat{\lambda} = \rho - \hat{\theta}$, and $\hat{\tau} = \tau^c$.

(ii) For $\hat{\theta} > \theta^*$, $W = W^2$, $\hat{\lambda} = \rho - \tau^*$, and $\hat{\tau} = \tau^*$. *Q.E.D.*

Dynamic analysis of the rebellion threat. Suppose next that $t = 1, 2, \dots, +\infty$, and that agents and the CA apply the same discount factor β to future utilities. The assumption that $T = +\infty$ is important here; with a finite horizon, the gain from a successful rebellion would decrease over time, generating an artificial increase over time in the relative cost of rebellion (expressed relative to future benefits). We assume that the cost of rebellion is $\rho/(1-\beta)$: while rebellion is a one-shot activity, we normalize its per-period cost to be ρ to facilitate the comparison with the static legitimacy model. The willingness to pay to keep one's identity is still θ per period. We focus on Markov perfect equilibria (MPE).

PROPOSITION 5—Far-sighted players and decreasing resistance: *Let* $\tau^* \equiv \max(\tau^m, \tilde{\tau}, \rho - \lambda)$ *and* $\theta^* \equiv R^c(\tau^c) - R^c(\tau^*) + \tau^* \in (\tau^*, \tau^c)$. *Under Assumptions* 1 *and* 2, *we have the following cases*:

(i) *If* $\hat{\theta} < \theta^*$, *the marginal rebel* $\hat{\theta}$ *converts at date* 1. *In the CA's optimal MPE, the CA back-loads the uniform tax, charging a low uniform tax at date* 1 *and raising the uniform tax to* $\min\{\lambda, \rho\}$ *once the threat of rebellion has subsided. The discriminatory tax is equal to* τ^c *in all periods.*

If $\hat{\theta} > \theta^*$ *and* $\rho - \tau^* \leq \lambda$, *the marginal rebel* $\hat{\theta}$ *never converts. The discriminatory tax and the uniform tax are equal to* $\tau_t = \tau^*$ *and* $\lambda_t = \rho - \tau^*$ *for all* t.

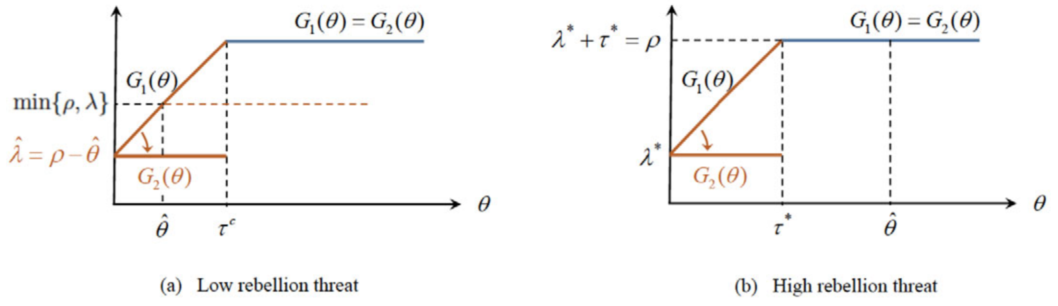

FIGURE 2.—Time-decreasing resistance. Note that $G_t(\theta)$ is type θ's gain from a successful rebellion at date t.

Despite the lack of commitment, the CA's per-period welfare is in both cases the same as in the static model, namely $\rho-\tau^*+R^c(\tau^*)$ *for* $\hat{\theta}\le\theta^*$ *and* $\rho-\hat{\theta}+R^c(\tau^c)$ *for* $\hat{\theta}\ge\theta^*$.

(ii) *The MPE maximizing the CA's payoff* (*characterized in part* (i)) *is, furthermore, coalition-proof à la Bernheim–Peleg–Whinston* (1987) *if* $\tau^c\ge\beta\frac{\hat{\theta}+\lambda-\rho}{1-\beta}$ *when* $\hat{\theta}<\theta^*$ *or if* $\hat{\theta}>\theta^*$.

Intuition. Assume in a first step that all parties are myopic ($\beta=0$); in particular, each generation cares about its own welfare, but apostasy creates a linkage between periods as conversions apply to future generations. A key insight is that when the marginal rebel is a convert, the marginal rebel's incentive to rebel decreases over time, as depicted in Figure 2(a) in the two-period case. Earlier converts' gain from a successful rebellion is limited to the uniform tax and no longer includes the preservation of their foregone identity. Thus, suppose that the threat of rebellion is not too high: $\widehat{\theta}<\theta^*$. A myopic CA then selects $\{\widehat{\lambda}=\rho-\widehat{\theta},\widehat{\tau}=\tau^c\}$ at date 1. All types $\theta\le\tau^c$ including the marginal rebel convert at date 1. Because the marginal rebel cares only about the uniform tax from date 2 on, the no-rebellion constraint at date 2 and at any subsequent date t yields

$$\lambda_t=\min\{\rho,\lambda\}$$

and the date $t\ge2$ welfare becomes $\min\{\rho,\lambda\}+R^c(\tau^c)$. This is also the maximal welfare that can be obtained in any given period: The uniform tax cannot exceed $\min\{\rho,\lambda\}$ without triggering a rebellion, and $R^c(\tau^c)$ is the maximum adjusted revenue from the discriminatory tax. All conversions occur at date 1, as the discriminatory tax is constant at τ^c from date 1 on. But the uniform tax increases from $\lambda_1=\rho-\widehat{\theta}<\min\{\rho,\lambda\}$ to $\lambda_2=\lambda_3=\cdots=\min\{\rho,\lambda\}$ once the threat of rebellion has decreased. In particular, it increases to equal the extractive tax if $\lambda\le\rho$.

By contrast, when the marginal rebel is a nonconvert in the static analysis ($\widehat{\theta}>\theta^*$; see Figure 2(b)), the threat of rebellion remains the same over time. The CA in each period must still satisfy $\widehat{\lambda_t}+\widehat{\tau_t}\le\rho$ for each t, and so $\widehat{\tau_t}=\tau^*$ and $\widehat{\lambda_t}=\rho-\tau^*$ for all $t\ge1$. The equilibrium is stationary and replicates the static analysis in each period.

When agents are far-sighted ($\beta>0$), one might guess that the agents' resistance in this case would no longer subside over time, as they internalize the fact that not rebelling will lead to an increase in future taxes. Interestingly, this is not the case. The reason has to do with the difference in objectives between marginal and inframarginal agents when the marginal rebel is a convert; the marginal rebel is then concerned solely with the discounted flow of uniform taxes; by contrast, agents who do not convert are affected by both the uniform and the discriminatory discounted taxes, as is the ruler. The ruler can soft-pedal uniform taxes and back-load their flow so as to dissuade the marginal convert from rebelling. Put differently, he can divide and conquer the agent community. Once the resistance of the converts has been reduced, the ruler can then increase the tax burden. The proof of Proposition 5 can be found in Appendix C.

Timing of the tax reform. We obtain two corollaries in the simple context of myopic agents and rulers. These corollaries also hold when $\beta > 0$. Consider Proposition 5: The uniform tax is initially low to avoid a rebellion, and so a tax reform is not necessary or at least yields low benefits. Once the threat of rebellion has decreased, though, the uniform tax is optimally raised, which may require a tax reform if the initial cap was low.

