Online Appendices

"From Investiture to Worms:

European Development and the Rise of Political Authority"

Journal of Politics

Bruce Bueno de Mesquita   Ethan Bueno de Mesquita

A Formal Results on Interregna

There is an interregnum if a nominee is rejected. In any given period, this occurs with probability:

$$\Pr(\text{interregnum}) = F(\bar{\epsilon}^*(r^*)).$$

The expected length of an interregnum, conditional on one occurring, is

$$\text{Expected Length of Interregnum} = \frac{1}{1 - F(\bar{\epsilon}^*(r^*))}.$$

The comparative statics of these two quantities are the same.

At an interior r^*, increasing the income of a diocese has competing effects. On the one hand, as income increases, the Ruler becomes less inclined to accept nominees. On the other hand, the Church becomes keener to have its nominee accepted and, as such, offers nominees who are more aligned with the Ruler. Whether the probability and length of an interregnum increases or decreases as a function of income thus depends on whether the Ruler's demands or the Church's willingness to accommodate change more. As the next result shows, this depends on the relative weights that the Ruler and Church put on economic gain versus bishop alignment.

Proposition A.1 *In a stationary equilibrium, if r^* is interior, then the probability and expected length of an interregnum is strictly increasing in y if $\lambda_R < \lambda_C$, strictly decreasing in y if $\lambda_R > \lambda_C$, and constant in y if $\lambda_R = \lambda_C$.*

If r^* is a corner solution, the probability and expected length of an interregnum is strictly increasing in y.

Proof. Differentiating:

$$\frac{d\,\Pr(\text{interregnum})}{dy} = f(\bar{\epsilon}^*(r^*))\left(\frac{\partial\bar{\epsilon}^*(r^*)}{\partial y} + \frac{\partial\bar{\epsilon}^*(r^*)}{\partial y}\frac{\partial r^*}{\partial y}\right).$$

At a corner solution, this has the same sign as $\frac{\partial\bar{\epsilon}^*(r^*)}{\partial y}$, which is positive, by Lemma 4.

At an interior solution, we can substitute for $\frac{\partial r^*}{\partial y}$ from Equation 10. Doing so, this derivative has the same sign as:

$$\frac{\partial\bar{\epsilon}^*(r^*)}{\partial y} - \frac{\partial\bar{\epsilon}^*(r^*)}{\partial y}\left(\frac{\left(\frac{f}{1-F}\right)'(\bar{\epsilon}^*(r^*))\frac{\partial\bar{\epsilon}^*(r^*)}{\partial y} + \frac{2\delta\lambda_C(1-\delta F(\bar{\epsilon}^*(r^*)))f(\bar{\epsilon}^*(r^*))\frac{d\bar{\epsilon}^*(r^*)}{dy}}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)} + \frac{\lambda_C^2(1-\delta F(\bar{\epsilon}^*(r^*)))^2}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)^2}}{\left(\frac{f}{1-F}\right)'(\bar{\epsilon}^*(r^*))\frac{d\bar{\epsilon}^*(r^*)}{dr} + \frac{2\delta\lambda_C(1-\delta F(\bar{\epsilon}^*(r^*)))f(\bar{\epsilon}^*(r^*))\frac{d\bar{\epsilon}^*(r^*)}{dr}}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)} + \frac{\lambda_C^2(1-\delta F(\bar{\epsilon}^*(r^*)))^2}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)^2}}\right).$$

Cross multiplying and rearranging this has the same sign as:

$$\frac{\partial\bar{\epsilon}^*(r^*)}{\partial y}\lambda_C^2 - \frac{\partial\bar{\epsilon}^*(r^*)}{\partial r}\lambda_C(1-\lambda_R).$$

Substituting for $\frac{\partial\bar{\epsilon}^*(r^*)}{\partial y}$ and $\frac{\partial\bar{\epsilon}^*(r^*)}{\partial r}$ from Lemma 4, this has the same sign as:

$$(1-\lambda_R)\lambda_C - \lambda_R(1-\lambda_C),$$

as required. The argument for expected length is analogous. ■

Proposition A.1 shows that, according to theory, the relationship between diocesan income and the occurrence or length of interregna could go either way. Without measures of λ_R and λ_C, we don't know whether they should be positively or negatively associated.

We pause to note that it is not ex ante obvious which should be larger, λ_R or λ_C. One might think the Church must care more than lay rulers about religious policy. But recall from Section 2 that bishops were important political figures with considerable sway over a broad range of issues in a lay ruler's domain. Thus, it is entirely possible that a lay ruler, with a more limited domain of secure political control, might care more about the alignment of local bishops than the central Church whose potential political power extended across all of Europe. Moreover, there might be heterogeneity among lay rulers and Church leaders regarding the relative importance of income and political control.

B Proofs of Results in Main Text

Proof of Lemma 2. Suppose the Pope uses a strategy that calls for proposing r in each period. Slightly abuse notation by writing the Church's strategy as simply r.