COROLLARY 4—Timing of tax reform: *Suppose that the uniform tax is initially capped by some level $\overline{\lambda}$ and that removing this cap, allowing any level of uniform tax up to the extraction level $\lambda > \overline{\lambda}$, generates some instantaneous cost $C > 0$ for the CA. Under Assumptions 1 and 2, we have the following cases*:

(i) *If the threat of rebellion is low ($\widehat{\theta} < \widetilde{\theta}$ for some $\widetilde{\theta}$)*[80] *and the cap on the uniform tax is binding ($\lambda < \rho - \widehat{\theta}$), then the tax reform occurs at date* 1 *if $\overline{\lambda} < \rho - \tau^c - C$ and at date* 2 *if $\overline{\lambda} \in (\rho - \tau^c - C, \min\{\rho, \lambda\} - C)$ (it never occurs if c if higher*).

(ii) *If the threat of rebellion is high ($\widehat{\theta} > \widetilde{\theta}$), then the tax reform, if it ever occurs, always occurs at date* 1.

Second, we have assumed for simplicity that the CA is well informed about the threat of rebellion. As a consequence, rebellions constrain the tax system but do not occur on the equilibrium path. With imperfect information about the threat of rebellion, rebellions in general will occasionally occur in equilibrium. When there is little uncertainty, rebellions will be rare. To obtain results about the composition of the rebel group after date 1 (at date 1 all start nonconverts, so only nonconverts can rebel), we consider the limit of distributions of the rebellion parameters ρ and $\widehat{\theta}$ converging to the certainty case.[81] The intuition behind the following proposition can be grasped from Figure 2(a) and (b). Suppose for instance that $\min\{\rho, \lambda\} = \rho$ and that the marginal rebel is a convert; a small overestimation of the cost of rebellion will lead converts with types in roughly $[\widehat{\theta}, \tau^c]$ and nonconverts with types $\theta \geq \tau^c$ to join the rebellion. Compare this with the case in which the marginal rebel is a nonconvert. Then a small overestimation of the level of ρ will lead (almost) only nonconverts to rebel.

COROLLARY 5—Composition of rebel group: *When the uncertainty about the cost of rebellion is small, at date* 1, *only nonconverts rebel when a rebellion occurs. Later on,* (i) *if the threat of rebellion is low ($\widehat{\theta} < \theta^*$), actual rebellions involve both converts and nonconverts*; (ii) *if the threat of rebellion is high ($\widehat{\theta} > \theta^*$), actual rebellions involve almost only nonconverts (the fraction of rebels who are converts tends to* 0 *as the uncertainty vanishes*).

[80] In general, $\widetilde{\theta}$ differs from θ^*, as the cap affects the welfare in the two regions.

[81] Two comments are in order here. First, we keep the analysis informal. The notion of vanishing uncertainty is the same as in Nash's celebrated noncooperative Nash demand game when the uncertainty about the size of the endowment vanishes. Second, the uncertainty could affect parameters other than ρ and $\widehat{\theta}$ without changing the analysis.

A.5. *Implication of Theory Under Binary Coding*

We here investigate the effect of the binary measurement of c_t on the probability of poll tax hikes and conversion waves under ruler t. We think of our binary measure, $\hat{c}_t$, as truncation at some level c^*: $\hat{c}_t = 1$ if $c_t \geq c^*$ and $\hat{c}_t = 0$ otherwise. Let $c^{t-1} \equiv \max_{1\leq k\leq t-1} c_k$ and let $\hat{c}^{t-1} \equiv \max_{1\leq k\leq t-1} \hat{c}_k$, the associated binary variable. Suppose that c_t is an independent random draw from a distribution $G(c_t)$. Let $n^c_{t-1} \equiv \sum_{1\leq k\leq t-1} \hat{c}_k$ denote the number of realizations $\hat{c}_k = 1$ up to $t-1$. We have for $n^c_{t-1} \geq 1$, $E[F_t - F_{t-1}] = \hat{c}_t \int_{c^*}^{+\infty}[\int_{c^{t-1}}^{+\infty}[F(\tau^a(c_t)) - F(\tau^a(c^{t-1}))]\,dG(c_t)]\frac{d}{dc^{t-1}}[(\frac{G(c^{t-1})-G(c^*)}{1-G(c^*)})^{n^c_{t-1}(c^{t-1})}]$. So in reduced form, $E[F_t - F_{t-1}] = \hat{c}_t W(n^c_{t-1})$, where W is a decreasing function converging to 0 as n^c_{t-1} goes to infinity. Similarly, the probability of a tax hike is $P_t = \hat{c}_t \int_{c^*}^{+\infty}[1 - G(c^{t-1})]\frac{d}{dc^{t-1}}[(\frac{G(c^{t-1})-G(c^*)}{1-G(c^*)})^{n^c_{t-1}(c^{t-1})}]$, and satisfies the same properties as $E(F_t - F_{t-1})$. To sum up, the probability of poll tax rises and conversion waves is increasing in $\hat{c}_t$ and is decreasing in n^c_{t-1}.

A.6. *Persecutions*

Agency Model of Persecutions

Consider a CA with identity c and a LA with identity c_i. Express the cost of persecution borne by a nonconvert in district i, $p_i \geq 0$, in terms of money, so that the agents' total cost of keeping their identity is $\tau_i + p_i$. Persecution does not bring any cash; it only serves to deter the agents from keeping their identity. The CA chooses the level of acceptable persecutions and the LA then collects taxes.[82]

COROLLARY 6—Agency and persecutions: *Consider an economy with parameter sequence $\{c_t, c_{it}, B_t, \lambda_t\}_{t\geq 1}$. Then the tax base shrinks and the discriminatory tax increases over time.*

(i) *Persecutions do not occur as long as the CA's identity is not much stronger than the LAs' identity: There exists a function c^* satisfying $c^*(\tilde{c}) > \tilde{c}$ for all $\tilde{c}$ such that there are no persecutions ($p_{it} = 0$) if and only if $c_t \leq c^*(c_{it})$.*

(ii) *The ruler is more likely to allow persecutions in districts with the weakest identity.*

PROOF: For the sake of the argument, suppose that the LA is soft ($c_i > 0$) rather than counter-attitudinal (the same reasoning works in the latter case). For a given p_i, the implementable set is $[\tau^m_i(p_i), \tau^a_i(c_i, p_i)]$, where $\tau^a_i(c_i, p_i)$ is the LA's preferred discriminatory tax, which solves max $(\tau_i - c_i)[1 - F(\tau_i + p_i - r_i)]$ (and $\tau^m_i(p_i)$ solves the same program for $c_i = 0$). So $\tau^a_i(c_i, 0) = \tau^a_i(c_i)$, the discriminatory tax in the persecution-free environment. It is easily shown that persecutions reduce the discriminatory tax as the LA absorbs a fraction of its effect: $\frac{\partial\tau^a_i(c_i,p_i)}{\partial p_i} \in (-1, 0)$. The CA's payoff when the LA sets $\tau^a_i(c_i, p_i)$ (which is the tax in the implementable set that the CA prefers) is $W_i = \lambda_i + [\tau^a_i(c_i, p_i) - c][1 - F(\tau^a_i(c_i, p_i) + p_i - r_i)]$. Simple computations show that $\frac{\partial W_i}{\partial p_i}\mid_{p_i=0} \propto \frac{\partial\tau^a_i(c_i,p_i)}{\partial p_i}\mid_{p_i=0} [c - c_i] + c - \tau^a_i(c_i, 0)$. For $c = c_i$, the first term on the right-hand side (RHS) is equal to 0, while the second term is strictly negative. The RHS is strictly increasing in c, and for c sufficiently large, the CA can guarantee itself λ_i by choosing an infinite level of persecutions and gets strictly less than λ_i when