From Equation 1, if a bishop of type r_t is proposed in period t, the Ruler accepts if:

$$\epsilon_t \geq \lambda_R \left(\frac{r_t}{1-\delta} - q \right) + (1-\lambda_R)y - c + \delta \max_{s_R} V_R(s_R, r).$$

Notice, since the Church's strategy is stationary, this condition is the same in all periods t. The one-shot-deviation principle thus establishes that the Ruler's strategy is stationary. ■

Proof of Lemma 3. Using the argument in the proof of Lemma 2, if the Ruler's strategy is a best response to a stationary strategy, r, by the Church, then it is an $\overline{\epsilon}^*(\cdot)$ satisfying:

$$\overline{\epsilon}^*(r_t) = \lambda_R \left(\frac{r_t}{1-\delta} - q \right) + (1-\lambda_R)y - c + \delta \max_{s_R} V_R(\overline{\epsilon}^*(\cdot), r), \tag{3}$$

for each r_t. If the Church conjectures that the Ruler is using such a strategy, the Church's expected utility from proposing r_t in period t and using a stationary strategy in which it proposes $\hat{r}$ in all other periods is:

$$(1 - F(\bar{\epsilon}^*(r_t))) \left(\frac{\lambda_C r_t + (1-\lambda_C)y}{1-\delta} \right) + F\left(\bar{\epsilon}^*(r_t)\right)\left(\lambda_C q + \delta V_C(\bar{\epsilon}^*(\cdot), \hat{r})\right).$$

The one-shot-deviation principle implies that the best response in period t must maximize this expected utility. Moreover, since this problem is the same in every period, the Church's strategy must be stationary as long as the optimum is unique.

To see that the optimum is unique, first suppose it is interior. Then it satisfies the following first-order condition:

$$\frac{f}{1-F}(\bar{\epsilon}^*(r_t^*))\frac{d\bar{\epsilon}^*(r_t^*)}{dr_t} = \frac{1-\delta}{\lambda_C}\left(\frac{\lambda_C r_t^* + (1-\lambda_C)y}{1-\delta} - \left(\lambda_C q + \delta V_C(\bar{\epsilon}^*(\cdot), \hat{r})\right)\right)^{-1}.$$

From Equation 3, we have

$$\frac{d\bar{\epsilon}^*(r_t^*)}{dr_t} = \frac{\lambda_C}{1-\delta}.$$

We can, thus, rewrite the first-order condition as:

$$\frac{f}{1-F}(\bar{\epsilon}^*(r_t^*)) = \left(\frac{1-\delta}{\lambda_C}\right)^2\left(\frac{\lambda_C r_t^* + (1-\lambda_C)y}{1-\delta} - \left(\lambda_C q + \delta V_C(\bar{\epsilon}^*(\cdot), \hat{r})\right)\right)^{-1}. \tag{4}$$

The right-hand side of Equation 4 is strictly decreasing in r_t^* and the log-concavity of f implies that the left-hand side is strictly increasing, so there is a unique solution to the first-order condition.

For this to be a stationary best response, we need $r_t^* = \hat{r}$. Hence, there will be a unique, stationary, interior best response if there is a unique $r^* \in [\underline{r}, \overline{r}]$ that satisfies:

$$\frac{f}{1-F}(\bar{\epsilon}^*(r^*)) = \left(\frac{1-\delta}{\lambda_C}\right)^2\left(\frac{\lambda_C r^* + (1-\lambda_C)y}{1-\delta} - \left(\lambda_C q + \delta V_C(\bar{\epsilon}^*(\cdot), r^*)\right)\right)^{-1}. \tag{5}$$

Using the standard recursive approach, if the Church uses a stationary strategy, r, its continuation value for the game is implicitly defined by:

$$V_C(\bar{\epsilon}^*(\cdot), r) = (1 - F(\bar{\epsilon}^*(r))) \left(\frac{\lambda_C r + (1-\lambda_C)y}{1-\delta}\right) + F\left(\bar{\epsilon}^*(r)\right)\left(\lambda_C q + \delta V_C(\bar{\epsilon}^*(\cdot), r)\right).$$

Rearranging, this yields:

$$V_C(\bar{\epsilon}^*(\cdot), r) = \frac{(1 - F(\bar{\epsilon}^*(r)))\left(\frac{\lambda_C r + (1-\lambda_C)y}{1-\delta}\right) + F\left(\bar{\epsilon}^*(r)\right)\lambda_C q}{1 - \delta F(\bar{\epsilon}^*(r))}. \tag{6}$$

Substituting this into Equation 5, a stationary, interior best response exists if there is an $r^* \in [\underline{r}, \overline{r}]$ satisfying:

$$\frac{f}{(1-F)}(\bar{\epsilon}^*(r^*)) = \frac{\lambda_C(1 - \delta F(\bar{\epsilon}^*(r^*)))}{\lambda_R(1-\delta)}\left(\frac{1}{\lambda_C(r^* - q) + (1-\lambda_C)y}\right), \tag{7}$$

as required in the statement of the Lemma. The left-hand side of Equation 7 is increasing and the right-hand side is decreasing, so r^* is unique if it exists.