[82] Persecutions under the Arab Caliphate were ordered by the CA (caliph or governor).

choosing $p_i = 0$. Finally, the cutoff level c_i^* is defined by $\frac{\partial \tau_i^a(c_i, p_i)}{\partial p_i}\,|_{p_i=0}\,[c_i^* - c_i] + c_i^* - \tau_i^a(c_i) = 0$. *Q.E.D.*

Signaling Model of Persecutions

Consider a ruler with unknown religiosity $c \in \{c_L, c_H\}$ with $c_L < c_H$. Make the extreme assumption that the poll tax rate or enforcement, τ, is unobserved by the Muslims, while the level of persecution, p, is perfectly observed.

The ruler is image-concerned and has payoff

$$(\tau - c)\big[1 - F(\tau + p)\big] + \mu \hat{c},$$

where $\hat{c} \equiv E[c|p]$ is the rulers' estimated religiosity conditional on the Muslims' information (p) and μ is the intensity of image concerns. Without loss of generality, we have ignore the additive, nondiscriminatory tax.

We will look for a separating equilibrium $\{\tau_k, p_k\}_{k\in\{L,H\}}$. As usual, the low type behaves as under full information (i.e., as if there were no image concerns: $p_L = 0$). And so let

$$W_L \equiv \max_{\{\tau_L\}} (\tau_L - c_L)\big[1 - F(\tau_L)\big] + \mu c_L$$

denote the low type's separating equilibrium payoff. As for the high type, $\{\tau_H, p_H\}$ is given by the least-cost-separating policy

$$\max_{\{\tau_H, p_H\}} (\tau_H - c_H)\big[1 - F(\tau_H + p_H)\big] + \mu c_H$$

such that

$$\max_{\{\tau\}} (\tau - c_L)\big[1 - F(\tau + p_H)\big] + \mu c_H \le W_L. \tag{A.2}$$

Let us define, letting $\hat{\tau} \equiv \tau + p$,

$$w(c, p) \equiv \max_{\{\hat{\tau}\}} (\hat{\tau} - p - c)\big[1 - F(\hat{\tau})\big].$$

For the separating equilibrium to exist, the sorting condition must be satisfied. This is indeed the case, using the envelope theorem

$$\frac{\partial^2 w}{\partial c \partial p} = f(\hat{\tau}) \frac{\partial \hat{\tau}}{\partial c} > 0,$$

where $\partial \hat{\tau} / \partial c > 0$ results from revealed preference.

The "least-cost-separating equilibrium" corresponds to the value $p_H = p_H^*$, where p_H^* satisfies (A.2) with equality.

Finally, let us see whether the separating equilibrium is consistent with $\tau_H > \tau_L$ (covariation of τ and p). Suppose a uniform distribution on $[0, 1]$: $F(\theta) = \theta$. Then $\tau_H - \tau_L = [(c_H - c_L) - p_H^*]/2$ and p_H^* is given by $\frac{(1-c_L)^2 - (1-c_L-p_H^*)^2}{2} = \mu(c_H - c_L) = \mu \Delta c$. Fix c_L and increase Δc.

Note that for (A.2) to have a solution, it must be the case that $p_H^* + c_L \le 1$, which requires μ not too large. Note also that $(1 - c_L - p_H^*)\, dp_H = \mu d(\Delta c)$. And so

$$\frac{d(\tau_H - \tau_L)}{d(\Delta c)} = \left[1 - \frac{\mu}{1 - c_L - p_H^*}\right] / 2$$

is equal to 1/2 for $\mu = 0$ (for which $p_H^* = 0$ as well) and remains positive as long as μ is not too large. Because at $\mu = 0$, $\tau_H = \tau_L$, then $\tau_H > \tau_L$ in this range.

APPENDIX B: EMPIRICS

B.1. *Data Sources*

B.1.1. *Cross-Sectional Analysis*

- Identity strength of local authorities (c_i): Locations of Arab tribes that settled in Egypt in 700–969 are constructed from al-Barri (1992), a secondary source that draws on *al-bayan wal-i'rab 'amman fi ard misr min al-a'rab* (*Arab Tribes in Egypt*) by al-Maqrizi (died in 1442).
- Identity strength of Copts (r_{ji}): The list of the Holy-Family-visit villages is from Anba Bishoy (1999) and Gabra (2001); both are based on the apocryphal book *Vision of Theophilus* in Mingana (1931). The list of pre-641 Coptic saints and martyrs is from the Coptic *Synaxarium*, *Le Synaxaire arabe-jacobite* translated by Basset (Basset (1907)).
- Proportion of converts (F_{ji}): The list of Coptic churches and monasteries circa 1200 is from *History of Churches and Monasteries* (abul-Makarim (1984)).
- Discriminatory tax rate (τ_{hi}): The individual-level poll tax payments are from Morimoto (1981, pp. 67–79, 85–87) for Greek papyri and the Arabic Papyrology Database for Arabic papyri.
- Total tax transfer (T_{ji}): The village-level data on total tax transfer (*'ibra*) per unit of land are from Ibn al-Ji'an (1898).
- Byzantine-period kura-level controls: The natural logarithm of urban population circa 300 is based on Wilson (2011, pp. 185–187). Byzantine military garrisons circa 600 are constructed from Maspero (1912). Autopract estates circa 600 are constructed from Hardy (1931).
- Geographic village-level controls are from the Food and Agriculture Organization's Global Agro-Ecological Zones (FAO-GAEZ) Data Portal 3.0.1. Crop suitability indices are under irrigation and intermediate input level. Population is from the 1897 population census (Ministère des finances (1898)).

B.1.2. *Time-Series Analysis*

- Poll tax hikes and conversion waves are based on *The Chronicle of John, the Bishop of Nikiu* for 641–661 (John of Nikiu (1916)) and *History of the Patriarchs of the Coptic Church of Alexandria* for 661–1170 (al-Muqaffa (1910, 1943)).
- Identity strength of central authority (c_t): Caliph-level piety (not drinking alcohol) is based on Sirhan (1978) for 641–868, al-Dhahabi's *The Lives of Noble Figures* (al-Dhahabi (1982)) for 868–969, and al-Maqrizi's *History of the Fatimid Caliphs* (al-Maqrizi (1996)) for 969–1170. Governor-level hostility toward nonconvert Copts is based on *The Chronicle of John, the Bishop of Nikiu* for the Rashidun period (641–661) and *History of the Patriarchs of the Coptic Church of Alexandria* for 661–1170.
- Control variables: The yearly number of foreign attacks is constructed from Mikaberidze (2011). The yearly occurrence of a Nile adverse shock is constructed from Chaney (2013).