It is straightforward from the first-order condition that if y is sufficiently large or small, then there is not an r^* satisfying Equation 7. Define $\overline{y}$ such that

$$\frac{f}{(1-F)}(\bar{\epsilon}^*(\underline{r})) = \frac{\lambda_C(1 - \delta F(\bar{\epsilon}^*(\underline{r})))}{\lambda_R(1-\delta)}\left(\frac{1}{\lambda_C(\underline{r} - q) + (1-\lambda_C)\overline{y}}\right).$$

And define $\underline{y}$ such that

$$\frac{f}{(1-F)}(\bar{\epsilon}^*(\overline{r})) = \frac{\lambda_C(1 - \delta F(\bar{\epsilon}^*(\overline{r})))}{\lambda_R(1-\delta)}\left(\frac{1}{\lambda_C(\overline{r} - q) + (1-\lambda_C)\underline{y}}\right).$$

Then r^* exists for any $y \in [\underline{y}, \overline{y}]$ and does not otherwise.

Now consider $y \notin [\underline{y}, \overline{y}]$. Precisely the argument given above implies that $\overline{r}$ is a stationary best response if $y < \underline{y}$ and $\underline{r}$ is a stationary best response if $y > \overline{y}$.

■

The following Lemma will be useful throughout:

Lemma 4

$$\frac{\partial \bar{\epsilon}^*(r^*)}{\partial r^*} = \frac{\lambda_R}{1-\delta F(\bar{\epsilon}^*(r^*))}.$$

$$\frac{\partial \bar{\epsilon}^*(r^*)}{\partial y} = \frac{1-\lambda_R}{1-\delta F(\bar{\epsilon}^*(r^*))}$$

$$\frac{\partial \bar{\epsilon}^*(r^*)}{\partial c} = \frac{-1}{1-\delta F(\bar{\epsilon}(r^*))}.$$

Proof. From Lemma 2, the cutoff rule that the Ruler uses in response to a nomination r^* is implicitly defined by:

$$\bar{\epsilon}^*(r^*) = \lambda_R \left(\frac{r^*}{1-\delta} - q \right) + (1-\lambda_R)y + \delta V_R(\bar{\epsilon}^*(\cdot), r^*).$$

Now, use the standard recursive approach to calculate $V_R(\bar{\epsilon}^*(\cdot), r^*)$. First, write:

$$V_R(\bar{\epsilon}^*(\cdot), r^*) = (1 - F(\bar{\epsilon}^*(r^*))) \left(\frac{-\lambda_R r^*}{1-\delta} + \mathbb{E}[\epsilon \mid \epsilon \geq \bar{\epsilon}^*(r^*)] \right) + F\left(\bar{\epsilon}^*(r^*)\right)\left(-\lambda_R q + (1-\lambda_R)y + \delta(\bar{\epsilon}^*(\cdot), r^*)\right).$$

Now rearrange to get:

$$V_R(\bar{\epsilon}^*(\cdot), r) = \frac{(1 - F(\bar{\epsilon}^*(r))) \left(\frac{-\lambda_R r}{1-\delta} + \mathbb{E}[\epsilon \mid \epsilon \geq \bar{\epsilon}^*(r)] \right) + F\left(\bar{\epsilon}^*(r)\right)\left(-\lambda_R q + (1-\lambda_R)y\right)}{1-\delta F(\bar{\epsilon}^*(r))}. \quad (8)$$

Substituting for $V_R(\bar{\epsilon}^*(\cdot), r^*)$ from Equation 8, noting that we can write

$$\mathbb{E}[\epsilon \mid \epsilon \geq \bar{\epsilon}^*(r^*)] = \int_{\bar{\epsilon}^*(r^*)}^{\infty} \tilde{\epsilon} \frac{f(\tilde{\epsilon})}{1 - F(\bar{\epsilon}^*(r^*))} \, d\tilde{\epsilon},$$

and simplifying, $\overline{\epsilon}^*(r^*)$ is given by:

$$(1-\delta F(\overline{\epsilon}^*(r^*)))\overline{\epsilon}^*(r^*) - \delta \int_{\overline{\epsilon}^*(r^*)}^{\infty} \tilde{\epsilon} f(\tilde{\epsilon})\, d\tilde{\epsilon} = \lambda_R(r^* - q) + (1-\lambda_R)y. \tag{9}$$

Now the result follows immediately by implicitly differentiating Equation 9. ■

Proof of Proposition 2. Implicitly differentiating Equation 7, we have that at an interior solution:

$$\frac{dr^*}{dy} = -\frac{\left(\frac{f}{1-F}\right)'(\overline{\epsilon}^*(r^*))\frac{d\overline{\epsilon}^*(r^*)}{dy} + \frac{2\delta\lambda_C(1-\delta F(\overline{\epsilon}^*(r^*)))f(\overline{\epsilon}^*(r^*))\frac{d\overline{\epsilon}^*(r^*)}{dy}}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)} + \frac{\lambda_C^2(1-\delta F(\overline{\epsilon}^*(r^*)))^2}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)^2}}{\left(\frac{f}{1-F}\right)'(\overline{\epsilon}^*(r^*))\frac{d\overline{\epsilon}^*(r^*)}{dr} + \frac{2\delta\lambda_C(1-\delta F(\overline{\epsilon}^*(r^*)))f(\overline{\epsilon}^*(r^*))\frac{d\overline{\epsilon}^*(r^*)}{dr}}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)} + \frac{\lambda_C^2(1-\delta F(\overline{\epsilon}^*(r^*)))^2}{(1-\delta)\lambda_R(\lambda_C(r^*-q)+(1-\lambda_C)y)^2}} < 0, \tag{10}$$

where the inequality follows two facts. First, log-concavity of f implies that $\left(\frac{f}{1-F}\right)'(\overline{\epsilon}^*(r^*)) > 0$. Second, Lemma 4 shows that $\frac{\partial\overline{\epsilon}^*(r^*)}{\partial y} > 0$. ■