B.1.3. *Descriptive Figures*

Figure 3 shows the spatial distribution of the cross-sectional outcomes and main regressors. Figures 4 and 5 show the evolution of poll tax hikes, conversion waves, and the total poll tax revenue in 641–1170.

B.2. *Measuring the Proportion of Converts (F_{ji})*

We measure the proportion of converts (F_{ji}) by a village-level dummy variable $= 1$ if there is no Coptic church or monastery in village j located within kura i circa 1200. Our measure is valid under the following assumptions: (a) the universe of villages is observed in 641 (no post-641 villages), (b) every village had at least one Coptic church or monastery in 641, (c) the nonexistence of Coptic churches and monasteries in a village in 1200 is the result of the conversion of the vast majority of its population in 641–1200, which led to the desertion of the churches and monasteries or their transformation into mosques, rather than any nonconversion cause (e.g., abandoning a church for financial reasons), (d) the list of churches and monasteries in 1200 is complete, and (e) there is no differential movement of converts and nonconverts across villages.

These assumptions are supported by a number of observations. In support of (a), we define the universe of villages based on the 1315 cadastre (Ibn al-Ji‘an (1898)).[83] Most of these villages existed before 641 (Ramzi (1954)). As a robustness check, we further restrict our analysis to a subset of villages mentioned in Byzantine-period sources, that was compiled by the French archaeologist Amélineau (1850–1915) (Amélineau (1893)).[84] The results are qualitatively similar. In support of (b), Amélineau's villages are quite large (mean population in 1897 is 5900, compared to 2700 in non-Amélineau villages). Hence, they are most likely to have had at least one church or monastery in 641. In support of (c), our measure is negatively correlated ($\rho = -0.29$) with the actual number of nonconvert Coptic households in 1245 among villages in Fayum kura, based on al-Nabulsi's Fayum cadastre.[85] We also use the individual-level religious affiliation in Egypt's first population censuses in 1848 and 1868 as a robustness check, finding similar results. In support of (d), Abul-Makarim's list is the most complete enumeration of churches and monasteries in medieval Egypt. It has more entries, and geographic coverage than any other list. We obtain similar results if we use al-Maqrizi's list of churches and monasteries circa 1500 as a robustness check.[86] In support of (e), (i) rural–rural migration was outlawed—papyrological administrative records reveal that "fugitives" who fled their villages were forced to go back—and (ii) (tax-induced) rural–urban migration is unlikely because cities were controlled by Arab LAs.[87]

[83]While the earliest extant comprehensive list of Egyptian villages dates to the 1298 cadastre, we chose to digitize the 1315 cadastre instead, because it has information on land area and total tax revenue.

[84]This is not an exhaustive list of pre-641 villages, though; it only includes villages that were large enough to be mentioned in the Byzantine sources.

[85]The number of Coptic households in the kura of Fayum is constructed from Rapoport (2018) based on the 1245 cadastre of Fayum in *Tarikh al-Fayum (History of Fayum)* by al-Nabulsi (died circa 1250).

[86]This is constructed from al-Maqrizi's *al-Mawa‘iz wal-I‘tibar fi Zhikr al-Khitat wal-'Athar* (*Sermons and Considerations in Examining Plans and Monuments*) (al-Maqrizi (2002)).

[87]In 1848, when mobility restrictions and the poll tax were both still enforced, the proportion of rural–rural cross-kura immigrants is not statistically different between Muslims and Copts (5.7% versus 6.1%).

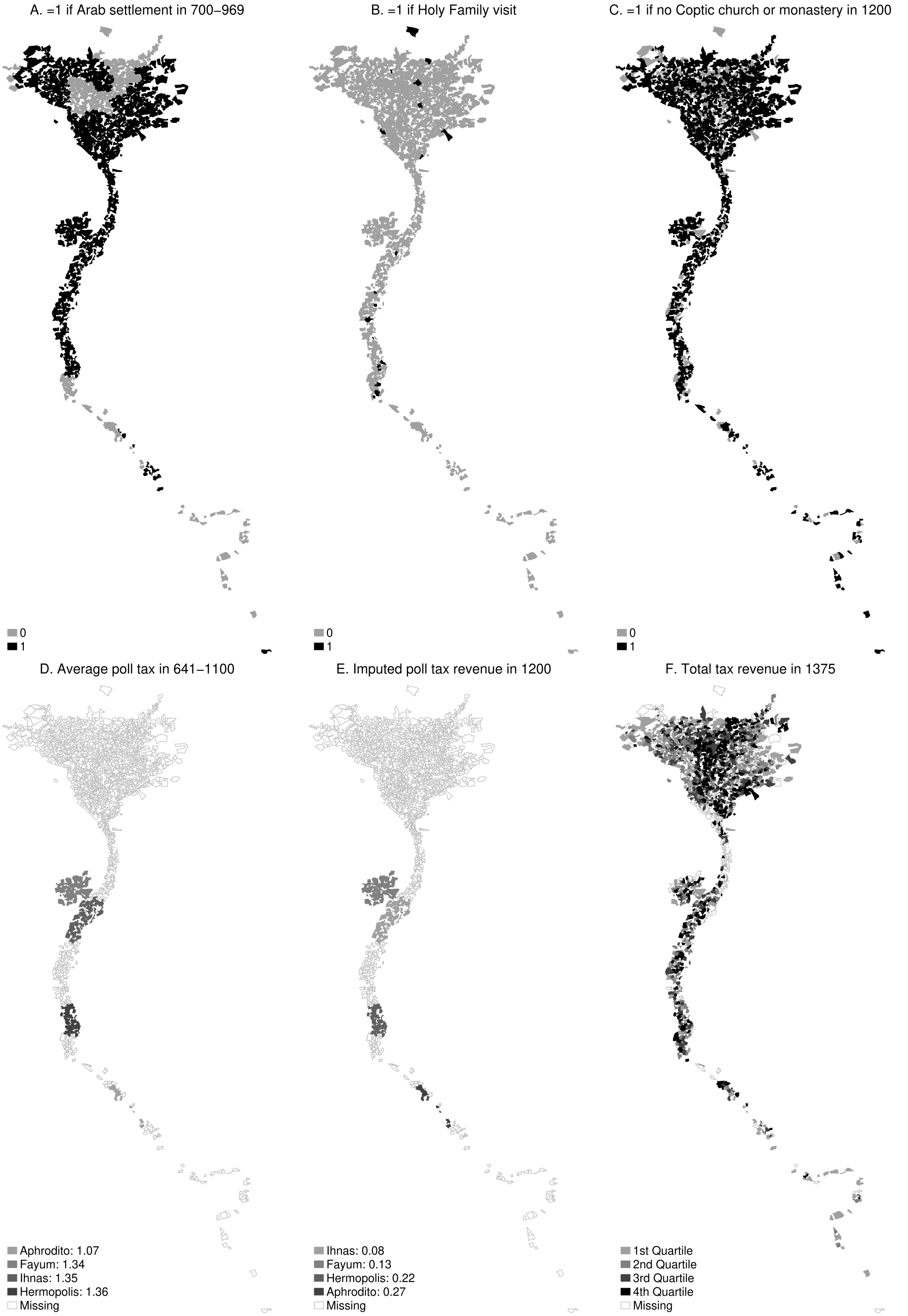

FIGURE 3.—Cross-sectional spatial heterogeneity in determinants and outcomes. The map shows 1782 villages in the 1315 cadastre, which defines our universe of villages, using the boundaries of villages in the 2006 population census. Nile delta refers to the northern triangle on the map. Nile valley covers the whole region to the south of the delta. Source: See Appendix B.1.

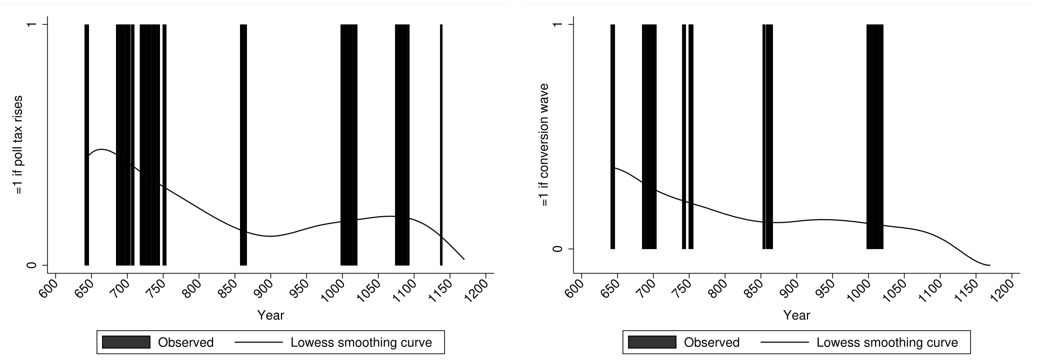

FIGURE 4.—Poll tax hikes and conversion waves in 641–1170. Notes. 641–750: Rashidun and Umayyads; 750–868: First Abbasid Period; 868–969: Tulunids, Second Abbasid Period, Ikhshidids; 969–1170: Fatimids. Source: See Appendix B.1.

B.3. *Measuring Total Tax Revenue (T_{ji})*

Tax transfer per unit of land ($\tilde{T}_{ji}$) is equal to T_{ji} only if population per unit of land and yield per unit of land are both held constant for all j. Letting q_{ji} denote the amount of land, z_{ji} denote the average yield per unit of land, and n_{ji} denote the number of inhabitants, total tax transfer is thus $T_{ji}^{\text{Tot}} = q_{ji} z_{ji} \lambda_{ji} + n_{ji} \tau_{ji}(1 - F_{ji})$. In the theory, we normalized $q_{ji} = z_{ji} = n_{ji} = 1$. We observe $\tilde{T}_{ji} = \frac{T_{ji}^{\text{Tot}}}{q_{ji}} = \lambda_{ji} z_{ji} + \frac{n_{ji}}{q_{ji}} R_{ji}$. Hence, $\tilde{T}_{ji} = T_{ji}$ only if z_{ji} and $\frac{n_{ji}}{q_{ji}}$ are the same for all j. Empirically, we control for z_{ji} by the FAO-GAEZ cereals suitability index and control for $\frac{n_{ji}}{q_{ji}}$ by the population size in 1897 divided by land area in 1315.

B.4. *Supplemental Tables*

This section presents Tables V–X that are referenced in the paper.

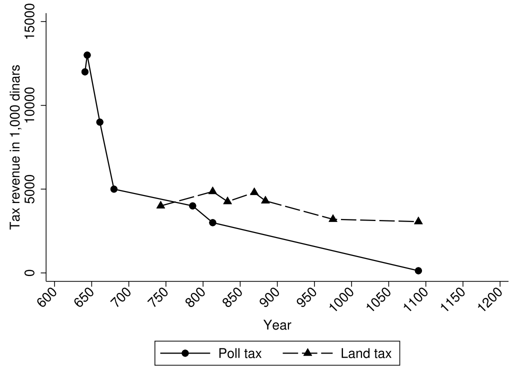

FIGURE 5.—Total poll and land tax revenues in 641–1170. Sources: Courbage and Fargues (1997); poll tax revenue in 1090, Mahmoud (2009).

TABLE V

RELEVANCE AND EXOGENEITY OF THE INSTRUMENTAL VARIABLES[a,b]

	(1) = 1 if Arab Settlement in 700–969	(2) = 1 if Holy Family Visit	(3) Log (Urban Population) in 300	(4) = 1 if Byzantine Garrison in 600	(5) FAO-GAEZ Cereals Suitability	(6) Mean Temperature	(7) Mean Temperature Range	(8) Mean Slope	(9) Mean Rainfall
Kura's distance to Arish (km)	0.011	0.005	−0.005	0.008	−0.001	0.000	0.008	0.002	−0.198
	(0.004)	(0.006)	(0.001)	(0.005)	(0.000)	(0.003)	(0.004)	(0.003)	(0.287)
= 1 if kura borders desert	3.715	1.168	−1.004	2.223	−0.130	−1.982	0.529	0.093	−36.071
	(1.020)	(1.419)	(0.492)	(1.211)	(0.112)	(0.668)	(1.058)	(0.677)	(67.489)
Bordering desert × dist. Arish	−0.012	−0.006	0.006	−0.008	0.000	0.009	0.001	0.000	0.053
	(0.004)	(0.006)	(0.002)	(0.005)	(0.000)	(0.003)	(0.004)	(0.003)	(0.288)
Observations	42	42	42	42	42	42	42	42	42
R^2	0.376	0.134	0.109	0.061	0.377	0.828	0.776	0.383	0.379

[a] White–Huber robust standard errors are given in parentheses. A constant term is included in all regressions.

[b] Sources: See Appendix B.1.

TABLE VI

LOCAL DETERMINANTS OF CONVERSIONS TO ISLAM IN 641–1200 AND TOTAL TAX TRANSFER IN 1375: NO REGION FIXED EFFECTS[a,b]

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
(a) Dependent Variable: = 1 if no Coptic Church or Monastery in Village j Circa 1200									
= 1 if Arab settlement (c_i)	0.08		0.08	0.08	0.08	0.13	0.12	0.13	0.12
	(0.03)		(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.06)	(0.05)
= 1 if HF visit (r_{ji})		−0.59	−0.58	−0.59	−0.62		−0.58	−0.59	−0.62
		(0.08)	(0.08)	(0.08)	(0.09)		(0.08)	(0.08)	(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	1782	1782	1782	1751	1782	1782	1782	1751
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	17.33	16.40	16.65
(b) Dependent Variable: Tax Transfer (ʻibra) in Army Dinars per Unit of Land in 1375									
= 1 if Arab settlement (c_i)	−0.13		−0.13	−0.30	−0.24	−0.46	−0.45	−0.57	−0.36
	(0.30)		(0.30)	(0.28)	(0.21)	(0.35)	(0.34)	(0.33)	(0.30)
= 1 if HF visit (r_{ji})		0.96	0.96	0.86	0.73		0.97	0.86	0.73
		(0.41)	(0.42)	(0.46)	(0.53)		(0.43)	(0.46)	(0.53)
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	0.01	0.05	−0.00	−0.00	0.01	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	16.42	16.17	14.65

[a]Tax transfer (ʻibra) is in army dinars (≈13.3/20 dinars) per feddan (= 6368 square meters). Standard errors clustered at the kura level are given in parentheses. Byzantine-period kura-level controls are (i) the logarithm of urban population in kura i circa 300 and (ii) a dummy variable equal to 1 if there was a Byzantine garrison in kura i circa 600. Geographic village-level controls are (iii) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (iv) mean temperature, (v) temperature range, (vi) slope, and (vii) rainfall. Population per unit of land is (viii) the population in 1897 divided by land area in 1315. A constant is included in all regressions.