Proof of Proposition 3. The Ruler's ex ante expected welfare is:

$$V_R(\overline{\epsilon}^*(r^*), r^*) = \frac{(1-F(\overline{\epsilon}^*(r^*)))\left(\frac{-\lambda_R r^*}{1-\delta} + \mathbb{E}[\epsilon \mid \epsilon > \overline{\epsilon}^*(r^*)]\right) + F\left(\overline{\epsilon}^*(r^*)\right)(-\lambda_R q + (1-\lambda_R)y)}{1-\delta F(\overline{\epsilon}^*(r^*))}.$$

Differentiating, we have:

$$\begin{aligned}\frac{dV_R(\overline{\epsilon}^*(r^*), r^*)}{dy} &= \frac{1}{(1-\delta F(\overline{\epsilon}^*(r^*)))^2}\Bigg[\left(f(\overline{\epsilon}^*(r^*))\left(\frac{\partial\overline{\epsilon}^*(r^*)}{\partial y} + \frac{\partial\overline{\epsilon}^*(r^*)}{\partial r}\frac{\partial r^*}{\partial y}\right)(1-\delta F(\overline{\epsilon}^*(r^*)))\right)\\
&\times\left(\lambda_R\left(\frac{r^2}{1-\delta} - q\right) + (1-\lambda_R)y - \overline{\epsilon}^*(r^*) - (1-F\left(\overline{\epsilon}^*(r^*)\right))\frac{\lambda_R}{1-\delta}\frac{dr^*}{dy} + F(\overline{\epsilon}^*(r^*))(1-\lambda_R)\right)\\
&+\delta f(\overline{\epsilon}^*(r^*))\frac{d\overline{\epsilon}^*(r^*)}{dy}\left((1-F(\overline{\epsilon}^*(r^*)))\left(\frac{-\lambda_R r^*}{1-\delta} + \mathbb{E}[\epsilon \mid \epsilon > \overline{\epsilon}^*(r^*)]\right) + F\left(\overline{\epsilon}^*(r^*)\right)(-\lambda_R q + (1-\lambda_R)y)\right)\Bigg]\end{aligned}$$

From Equation 9, we can write:

$$\overline{\epsilon}^*(r^*) = \frac{\lambda_R(r^* - q) + (1-\lambda_R)y + \delta \int_{\overline{\epsilon}^*(r^*)}^{\infty} \epsilon f(\epsilon)\, d\epsilon}{1 - \delta F(\overline{\epsilon}^*(r^*))}.$$

Making this substitution and canceling like terms, the derivative reduces to:

$$\frac{dV_R(\overline{\epsilon}^*(r^*), r^*)}{dy} = \frac{F(\overline{\epsilon}^*(r^*))(1-\lambda_R) - (1 - F(\overline{\epsilon}^*(r^*)))\frac{\lambda_R}{1-\delta}\frac{dr^*}{dy}}{1 - \delta F(\overline{\epsilon}^*(r^*))}.$$

The derivative has the same sign as its numerator. The result now follows from the fact that, as shown in Proposition 2, $\frac{dr^*}{dy} \leq 0$.

The Church's ex ante expected welfare is:

$$V_C(\overline{\epsilon}^*(r^*), r^*) = \frac{(1 - F(\overline{\epsilon}^*(r^*)))\left(\frac{\lambda_C r^* + (1-\lambda_C)y}{1-\delta}\right) + F\left(\overline{\epsilon}^*(r^*)\right)\lambda_C q}{1 - \delta F(\overline{\epsilon}^*(r^*))}.$$

Differentiating, we have

$$\begin{aligned}
\frac{dV_C(\overline{\epsilon}^*(r^*), r^*)}{dy} &= \frac{1}{(1 - \delta F(\overline{\epsilon}^*(r^*)))^2}\Big[(1 - \delta F(\overline{\epsilon}^*(r^*))) \\
&\times \left(-f(\overline{\epsilon}^*(r^*)) \left(\frac{\partial \overline{\epsilon}^*(r^*)}{\partial y} + \frac{\partial \overline{\epsilon}^*(r^*)}{\partial r}\frac{\partial r^*}{\partial y} \right) \left(\lambda_C \left(\frac{r^*}{1-\delta} - q \right) + (1-\lambda_C)\frac{y}{1-\delta} \right) \right. \\
&\qquad \left. + (1 - F(\overline{\epsilon}^*(r^*))) \left(\frac{\lambda_C \frac{\partial r^*}{\partial y} + (1-\lambda_C)}{1-\delta} \right) \right) \\
+\delta f(\overline{\epsilon}^*(r^*)) &\left(\frac{\partial \overline{\epsilon}^*(r^*)}{\partial y} + \frac{\partial \overline{\epsilon}^*(r^*)}{\partial r}\frac{\partial r^*}{\partial y} \right) \left((1 - F(\overline{\epsilon}^*(r^*))) \left(\frac{\lambda_C r^* + (1-\lambda_C)y}{1-\delta} \right) + F\left(\overline{\epsilon}^*(r^*)\right)\lambda_C q \right) \Big]
\end{aligned}$$

This can be rewritten:

$$\begin{aligned}
\frac{dV_C(\overline{\epsilon}^*(r^*), r^*)}{dy} &= \frac{1}{(1 - \delta F(\overline{\epsilon}^*(r^*)))^2}\left[\frac{(1 - \delta F(\overline{\epsilon}^*(r^*)))\,(1 - F(\overline{\epsilon}^*(r^*)))(1-\lambda_C)}{1-\delta} \right. \\
&\qquad - f(\overline{\epsilon}^*(r^*))\frac{\partial \overline{\epsilon}^*(r^*)}{\partial y}\Big(\lambda_C(r^* - q) + (1-\lambda_C)y \Big) \\
&\left. + \frac{\partial r^*}{\partial y}\left((1 - \delta F(\overline{\epsilon}^*(r^*)))\,(1 - F(\overline{\epsilon}^*(r^*)))\frac{\partial \overline{\epsilon}^*(r^*)}{\partial r} - f(\overline{\epsilon}^*(r^*))\Big(\lambda_C(r^* - q) + (1-\lambda_C)y \Big) \right) \right].
\end{aligned}$$

This derivative has the same sign as the term in square brackets.

Now note that, by the Envelope Theorem, the term on the third-line is equal to zero. To see this, note that in the event that r^* is a corner solution, $\frac{\partial r^*}{\partial y} = 0$. In the event that r^* is interior, the first-order condition implies that the term in parentheses is 0.

Thus, the derivative has the same sign as:

$$\frac{(1-\delta F(\bar{\epsilon}^*(r^*)))\,(1-F(\bar{\epsilon}^*(r^*)))(1-\lambda_C)}{1-\delta} - f(\bar{\epsilon}^*(r^*))\frac{\partial \bar{\epsilon}^*(r^*)}{\partial y}\Big(\lambda_C(r^*-q)+(1-\lambda_C)y\Big).$$

Substituting for $\frac{\partial \bar{\epsilon}^*(r^*)}{\partial y}$, the derivative has the same sign as:

$$\frac{(1-\delta F(\bar{\epsilon}^*(r^*)))\,(1-F(\bar{\epsilon}^*(r^*)))(1-\lambda_C)}{1-\delta} - \frac{f(\bar{\epsilon}^*(r^*))\Big(\lambda_C(r^*-q)+(1-\lambda_C)y\Big)(1-\lambda_R)}{1-\delta F(\bar{\epsilon}^*(r^*))}. \tag{11}$$

Now we divide the analysis into several lemmas. First, focus on the case of an interior r^*.

Lemma 5 *For any $y \in (\underline{y}, \overline{y})$, the Church's welfare is strictly increasing in y if $\lambda_R > \lambda_Y$, strictly decreasing in y if $\lambda_R < \lambda_Y$ and constant in y if $\lambda_R = \lambda_Y$.*

Proof of Lemma 5.

From the first-order condition, we have that:

$$f(\bar{\epsilon}^*(r^*))\Big(\lambda_C(r^*-q)+(1-\lambda_C)y\Big) = \frac{(1-F(\bar{\epsilon}^*(r^*)))(1-\delta F(\bar{\epsilon}^*(r^*)))\lambda_C}{\lambda_R(1-\delta)}.$$

Substituting this into Equation 11, at an interior r^*, $\frac{dV_C(\bar{\epsilon}^*(r^*),r^*)}{dy}$ has the same sign as:

$$\frac{(1-\delta F(\bar{\epsilon}^*(r^*)))\,(1-F(\bar{\epsilon}^*(r^*)))(1-\lambda_C)}{1-\delta} - \frac{(1-\delta F(\bar{\epsilon}^*(r^*)))\,(1-F(\bar{\epsilon}^*(r^*)))(1-\lambda_R)\lambda_C}{(1-\delta)\lambda_R}.$$

Rearranging shows that at an interior r^*, $\frac{dV_C(\bar{\epsilon}^*(r^*),r^*)}{dy}$ has the same sign as:

$$\lambda_R - \lambda_C,$$

as required. ■

Now consider when r^* is a corner solution. We will establish the result in two steps. First, we show that, when r is fixed, the Church's welfare is decreasing in y if and only if y is sufficiently large.

Lemma 6 *Fix an r. Then there exists a $\hat{y}$ such that $V(\bar{\epsilon}^*(r), r)$ is strictly decreasing in y if $y > \hat{y}$ and strictly increasing in y if $y < \hat{y}$.*

Proof of Lemma 6. Rearranging Equation 11, $V(\bar{\epsilon}^*(r), r)$ is increasing if

$$\frac{1-\lambda_C}{(1-\delta)(1-\lambda_R)} \frac{(1-\delta F((\bar{\epsilon}^*(r)))\,(1-F((\bar{\epsilon}^*(r)))}{f((\bar{\epsilon}^*(r))} > \lambda_C(r-q) + (1-\lambda_C)y,$$

decreasing if the sign is reversed, and constant at equality. It is straightforward that the right-hand side is increasing and going to infinity in y. Hence, it suffices to show that the left-hand side is decreasing in y.