[b]Sources: See Appendix B.1.

TABLE VII

LOCAL DETERMINANTS OF CONVERSIONS TO ISLAM IN 641–1200: ROBUSTNESS CHECKS[a,b]

	Amélineau's Byzantine-Era Villages		= 1 if no Church or Mon. Circa 1500		= 1 if Copt in 1848–1868		Control Interaction ($c_i \times r_{ji}$)		Control Autopract Circa 600		Alternative Measure of r_{ji} (Saint–Martyr)		SARAR Model	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS	(9) OLS	(10) 2SLS	(11) OLS	(12) 2SLS	(13) OLS	(14) 2SLS
	Dependent Variable: = 1 if no Coptic Church or Monastery in Village j Circa 1200 (except in columns 3–6)													
= 1 if Arab settlement (c_i)	0.16	0.08	0.04	0.03	−0.03	−0.05	0.07	0.09	0.10	0.07	0.07	0.08	0.06	0.10
	(0.08)	(0.15)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.05)	(0.07)	(0.19)	(0.03)	(0.05)	(0.02)	(0.03)
= 1 if HF visit (r_{ji})	−0.60	−0.59	−0.27	−0.27	0.12	0.13	−0.64	−0.50	−0.64	−0.64			−0.62	−0.62
	(0.08)	(0.07)	(0.08)	(0.07)	(0.04)	(0.04)	(0.18)	(0.28)	(0.11)	(0.11)			(0.08)	(0.08)
$c_i \times r_{ji}$							0.03	−0.17						
							(0.21)	(0.35)						
= 1 if autopract circa 600									−0.01	−0.01				
									(0.07)	(0.07)				
= 1 if saint–martyr in village											−0.51	−0.51		
											(0.09)	(0.09)		
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Byzantine controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs. (villages)	157	157	1751	1751	16,195	16,195	1751	1751	575	575	1748	1748	1730	1730
Clusters (kuras)	37	37	42	42	42	42	42	42	21	21	42	42		
R^2	0.15	0.14	0.12	0.12	0.09	0.09	0.06	0.06	0.09	0.09	0.05	0.05		
Mean dep. var. in control	0.54	0.54	0.95	0.95	0.12	0.12	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat.		16.33		19.32		5.33		10.99		6.90		19.38		

[a] Robust standard errors clustered at the kura level are given in parentheses. Regions, Byzantine-period controls, and geographic controls are defined as in Table I. In column 8, 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 (i) kura's distance to Arish (DistancetoArish$_i$) (ii) = 1 if kura borders desert (BorderDesert$_i$), (iii) DistancetoArish$_i$ × BorderDesert$_i$, (iv) r_{ji} × DistancetoArish$_i$, (v) r_{ji} × BorderDesert$_i$, (vi) r_{ji} × DistancetoArish$_i$ × BorderDesert$_i$. In column 10, the excluded instrument in the first-stage regression for Arab settlement (c_i) is DistancetoArish$_i$. Column 13 reports the results of estimating a spatial autoregressive model with spatial autoregressive standard errors (SARAR) with inverse distance weighting matrix estimated using generalized spatial two-stage least squares (GS2SLS) (STATA command spreg). Column 14 reports the results of estimating a SARAR model with endogenous variables (STATA command spivreg). The full results are shown in Appendix D.

[b] Sources: See Appendix B.1. Columns 1 and 2, Byzantine-era villages listed in Amélineau (1893); columns 3 and 4, list of Coptic churches and monasteries circa 1500 in al-Maqrizi (2002); columns 5 and 6, 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 no missing information on age, religion, occupation, and district of origin.

TABLE VIII

EVALUATING THE NATIONAL REPRESENTATIVENESS OF THE POLL TAX SUBSAMPLE[a,b]

(a) Village-Level Variables

	Villages Out of Poll-Tax Sample			Villages in Poll-Tax Sample			
	N	Mean	SD	N	Mean	SD	Diff.
= 1 if no church or monastery in 1200	1587	0.83	0.37	195	0.89	0.31	0.061
= 1 if no church or monastery in 1500	1587	0.98	0.15	195	0.94	0.24	−0.038
ibra per feddan in 1375	1335	3.22	2.26	176	4.01	6.40	0.793
ibra per feddan in 1477	1335	2.76	1.95	176	3.51	6.45	0.749
= 1 if on HF route	1587	0.01	0.11	195	0.03	0.16	0.014
= 1 if pre-641 Coptic saint or martyr	1583	0.01	0.12	195	0.03	0.17	0.016
FAO-GAEZ cereals suitability index	1560	0.68	0.10	191	0.66	0.10	−0.024
Mean temperature	1560	20.98	0.82	191	21.88	0.30	0.899
Mean temperature range	1560	14.17	1.04	191	16.34	0.23	2.167
Mean slope	1560	3.43	0.61	191	3.90	0.63	0.467
Mean rainfall	1560	50.26	33.27	191	6.43	3.31	−43.832

(b) Kura-Level Variables

	Kuras Out of Poll-Tax Sample			Kuras in Poll-Tax Sample			
	N	Mean	SD	N	Mean	SD	Diff.
= 1 if Arab settlement in 700–969	38	0.63	0.49	4	0.75	0.50	0.118
Log (urban population) in 300	38	10.00	0.73	4	10.57	0.72	0.570
= 1 if Byzantine garrison in 600	38	0.42	0.50	4	1.00	0.00	0.579
Kura's distance to Arish (km)	38	354.07	148.34	4	425.86	83.63	71.792
= 1 if kura borders desert	38	0.76	0.43	4	1.00	0.00	0.237

(c) Dep. Var. = 1 if no Coptic Church or Monastery in Village j Circa 1200: Poll Tax Subsample

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
= 1 if Arab settlement (c_i)	0.15		0.17	0.24	0.64	0.33	0.24	0.26	0.85
	(0.56)		(0.58)	(0.17)	(0.35)	(0.21)	(0.13)	(0.22)	(0.34)
= 1 if HF visit (r_{ji})		−0.51	−0.52	−0.50	−0.66		−0.52	−0.50	−0.66
		(0.78)	(0.78)	(0.69)	(0.92)		(0.77)	(0.76)	(0.83)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs. (villages)	195	195	195	195	191	195	195	195	191
Clusters (kuras)	4	4	4	4	4	4	4	4	4
R^2	0.01	0.07	0.08	0.08	0.12	−0.01	0.08	0.08	0.12
Mean dep. var. in control	0.75	0.91	0.75	0.75	0.75	0.75	0.75	0.75	0.75
KP Wald F-stat.						1.11	1.17	395.65	910.85