To see this, first note that log-concavity of f implies log-concavity of $1-F$. Thus, for any x, we have:

$$-f'(x)(1-F(x)) < f(x)^2. \tag{12}$$

Differentiating, the left-hand side is decreasing in y if:

$$\frac{\partial \bar{\epsilon}^*(r)}{\partial y}\left(\frac{-\delta f(\bar{\epsilon}^*(r))^2\,(1-F(\bar{\epsilon}^*(r))) - f(\bar{\epsilon}^*(r))^2\,(1-\delta F(\bar{\epsilon}^*(r))) - f'(\bar{\epsilon}^*(r))\,(1-F(\bar{\epsilon}^*(r)))\,(1-\delta F(\bar{\epsilon}^*(r)))}{f(\bar{\epsilon}^*(r))^2}\right) < 0$$

From Lemma 4, $\frac{\partial \bar{\epsilon}^*(r)}{\partial y} > 0$, so this inequality holds if and only if the fraction in parentheses is negative. Rearranging, this is equivalent to:

$$-f'(\bar{\epsilon}^*(r))\,(1-F(\bar{\epsilon}^*(r))) < f(\bar{\epsilon}^*(r))^2\left(1 + \frac{\delta\,(1-F(\bar{\epsilon}^*(r)))}{1-\delta F(\bar{\epsilon}^*(r))}\right),$$

which follows from Condition 12. ■

Finally, we show that the location of $\hat{y}$ is as in the statement of the proposition.

Lemma 7 *At any $y \in \{\underline{y}, \overline{y}\}$, the Church's welfare is strictly increasing in y if $\lambda_R > \lambda_Y$, strictly decreasing in y if $\lambda_R < \lambda_Y$, and constant in y if $\lambda_R = \lambda_Y$.*

Proof of Lemma 7. Rearranging Equation 11, for a fixed r, $\frac{dV(\overline{\epsilon}^*(r),r)}{dy}$ has the same sign as

$$\frac{1-\lambda_C}{(1-\delta)(1-\lambda_R)}\left(1-\delta F((\overline{\epsilon}^*(r))\right)\left(1-F((\overline{\epsilon}^*(r))\right)-f((\overline{\epsilon}^*(r))\left(\lambda_C(r-q)+(1-\lambda_C)y\right). \quad (13)$$

Using the definitions of $\overline{y}$ and $\underline{y}$, at either of these values, the second term of Condition 13 is equal to

$$\frac{\lambda_C}{\lambda_R(1-\delta)}(1-\delta F((\overline{\epsilon}^*(r)))(1-F((\overline{\epsilon}^*(r))),$$

where either $r=\underline{r}$ and $y=\overline{y}$, or $r=\overline{r}$ and $y=\underline{y}$. Substituting this into Condition 13, we have that at either of these values of y, $\frac{dV(\overline{\epsilon}^*(r),r)}{dy}$ has the same sign as

$$\frac{1-\lambda_C}{(1-\delta)(1-\lambda_R)}\left(1-\delta F((\overline{\epsilon}^*(r))\right)\left(1-F((\overline{\epsilon}^*(r))\right)-\frac{\lambda_C}{\lambda_R(1-\delta)}(1-\delta F((\overline{\epsilon}^*(r)))(1-F((\overline{\epsilon}^*(r))).$$

Rearranging one more time, $\frac{dV(\overline{\epsilon}^*(r),r)}{dy}$ has the same sign as

$$\lambda_R-\lambda_C,$$

as required. ■

From Lemma 6, when r is fixed, there is a $\hat{y}(r)$ such that $V(\overline{\epsilon}^*(r),r)$ is strictly increasing in y up to $\hat{y}(r)$ and then strictly decreasing. f From Lemma 7, if $\lambda_R > \lambda_Y$, then $V(\overline{\epsilon}^*(\overline{r}),\overline{r})$ is increasing at $y=\underline{y}$. Hence, $\hat{y}(\overline{r}) > \underline{y}$, so the Church's welfare is increasing for all $y \leq \underline{y}$. Moreover, by Lemma 5, $V(\overline{\epsilon}^*(r^*),r^*)$ is increasing for $y \in (\underline{y},\overline{y})$. Finally, by Lemma 7,

$V(\bar{\epsilon}^*(\underline{r}), \underline{r})$ is increasing at $y = \overline{y}$. Hence $\hat{y}(\underline{r})\overline{y}$. Thus, $V(\bar{\epsilon}^*(r^*), r^*)$ is strictly increasing in y up to $\hat{y}(\underline{r})) > \overline{y}$ and then strictly decreasing.

From Lemma 7, if $\lambda_R < \lambda_Y$, then $V(\bar{\epsilon}^*(\overline{r}), \overline{r})$ is strictly decreasing at $y = \underline{y}$. Hence, $\hat{y}(\overline{r}) < \underline{y}$. Moreover, by Lemma 5, $V(\bar{\epsilon}^*(r^*), r^*)$ is strictly decreasing for $y \in (\underline{y}, \overline{y})$. Finally, by Lemma 7, $V(\bar{\epsilon}^*(\underline{r}), \underline{r})$ is decreasing at $y = \overline{y}$. Thus, $V(\bar{\epsilon}^*(r^*), r^*)$ is increasing in y up to $\hat{y}(\overline{r}) < \underline{y}$ and then decreasing.

From Lemma 7, if $\lambda_R = \lambda_Y$, then $V(\bar{\epsilon}^*(\overline{r}), \overline{r})$ is constant at $y = \underline{y}$. Hence, $\hat{y}(\overline{r}) = \underline{y}$. Moreover, by Lemma 5, $V(\bar{\epsilon}^*(r^*), r^*)$ is constant in y, for $y \in (\underline{y}, \overline{y})$. Finally, by Lemma 7, $V(\bar{\epsilon}^*(\underline{r}), \underline{r})$ is constant at $y = \overline{y}$. Hence, $\hat{y}(\underline{r}) = \overline{y}$. Thus, $V(\bar{\epsilon}^*(r^*), r^*)$ is strictly increasing in y up to $\underline{y}$, constant for $y \in [\underline{y}, \overline{y}]$, and strictly decreasing for $y > \overline{y}$. ■

C Data Sources

Here we describe the data collection and sources. The replication data and do-file can be accessed on the authors' webpage.

C.1 Bishop types

To evaluate bishop alignment, we scraped web sites for each European diocese. The data regarding bishops can be found by searching Wikipedia for European Roman Catholic bishops. Such a search will lead to `https://en.wikipedia.org/wiki/Category:Roman_Catholic_bishops_in_Europe` from which one can then choose each country in turn and each bishop in turn. Bishops coded in black font have no biographical information. Those coded in blue do have biographical information. Those coded in red may have biographies forthcoming in the future. In addition to Wikipedia, we also scraped information on individual bishops from `http://www.catholic-hierarchy.org/bishop/` or, equivalently, the Catholic hierarchy site by country. In ambiguous cases additional websites relevant to the individual bishop were also searched although they rarely turned up information not already covered by Catholic Hierarchy or Wikipedia.

To create a preliminary coding of each bishop's type, biographical texts were scanned as follows:

A bishop was given a preliminary coding of Church-aligned if the biographical text included any of the following terms (with the appearance of multiple terms coded as well and with checks both for uppercase and lowercase entries):

> Archbishop, Benedictine, monk, Bishop, Bishop-elect, Cantor, deacon, Domscholaster, abbey, abbot, abbott, abott, arch-deacon, archdeacon, canon, cardinal, cathedral, champlain, chaplain, choirmaster, church, clergy, cleric, deacon, dean, elected, friar, hermit, cathedral, missionary, monastery, monk, monk/silversmith, papal, Pope, pope, preacher, prebend, prebendary, precenter, precentor, priest, priests, prior, proctor, rector, religious, sacrist, sub-dean, theologian, vicar, bishoprics, hermit.

A bishop was given a preliminary coding as lay-aligned if the biographical text included any of the following terms (with the appearance of multiple terms coded as well and with checks both for uppercase and lowercase entries):

> Governor, academic, ambassador, archchancellor, archduke, architect, artist, scholar, auditor, chancellor, chancery, coadjutor, diplomat, composer, count, diplomacy, doctor, duke, prince, exchequer, goldsmith, judge, government, keeper, king, kings, secretary, knight, law, writer, vice-chancellor, vice, lawyer, treasurer, privy, master, military, military/chancellor, noble, office, poet, politician, professor, advisor, council, councillor, justice, notary, official, physician, steward, scholar, secretary, secular, statesman, teacher, treasurer, prince-bishop, imperial.

After this preliminary coding, multiple coders hand read the individual text in cases for which the criteria yielded ambiguous or no coding. Coding was considered to be ambiguous if there was not a clear preponderance of one set of terms (either “lay” or “religious”) over the other. Prince-bishops, for example, were initially triggered as “lay” but if their biographical text as well triggered “religious” markers then their biographies were read and hand coded. In the case of prince-bishops this constituted the case for most of them. Ambiguities arise as well with certain terms or phrases that while important did not inherently indicate that the bishop was likely to be expected to be “lay” or “religious” oriented. For instance, while law could imply secular or religious alignment, phrases such as canon or canon law would indicate the individual was more likely in the religious domain whereas Roman law or just “law” would more likely have indicated a lay occupation. About 90 percent of the codings from the list of words were unambiguous (random checks were performed) so about 10 percent of the codings required close individual readings.

C.2 Trade Data

Trade data were downloaded from `http://www.ciolek.com/owtrad.html` with all routes designated as major and involving a European starting or ending point coded for the inclusive years specified on the website. The online appendix contains a spreadsheet with links to the routes extracted from the Old World TradeRoutes Project.

Additional trade route data both for continental Europe and for today’s United Kingdom were coded using the detailed trade route map found at `https://easyzoom.com/imageaccess/ec482e04c2b240d4969c14156bb6836f`. We also consulted the Wikipedia entries for each diocese in the United Kingdom for indications of medieval trading activity and their timespans.

D Additional Figures and Tables

Figure C.1 shows the number of bishops for whom we observe alignment for each half-century, for covered and uncovered dioceses.