[a]"Diff." reports the slope in the regression: $y = \alpha_1 + \alpha_2 \text{polltaxsample}_i + \epsilon$, where y is the outcome of village j located in kura i in panel (a) or the outcome of kura i in panel (b), and $\text{polltaxsample}_i = 1$ if kura i is in the poll tax subsample. Standard errors are clustered at the kura level in panel (a) and are White–Huber SEs in panel (b). Panel (c): The IV in columns 6–9 is the kura's distance to Arish. The p-values are given in parentheses, estimated by clustering standard errors at the kura level, using wild cluster restricted bootstrap for OLS and wild restricted efficient clustered bootstrap for IV, with Webb weights and 999,999 replications. A constant is included in all regressions. Controls are defined as in Table I.

[b]Sources: See Appendix B.1.

TABLE IX

LOCAL DETERMINANTS OF THE TOTAL TAX TRANSFER IN 1375: ROBUSTNESS CHECKS[a,b]

	Tax Transfer in 1477		Province FE Two-Way Clustering		Control Interaction ($c_i \times r_{ji}$)		Control Mamluk LA in 1375		Alternative Measure of r_{ji} (Saint–Martyr)		SARAR Model	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS	(9) OLS	(10) 2SLS	(11) OLS	(12) 2SLS
Dependent Variable: Tax Transfer (‘ibra) in Army Dinars per Unit of Land in 1375 (except in columns 1 and 2)												
= 1 if Arab settlement (c_i)	−0.39	−0.65	−0.46	−0.33	−0.50	−0.81	−0.45	−0.68	−0.45	−0.69	−0.19	−0.12
	(0.16)	(0.26)	(0.10)	(0.12)	(0.20)	(0.33)	(0.18)	(0.27)	(0.18)	(0.28)	(0.21)	(0.26)
= 1 if HF visit (r_{ji})	0.68	0.69	0.84	0.83	0.90	1.04	0.82	0.82			0.86	0.86
	(0.53)	(0.53)	(0.45)	(0.38)	(1.08)	(1.30)	(0.45)	(0.45)			(0.70)	(0.70)
$c_i \times r_{ji}$					0.19	0.01						
					(1.19)	(1.42)						
= 1 if LA in 1375 Mamluk							0.70	0.70				
							(0.18)	(0.18)				
= 1 if saint–martyr (r_{ji})									0.54	0.54		
									(0.65)	(0.64)		
Region FE?	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE?	No	No	Yes	Yes	No	No	No	No	No	No	No	No
Byzantine controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population per unit of land?	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Obs. (villages)	1485	1485	1467	1467	1486	1486	1460	1460	1482	1482	1456	1456
Clusters (kuras)	40	40	40	40	40	40	40	40	40	40		
Clusters (provinces)			19	19								
R^2	0.07	0.07	0.10	0.10	0.04	0.04	0.08	0.08	0.06	0.06		
Mean dep. var. in control	2.89	2.89	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.42	3.42
KP Wald F-stat.		19.94		58.59		10.11		20.06		20.07		

[a]Army dinar ≈13.3/20 dinars. Feddan = 6368 square meters. Robust standard errors clustered at the kura level are given in parentheses. In columns 3 band 4, standard errors are clustered at both the kura and province level (STATA commands reghdfe and ivreg2), where provinces are defined according to the administrative division in the 1315/1375 cadastre. Regions, Byzantine-period controls, and geographic controls are defined as in Table I. Population per unit of land is the population in 1897 divided by land area in 1315. In column 6, the excluded instruments are (i) kura's distance to Arish (DistancetoArish$_i$) (ii) = 1 if kura borders desert (BorderDesert$_i$), (iii) DistancetoArish$_i$ × BorderDesert$_i$, (iv) r_{ji} × DistancetoArish$_i$, (v) r_{ji} × BorderDesert$_i$, (vi) r_{ji} × DistancetoArish$_i$ × BorderDesert$_i$. Column 11 is a spatial autoregressive model with spatial autoregressive standard errors (SARAR) with inverse distance weighting matrix estimated using generalized spatial two-stage least squares (GS2SLS) (STATA command spreg). Column 12 is a SARAR model with endogenous variables (STATA command spivreg). The full results are shown in Appendix D.

[b]Sources: See Appendix B.1.

TABLE X

TIME-SERIES DETERMINANTS OF POLL TAX HIKES ($\Delta\tau_t$) AND CONVERSION WAVES (ΔF_t) IN 641–1170: ROBUSTNESS CHECKS[a]

	Governor-Level Data Set			Control No. of Prev. Tax Hikes			Control $\hat{c}_t \times n^c_{t-1}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(a) Dependent Variable = 1 if a Poll Tax Hike Mentioned During the Reign of Ruler t									
= 1 if ruler's identity strong ($\hat{c}_t$)	0.53	0.53	0.50	0.21	0.16	0.17	0.15	0.20	−0.19
	(0.15)	(0.14)	(0.13)	(0.10)	(0.12)	(0.14)	(0.22)	(0.27)	(0.44)
No. of prev. strong identity rulers (n^c_{t-1})	−0.00	−0.02	−0.05				−0.02	−0.01	−0.03
	(0.01)	(0.02)	(0.04)				(0.01)	(0.03)	(0.03)
No. of previous poll tax hikes				−0.04	−0.07	−0.09			
				(0.02)	(0.04)	(0.07)			
$\hat{c}_t \times n^c_{t-1}$							0.01	0.00	0.03
							(0.01)	(0.01)	(0.02)
Ruler's start year		0.38	0.78		0.50	0.96		−0.27	−0.11
		(0.35)	(0.76)		(0.45)	(0.96)		(0.60)	(0.66)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs. (governors/caliphs)	121	121	121	64	64	64	64	64	64
Years	526	526	526	526	526	526	526	526	526
R^2	0.29	0.30	0.33	0.21	0.23	0.27	0.16	0.16	0.20
p-value (Breusch–Godfrey test)	0.01	0.00	0.00	0.04	0.02	0.01	0.06	0.04	0.01
Mean dep. var.	0.13	0.13	0.13	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 Ruler t									
= 1 if ruler's identity strong ($\hat{c}_t$)	0.47	0.47	0.46	0.19	0.26	0.31	0.32	0.64	0.29
	(0.19)	(0.19)	(0.19)	(0.13)	(0.19)	(0.20)	(0.33)	(0.41)	(0.67)
No. of prev. strong identity rulers (n^c_{t-1})	0.00	−0.00	0.01				−0.01	0.05	0.03
	(0.01)	(0.01)	(0.02)				(0.01)	(0.03)	(0.04)
No. of prev. poll tax hikes				−0.04	0.01	0.04			
				(0.01)	(0.05)	(0.07)			
$\hat{c}_t \times n^c_{t-1}$							−0.00	−0.02	−0.00
							(0.01)	(0.02)	(0.03)
Ruler's start year		0.05	−0.12		−0.81	−1.21		−1.47	−1.36
		(0.20)	(0.44)		(0.72)	(1.07)		(0.48)	(0.57)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs. (governors/caliphs)	121	121	121	64	64	64	64	64	64
Years	526	526	526	526	526	526	526	526	526
R^2	0.33	0.33	0.38	0.21	0.26	0.30	0.19	0.29	0.31
p-value (Breusch–Godfrey test)	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.00
Mean dep. var.	0.07	0.07	0.07	0.18	0.18	0.18	0.18	0.18	0.18