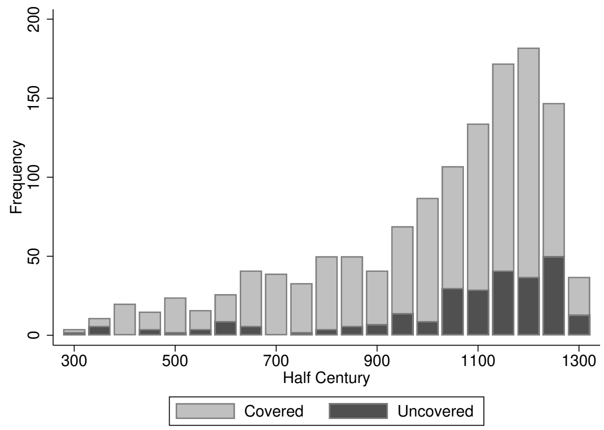

Figure C.1: Number of bishops for whom alignment is observed by half-century.

Table C.1 reports our standard regressions, but using missingness of bishop alignment data as the dependent variable. Columns correspond to different measures of wealth.

Table C.1: Correlates of missingness of bishop alignment data for various wealth measures.

	(1) Trade	(2) Caloric Potential	(3) Population 5,000	(4) Population 10,000	(5) Population 15,000
Concordats * Wealth	-0.01	-0.02	-0.03	-0.00	-0.06
	(0.04)	(0.04)	(0.04)	(0.07)	(0.10)
Concordats	-0.05*	-0.04	-0.04*	-0.05**	-0.05**
	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)
N	6895	6895	6895	6895	6895
half-century fixed effects	yes	yes	yes	yes	yes
diocese fixed effects	yes	yes	yes	yes	yes
crusades	yes	yes	yes	yes	yes
sample years	325-1309	325-1309	325-1309	325-1309	325-1309

Column labels indicate variable used to measure wealth. Standard errors clustered by diocese.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table C.2 shows that qualitatively similar results to Table 1 hold for the balanced panel.

Table C.3 shows the same regressions as for Table 1 clustering the standard errors by kingdom-by-half-century. Because the kingdom data is only available starting in 800, this changes the years and number of observations for all specifications.

Table C.2: Estimates of the difference in the effect of the Concordats on bishop alignment with the Church depending on wealth ($\hat{\beta}_2$) with a balanced panel of dioceses that exist for the entire 800–1309 time span. Rows correspond to different measures of wealth. Columns correspond to different fixed effects and controls. Entries are the coefficient estimate and standard error for the interaction of wealth and Concordats.

Wealth Measure	(1)	(2)	(3)	(4)	(5)
Trade	-0.08	-0.05	-0.12*	-0.09	-0.15*
	(0.06)	(0.09)	(0.07)	(0.06)	(0.08)
Caloric Potential	-0.11*	-0.19***	-0.15**	-0.01	-0.15*
	(0.06)	(0.07)	(0.06)	(0.09)	(0.08)
City 5,000	-0.03	-0.01	-0.11	-0.06	-0.06
	(0.06)	(0.09)	(0.07)	(0.06)	(0.08)
City 10,000	-0.09	0.03	-0.03	-0.13	-0.09
	(0.09)	(0.13)	(0.12)	(0.08)	(0.09)
City 15,000	-0.27***	0.23	-0.21	-0.18	-0.22***
	(0.09)	(0.14)	(0.15)	(0.16)	(0.08)
N	1182	644	898	1091	1182
half-century fixed effects	yes	yes	no	no	yes
diocese fixed effects	yes	yes	yes	yes	yes
crusades	yes	yes	yes	yes	yes
monarch fixed effects	no	yes	no	no	no
kingdom * half-century fixed effects	no	no	yes	no	no
grid * half-century fixed effects	no	no	no	yes	no
linear time trend * wealth	no	no	no	no	yes
sample years	800-1309	800-1309	800-1309	800-1309	800-1309

Standard errors clustered by diocese.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table C.3: Estimates of the difference in the effect of the Concordats on bishop alignment with the Church depending on wealth ($\hat{\beta}_2$), standard errors clustered by kingdom-half-century. Rows correspond to different measures of wealth. Columns correspond to different fixed effects and controls. Entries are the coefficient estimate and standard error for the interaction of wealth and Concordats.

Wealth Measure	(1)	(2)	(3)	(4)	(5)
Trade	-0.17***	-0.10	-0.16***	-0.21***	-0.27***
	(0.05)	(0.08)	(0.06)	(0.06)	(0.09)
Caloric Potential	-0.02	-0.21***	-0.19***	-0.15*	-0.02
	(0.08)	(0.06)	(0.06)	(0.08)	(0.12)
City 5,000	-0.12*	-0.09	-0.16***	-0.17**	-0.17*
	(0.06)	(0.08)	(0.06)	(0.08)	(0.10)
City 10,000	-0.13	0.00	-0.08	-0.14	-0.25
	(0.13)	(0.13)	(0.14)	(0.16)	(0.18)
City 15,000	-0.32	0.26*	-0.26	-0.24	-0.60*
	(0.22)	(0.13)	(0.23)	(0.23)	(0.31)
N	1241	853	1208	1172	1241
half-century fixed effects	yes	yes	no	no	yes
diocese fixed effects	yes	yes	yes	yes	yes
crusades	yes	yes	yes	yes	yes
monarch fixed effects	no	yes	no	no	no
kingdom * half-century fixed effects	no	no	yes	no	no
grid * half-century fixed effects	no	no	no	yes	no
linear time trend * wealth	no	no	no	no	yes
sample years	800-1309	800-1309	800-1309	800-1309	800-1309

Standard errors clustered by kingdom by half-century.
* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$