[a]Newey–West standard errors are given in parentheses, assuming that the error is both heteroskedastic and autocorrelated up to 15 lags (governors) and 11 lags (caliphs). Controls are (i) = 1 if foreign attack occurred and (ii) = 1 if an adverse Nile shock occurred. Ruler's start year is normalized to be an element of [0, 1]. Regressions are weighted by the length of ruler's tenure. The H_0 for the Breusch–Godfrey test is that there is no serial correlation up to 15 lags (governors) and 11 lags (caliphs). A constant is included in all regressions. Sources: See Appendix B.1.

REFERENCES

ABUL-MAKARIM, S. (1984): *Tarikh al-Kana'is wal-Adyera (History of Churches and Monasteries)*. Cairo: Dar al-Na'am. Anba Samuel (Ed.). [11]

AL-BARRI, A. K. (1992): *al-Qaba'il al-'Arabiya fi Misr fil Qurun al-Thalatha al-Ula lil Hijra (Arab Tribes in Egypt in the First Three Hijri Centuries)*. Cairo: Egyptian General Book Authority. [11]

AL-DHAHABI, S. (1982): *Siyar A'lam al-Nubala' (the Lives of Noble Figures)*. Beirut: al-Risala. [11]

AL-MAQRIZI, T. (1996): *Itti'az al-Ḥunafa' bi-Akhbar al-A'immah al-Fatimiyin al-Khulafa' (History of the Fatimid Caliphs)*. Cairo: Islamic Affairs Supreme Council. Jamal-el-Din al-Shayyal (Ed.). [11]

——— (2002): *al-Mawa'iz wal-I'tibar fi Zhikr al-Khitat wal-'Athar (Sermons and Considerations in Examining Plans and Monuments)*, Vol. 4, ed. by A. F. Sayyid. London: Al-Furqan Islamic Heritage Foundation. [12, 17]

AL-MUQAFFA, S. (1910): "History of the Patriarchs of the Coptic Church of Alexandria, Volume I," in *Patrologia Orientalis*, Vol. 1, 5, 10. Paris: Firmin-Didot Et Co. B. Evetts (Tr.). [11]

——— (1943): "History of the Patriarchs of the Coptic Church of Alexandria, Volume II: Part I-III, Volume III: Part I," Publications de la Société d'Archéologie Copte, Cairo. Yassa Abd al-Masih, Aziz Suryal Atiya, O.H.E. Burmester, Antoine Khater (Trs. and Eds.), 1943, 1948, 1959, 1968. [11]

AMÉLINEAU, E. (1893): *La Géographie de l'Egypte à l'Epoque Copte*. Paris: Imprimerie nationale. [12,17]

ANBA BISHOY, P. (1999): *The Flight of the Holy Family to Egypt*. Cairo: Nubar Printing House. [11]

BASSET, R. (1907): "Le Synaxaire Arabe-Jacobite (Redaction Copte)," in *Patrologia Orientalis*, ed. by R. Blundell, W. Newey, and T. Persson. Paris: Firmin-Didot, 1907–1929. [11]

CHANEY, E. (2013): "Revolt on the Nile: Economic Shocks, Religion, and Political Power," *Econometrica*, 81, 2033–2053. [11]

COURBAGE, Y., AND P. FARGUES (1997): *Christians and Jews Under Islam. I*. London-New York: B. Tauris Publishers. J. Mabro (Tr.). [14]

GABRA, G. (2001): *Be Thou There: The Holy Family's Journey in Egypt*. Cairo: American University in Cairo Press. [11]

HARDY, E. R. (1931): *The Large Estates of Byzantine Egypt*. New York: Columbia University Press. [11]

IBN AL-JI'AN (1898): *al-Tuhfa al-Saniyya bi Asma' al-Bilad al-Misriya (Book of Names of Egyptian Localities)*. Cairo: al-Matba'a al-Ahliya. [11,12]

JOHN OF NIKIU (1916): *The Chronicle of John, Bishop of Nikiu*. Williams and Norgate, London-Oxford. R. H. Charles (Tr.). [11]

MAHMOUD, A. A. (2009): *al-jizya fi misr 1713-1856 (The Poll Tax in Egypt in 1713-1856)*. Cairo: Supreme Council of Culture. [14]

MASPERO, J. (1912): *Organisation Militaire de l'Égypte Byzantine*. Paris: Bibliothèque de l'École des Hautes-Études. [11]

MIKABERIDZE, A. (2011): *Conflict and Conquest in the Islamic World: A Historical Encyclopedia [2 Volumes]: A Historical Encyclopedia*. Santa Barbara, CA: ABC-CLIO. [11]

MINGANA, A. (1931): "Vision of Theophilus," in *Woodbrooke Studies*, Vol. 3. Limited, Cambridge: W. Heffer and Sons, 1–92. [11]

MINISTÈRE DES FINANCES (1898): *Recensement Général de l'Egypte: 1 juin 1897*. Le Caire: Imprimerie nationale. [11]

MORIMOTO, K. (1981): *The Fiscal Administration of Egypt in the Early Islamic Period*. Kyoto: Dohosha. [11]

RAMZI, M. (1994) [1954]: *al-Qamus al-Gughrafi lil Bilad al-Misriya min 'Ahd Qudama' al-Misriyin ila Sanat 1945 (Geographic Dictionary of Egyptian Localities From the Time of Ancient Egyptians to 1945)*. Cairo: Egyptian General Book Authority. [12]

RAPOPORT, Y. (2018): *Rural Economy and Tribal Society in Islamic Egypt: A Study of al-Nabulusi's Villages of the Fayyum*. Turnhout: Brepols Publishers. [12]

SIRHAN, J. (1978): "al-Musamara wal-Munadama 'ind al-'Arab hatta al-Qarn al-Rabi'al-Hijri (the Boon-Companionship in the Arab World Until the 4th Century A.H.)," PhD thesis, American University of Beirut. Beirut. [11]

TIROLE, J. (2016): "From Bottom of the Barrel to Cream of the Crop: Sequential Screening With Positive Selection," *Econometrica*, 84 (4), 1291–1343. [2]

WILSON, A. (2011): "City Sizes and Urbanization in the Roman Empire," in *Settlement, Urbanization, and Population*, ed. by A. Bowman and A. Wilson. Oxford: Oxford University Press, 161–195. [11]

Co-editor Fabrizio Zilibotti handled this manuscript.

Manuscript received 29 April, 2019; final version accepted 10 March, 2021; available online 23 March, 2021.